

NBER WORKING PAPER SERIES

A SECOND CHANCE AT SUCCESS? EFFECTS OF COLLEGE GRADE FORGIVENESS  
POLICIES ON STUDENT OUTCOMES

Xuan Jiang  
Kelly Chen  
Zeynep K. Hansen  
Scott Lowe

Working Paper 29493  
<http://www.nber.org/papers/w29493>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
November 2021

Authors do not have any funding support for this research. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Xuan Jiang, Kelly Chen, Zeynep K. Hansen, and Scott Lowe. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Second Chance at Success? Effects of College Grade Forgiveness Policies on Student Outcomes  
Xuan Jiang, Kelly Chen, Zeynep K. Hansen, and Scott Lowe  
NBER Working Paper No. 29493  
November 2021  
JEL No. I21,I23

**ABSTRACT**

The increased popularity of college Grade Forgiveness policies, which allow students to retake classes and substitute the new grades for the previous grades in their GPA calculations, is controversial yet understudied. Our paper is the first to ask whether such policies benefit students and how. To answer these questions, we use student-level admissions and transcript data from a four-year public institution in the U.S. that underwent two major changes in its GPA policy. We find that Grade Forgiveness significantly incentivizes students, especially students with the strongest academic preparation, to take STEM courses and challenging courses and to enroll in more credits. The increased variations in within-term grades suggest that students may change their effort allocations between courses taken in the same semester and spend more effort on courses that promise a higher grade in return. We also find that repeaters whose first attempted grades are forgiven are more likely to persist in the failed subject and obtain better grades subsequently. Finally, we see an increase in graduation in STEM majors for students who were intensively exposed to this policy.

Xuan Jiang  
Department of Economics  
The Ohio State University  
1945 N High St  
Columbus, OH 43210  
jiang.445@osu.edu

Zeynep K. Hansen  
Boise State University  
1910 University Drive  
Boise, ID 83725-1620  
and NBER  
zeynepahansen@boisestate.edu

Kelly Chen  
Department of Economics  
College of Business and Economics  
Boise State University  
1910 University Dr.  
Boise, ID 83725  
kellychen@boisestate.edu

Scott Lowe  
Graduate College  
Boise State University  
1910 University Drive  
Boise, ID 83725-1110  
scottlowe@boisestate.edu

# 1 Introduction

Should college academic policies provide more forgiving or more strict standards to promote success in higher education? One widely-adopted more forgiving policy is Grade Forgiveness. It allows college students to retake courses in which they received low grades and utilize only the most recent grade in the calculation of their overall grade point average (GPA). Although Grade Forgiveness has never been studied, whether college students should be offered a chance to repeat courses and replace low grades (or first missteps) is debatable. Opponents of Grade Forgiveness policies believe that these policies diminish the importance of grades, which could make students less inclined to put forth the level of effort needed to succeed in the course the first time. They also worry that it might be bad use of academic resources for students to retake courses to improve grades. Proponents point out that just because it takes some students longer to grasp a concept, does not mean that they are less capable and should be eliminated from the process. Many early-stage college students, underrepresented minorities, and first-generation students struggle to thrive in college, both academically and socially, and this policy improves their chances of success.<sup>1</sup> At a theoretical level, it is unclear whether this policy leads students to reduce effort or if it provides protection for students to challenge themselves.

In this paper, we comprehensively evaluate Grade Forgiveness policies' impact on students choices and performance in current and subsequent courses, and graduation. "Grade Forgiveness" policies directly change the GPA calculation by increasing the expected benefit of repeating a course without changing the cost of repeating a course. Thus, we expect that, via increasing the probability of repeating, the policy will provide students with incentives to change their choices and behaviors across many dimensions. The research questions we ask are threefold.

---

<sup>1</sup>For example, see discussion in news articles from Atlantic: <https://www.theatlantic.com/education/archive/2018/06/college-grades-gpa/564095/> and <https://www.miamistudent.net/article/2012/04/mu-implements-grade-forgiveness>.

First, how does the policy impact students' curriculum choices, including the number of courses or credits that they enroll in, the difficulty of the courses that they take, and the time that they allocate across the courses that they take in the same semester? We are mostly interested in whether the policy incentivizes students to challenge themselves by taking harder courses or more credits, and whether students re-allocate time spent across courses taken in the same semester. Second, how does the policy affect repeaters' subsequent interest and performance in their repeated course? Finally, to what extent does the policy promote students' success or persistence in obtaining a degree?

We first present a simple conceptual framework to illustrate the changes in students' decision making in response to the change from a grade averaging policy to a "Grade Forgiveness" policy.<sup>2</sup> We then provide empirical evidence on Grade Forgiveness policy's impacts on students' outcomes using a rich dataset with student-level admission information and student-course-level transcript information from Boise State University (hereafter "the University"), the largest public four-year institution in the state of Idaho. A marked advantage of investigating Grade Forgiveness by using the University's data is that the University experienced a Grade Forgiveness policy implementation and reversal within recent years. Over our study period, the University had a "Grade Forgiveness" policy in place before 1990 and replaced it with an averaging policy in the fall of 1995. Within the averaging policy era, the repeated grade and the first attempt grade were averaged when calculating the student's cumulative GPA. The University re-implemented the "Grade Forgiveness" policy in the fall of 2001.<sup>3</sup>

We adopt a fixed effect model strengthened by the on-off-on policy variations to empirically identify the policy's intent-to-treat effects. We begin by showing that students are significantly more likely to repeat courses when the policy is in place. Specifically, the policy increases the likelihood of repeating by up to 2.23 percent-

---

<sup>2</sup>A usual counter GPA policy to the "Grade Forgiveness" policy is grade averaging policy, which counts both the repeated grade and the first-attempted grade by averaging them in the cumulative GPA calculation.

<sup>3</sup>The policy is not retroactive and only applies to the courses repeated during each policy period.

age points, which translates into a 64% increase from the baseline (the time period when grade-averaging policy is in place, between fall 1995 and fall 2001) repeating rate of 3.5%. This provides evidence on the mechanism of changes in students' curriculum choices. Contradicting opponents of Grade Forgiveness policies, we find that the policy significantly nudges students to challenge themselves: when grades can be forgiven, students tend to enroll in more credits, to enroll in more challenging courses, and to enroll in courses that have more stringent grading policies. For example, students are 9.5% more likely to take a Science, Technology, Engineering, and Mathematics (STEM) course. More importantly, for students who have not yet declared a major, the policy increases their likelihood of enrolling in a STEM course by 10.6%. We have also explored the policy's gender-specific effects and find that the policy has a significantly weaker effect on women in terms of course repetition and enrollment in STEM courses, which suggests that the perceived cost associated with course repetition is higher for women than for men.

Further, we investigate the policy's effects on the subsequent course choice and performance of the repeaters—students who have repeated at least one course. We are particularly interested in repeaters' likelihood of taking another course *in the same subject as the repeated course*. We do so by using an instrumental variable approach, where the probability of repeating a course is instrumented by the presence of the policy. We find that repeaters are 22 percentage points more likely to take another same-subject course in a subsequent semester after repeating. Moreover, repeaters who take another same-subject subsequent course are 30 percentage points less likely to get a failing grade (D or lower) under the policy, implying that the policy has a positive effect on encouraging on-the-margin students to persist and progress in challenging subjects.

By investigating students' performance, we find that grades within a term are more dispersed, echoing the theoretical insights: students re-allocate their time among courses in response to the policy in a way that favors the more promising courses. All the impacts above contribute to student's relatively long-term out-

comes indefinitely. We thus look at graduation (measured at the per-student-level) by comparing those who were more intensely treated by the policy in the early stage of their college career versus the rest. We find a null effect of the “Grade Forgiveness” policy on the probability of ever graduating from the University for non-transfer (first-time) students. We also find that the policy has a null effect on the on-time (6-year) graduation rate at the University. Among the sub-sample of students who graduated from the University, we see a positive and significant policy effect on obtaining a STEM degree. This implies that, overall, the policy nudges students to obtain challenging degrees without a significant “price,” i.e., without having them take longer to graduate.

To the best of our knowledge, we are the first to study Grade Forgiveness policies. Indeed, more forgiving or inclusive academic policies are under-studied, relative to more strict policies, such as remedial education (e.g., [Bettinger and Long \(2009\)](#)) and grade retention policies (e.g., [Tafreschi and Thiemann \(2016\)](#)). Unlike the mandatory remedial education and retention policies, which raise standards for the least prepared students, Grade Forgiveness policies act as an insurance by providing a chance for all students to voluntarily retake courses to improve grades. In a similar vein, there is a relatively larger, although still small, literature on retaking high-stakes tests, including the SAT, ACT, and college entrance exams. Retaking high-stakes tests substantially improves scores and increases four-year college enrollment rates, particularly for low-income and underrepresented minority students ([Goodman, Gurantz, and Smith, 2020](#); [Vigdor and Clotfelter, 2003](#)). Retaking high-stakes tests improves grades both through increased familiarity with the test and through actual learning ([Frisancho, Krishna, Lychagin, and Yavas, 2016](#)). While both are inclusive policies in higher education, the essential difference between retaking college courses and retaking college entrance exams is that the former impacts not only the repeated course but also students’ entire curriculum, performance and subsequent progress. Our paper, on the one hand, echoes the studies cited above on the positive consequences of retaking a course; on the other hand,

it provides the first evidence on changes in choices and behaviors of all students, including repeaters and non-repeaters, in response to Grade Forgiveness policy.

The findings in our paper are timely and policy-relevant as more and more universities are adopting this type of GPA policy as well as seeking new ways to mitigate the academic struggles of college students. Whether college students should be given a second chance in coursework is a fundamental topic in institutional research, as is whether bad grades should be forgiven. Understanding these outcomes is critical to having better insights on students' achievement, including their progress at every stage of college and final job outcomes. Grades are important and powerful signals for students' later success ([Bandiera, Larcinese, and Rasul, 2015](#); [Goodman, 2016](#); [Marx and Meeler, 2013](#)), as well as for students on the job market as employers frequently require candidates to submit their academic transcripts and many competitive positions require a minimum GPA ([Reshwan, 2016](#); [Tafreschi and Thiemann, 2016](#)). Although the literature has attempted to address how to promote students' success through grade-related policies, most existing studies have focused on different *grading policies'* effects on college students' success ([Minaya, 2020](#); [Ahn, Arcidiacono, Hopson, and Thomas, 2019](#)). However, modifying course grading policies at the institutional level is extremely difficult to justify, and even more difficult to implement. Professors and instructors may find it hard to accommodate a lenient grading policy without fundamentally changing the course material and the study goals for students with different backgrounds. The "Grade Forgiveness" policy we study, however, does not affect any particular course grading policy. It simply provides an option for students to fix their "failure" without requesting any changes to individual courses and may also make it easier for instructors to be "tough".

The existing literature on "Grade Forgiveness" policies has also documented that self-learning about ability through grades plays a very prominent role in college dropout decisions. Students that drop out between their first and second years would be largely reduced if no self-learning occurred about grade performance or

academic ability (Stinebrickner and Stinebrickner, 2012, 2014). Students may revise their beliefs about their own abilities in response to (low) grades, leading some students to leave difficult majors (Ahn et al., 2019; Astorne-Figari and Speer, 2019; Stinebrickner and Stinebrickner, 2013). The option to repeat courses, which most of the existing studies do not consider, would likely affect college dropout and major-switching decisions. We contribute to the literature by highlighting a key effect of this widely adopted but overlooked policy—nudging students to challenge themselves and promoting them to pursue and persist in difficult majors.

## 2 The Policy Background

“Grade Forgiveness” policies allow students to retake any courses in which they received low grades and utilize the most recent grade to calculate their overall GPAs. Recently, more and more universities have adopted this kind of policy, with varying requirements and restrictions. For example, the Ohio State University adopted a similar “Grade Forgiveness” policy in fall 2015, which allows students to petition to complete a second-course attempt and to replace the original grade<sup>4</sup>. In 2018, the Massachusetts Institute of Technology initiated an experimental freshman grading policy, which allows up to three science core General Institute Requirements (GIRs) to be graded on a Pass/No Record (P/NR) basis as part of their Committee on the Undergraduate Program (CUP) Experimental Grading Policy<sup>5</sup>. More universities are starting to implement similar policies—some have applied the grade forgiveness policies to the entire student body, some to only first-year students, some have restrictions on certain grades to be repeated (e.g., letter grades below D), and some have restrictions on the maximum number of courses that students can repeat.

Most land grant universities have adopted a similar Grade Forgiveness policy. The practice of this policy may vary across institutions regarding limits to the number of courses that are allowed to be repeated and whether the registration of the

---

<sup>4</sup><https://math.osu.edu/undergrad/non-majors/scheduling/repeating/forgiveness>

<sup>5</sup><https://registrar.mit.edu/classes-grades-evaluations/grades/grading-policies/experimentalgrading-policy/entering-fall-2018>



repeated course requires approval from advisors. The increased popularity of college “Grade Forgiveness” is often criticized by the media saying that colleges are trying to ensure their “customers” are satisfied. Although there is no survey or study on this policy’s influence on students to date, the popular press has captured different opinions on this practice from students, both pros and cons.<sup>6</sup>

## 2.1 Boise State University’s Policy

Boise State University (BSU) is a four-year public university located in the northwest United States, with an undergraduate population of approximately 22,000. During the observation period, BSU had the largest undergraduate enrollment in the State of Idaho and offers nearly 80 bachelor’s degrees across seven academic colleges: Arts & Sciences, Business & Economics, Education, Engineering, Health Sciences, Innovation & Design, and the School of Public Service.

BSU has a long history with “Grade Forgiveness” policies, referred to as “grade replacement” policies by the University. A grade replacement policy was implemented prior to 1970, which allows students who receive grades below “D” to retake a course to improve a grade and use the most recent grades for GPA calculation. Starting Fall 1988, students can retake a course to improve *any grades* and use the most recent grades for GPA calculation. This grade replacement policy was replaced with a grade averaging policy in Fall 1995 and then reverted back to the grade replacement policy in Fall 2001. This policy applies to all students in the University. During the averaging policy era, the only difference in course repeating, compared to the “Grade Forgiveness” policy eras, is that the repeated grade and the first-attempted grade were averaged when calculating the student’s cumulative GPA, which is typical at schools without a Grade Forgiveness policy.<sup>7</sup> The Fall 2001

---

<sup>6</sup>“I feel like it might be a bad thing. It would give kids an opportunity to feel like they don’t have to work as hard the first time around”—<https://www.thelantern.com/2015/08/freshman-forgiveness-now-undergrad-amnesty/>. “It will allow students who are new to the rigor of a college education a period of adjustment. This policy will serve as a safety net for students to recover from a course that they might not have been prepared for”—<https://www.miamistudent.net/article/2012/04/mu-implements-grade-forgiveness>.

<sup>7</sup>See [current official webpage describing policy changes here](#):

“grade forgiveness policy” allowed students to repeat any course and to replace the old grade with the most recent one.<sup>8</sup> Only the most recent grade is used in the calculation of the cumulative GPA. The university’s official statement on the policy is:

*“Courses repeated prior to Fall 1995 use a grade replacement policy. Only the most recent grade was used in calculating the cumulative GPA.”*

*“Courses repeated Fall 1995 through Summer 2001 used a grade averaging policy. Courses repeated will be averaged, using both grades in the calculation of the GPA.”*

*“Beginning Fall 2001 and on, courses repeated will use a grade replacement policy. Only the most recent grade will be used in the calculation of the cumulative GPA.”*

The University passed several policies in Fall 1995 in order to raise academic standards. Those policies restricted the possibility of improving grades by course repeating, including substituting the grade replacement policy with a grade averaging policy, and limiting the maximum number of times that a student can enroll in the same course. It should be noted that the University was served by different sets of Academic Standards Committees during the 5 years of policy changes. Starting Fall 2001, the University changed the course repeat policy back to a grade replacement policy.

According to the University’s Academic Standards Committee, they reverted back to a grade replacement policy for two primary reasons. First, it came to the faculty’s attention that the grade averaging policy treated the first and the second attempts as two independent grades for GPA calculations, rather than averaging the two grade points, which appeared to “penalize students to a greater extent than was first proposed” and “it [was] more difficult for students to raise their GPA”.<sup>9</sup>

<https://www.boisestate.edu/policy/academic-affairs-student/policy-name-course-repetition-gpa-relationship/>

<sup>8</sup>See Boise State University’s 2001-2002 undergraduate catalog describing the repeat policy on pages 23-24 on this webpage: <https://scholarworks.boisestate.edu/catalogs/67/>.

<sup>9</sup>For example, a student has three entries on their transcript: ECON 101 (grade B); ECON 102 (first time, grade D); ECON 102 (second time, grade C). Both courses offered 3 credits. Based on the

Second, most other colleges and universities in the state of Idaho implemented a grade replacement policy at that time, and “this has proved to be unfair to incoming transfer students” because those students took courses at their original institutions in good faith under the grade replacement rules.<sup>10</sup>

Students were well-informed of the grade replacement policy changes by academic advising, and were made aware of the policy changes in a timely manner. In addition, grade policies were published in each year’s undergraduate catalog, and in the Student Newspaper, the “Arbiter”. In the Arbiter, January 18, 1995, an article titled “New grade rules will greet students next fall” informed students of the academic policy changes that would take place in Fall 1995. In the Arbiter, August 30, 2001, an article titled “Grade replacement policy takes effect this semester” mentioned the policy change as “Students have a new tool this semester to improve their all-important grade point averages.”<sup>11</sup>

Figure 1 shows the average course repeat rate across calendar years. There is a clear drop in the course repeat rate around Fall 1995 when the University turned off the “Grade Forgiveness” policy. The frequency stayed at a low level until Fall 2001 when the University switched the grade forgiveness policy back on. The direct effect of tightening the course repeat policy is a decline in the average cumulative GPA. Figure 2 reveals that the undergraduate cumulative GPA gradually declined after Fall 1995 and then increased, responding to the policy change in Fall 2001.

The monetary cost of repeating a credit hour varies across time. For full-time students, who register for between 8 and 19 credits during Fall 1990–Summer 2008 or register for between 12 and 17 credits during Fall 2008–Summer 2016, the tuition is a flat rate. Thus for this group of student, repeating a course is free. For part-time students or students who register for more than the upper limit of the full-time credit hours (i.e., “overload”), the “sticker price” for each credit hour varies

---

grade averaging policy in force, the GPA will be  $\frac{3 \times 3 + 3 \times 1 + 3 \times 2}{3 + 3 + 3} = 2$ , rather than  $\frac{3 \times 3 + 3 \times \frac{1+2}{2}}{3+3} = 2.25$ .

<sup>10</sup>The decisions of the policy changes were mentioned in the files of the Academic Standards Committee.

<sup>11</sup>See [https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1985&context=student\\_newspapers](https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1985&context=student_newspapers) and [https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2194&context=student\\_newspapers](https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2194&context=student_newspapers).

from \$61.75 to \$297 during 1990–2016 (same for in-state and out-of-state students).<sup>12</sup> About 61% of BSU students are enrolled as full-time students, among which about 6% full-time students are “overloaded”<sup>13</sup>. Among the students who have repeated at least one course, about 80% repeated one course which typically accounts 3 credits. Thus, for a typical part-time student, the sticker price for repeating a typical course is around \$180–\$900.

### 3 Conceptual Framework

Conceptually, the “Grade Forgiveness” policy offers students costly insurance against a low grade. The policy increases the expected final grade but does not change the cost of repeating, when repeat happens, by changing the calculation of GPA relative to the alternative GPA policy—a GPA averaging policy. The cost of repeating consists of a time cost and possible additional mental cost. If the insurance value of using the “Grade Forgiveness” policy outweighs its cost, we would see it incentivizes students to take risks. Therefore, the policy has two sets of potential effects that are both interesting and important to be explored. First, the policy has an indirect effect on all students, who may or may not choose to retake a course, which captures the “indirect” effect of the policy on a variety of outcomes, through both choices and performance in courses. This effect functions through the possibility of repeating a course. Second, the policy has a direct or “intended” effect on students, who choose to retake a course. This set of effects, including the subsequent outcomes of the students who repeated courses, indicates whether the policy benefit its “targeted” group.

To better understand the “indirect” effect of the “Grade Forgiveness” policy, we lay out a simple theoretical framework to illustrate a series reactions of students to this policy, relative to the counter GPA policy—grade averaging policy. This frame-

---

<sup>12</sup>As a public institution, BSU also receives state funds to help subsidize the price of a degree for each student, based on eligibility. The sticker price should be considered as an upper bound of the tuition.

<sup>13</sup>“Overloaded” students are often high-performing and attempting multiple majors/minors.

work also provides insights for the mechanism of the policy's empirical evidence on course choice, time allocation, and performance.

We consider the environment in which a student makes choices on course taking and time allocation to maximize their utility, where both the difficulty level of courses and the expected grades determine utility. The intuition of this assumption is that students should value grades as well as the subject or material they study. Studies have shown that subjects gearing towards high-paying jobs are often harsh, associated with tough grading policies.

A student chooses course difficulty,  $d$ , and time spent on the course,  $t$ , which together define the grading policy as  $g(d, t)$ . The cost of taking a course is a function of course difficulty and time,  $c(d, t)$ . When the student does not repeat a course, the utility is simply determined by the difficulty level, the grading policy, and cost function:

$$U^{(\text{not repeat})} = d * g(d, t) - c(d, t)$$

When the student repeats a course after taking the course for the first time (thus revealed the first-attempt grade), the utility of repeating is defined differently under each policy:

$$U^{(\text{repeat})} = \begin{cases} d * \frac{E[g'] + g(d, t)}{2} - c(d, t) - E[c'], & \text{if Forgiveness} = 0 \\ d * E[g'] - c(d, t) - E[c'], & \text{if Forgiveness} = 1 \end{cases}$$

where  $E[g']$  is the expected grades of repeating and  $E[c']$  is the expected cost of the repeating. We assume that, under the Grade Forgiveness policy, when students can replace the repeated grade with the original grade, students do not care about the original grade. We can then write the utility function as the summation of the utility function of repeat and the utility function of not repeat integrated over a function of the first-attempt grade,  $f(g)$ . Intuitively, the student will repeat when the first-attempt grade below a cut-off grade,  $g^*$ , and will not repeat when the first-

attempt grade above the cut-off and below the maximum grade (denoted as "A" here).

$$U(g, c, d, t) = \int_0^{g^*} U^{(\text{repeat})} f(g) dg + \int_{g^*}^A U^{(\text{not repeat})} f(g) dg$$

Assume  $g$  and  $c$  are both twice continuously differentiable and  $g'(d) < 0, g''(d) < 0, g'(t) > 0, g''(t) < 0, c'(d) > 0, c''(d) < 0$ . The probability of repeating a course,  $f(g)$ , is decreasing in  $g$ :  $f'(g) < 0$ . To fix ideas, the level of difficulty the student chooses for a semester is affected by both the level of difficulty of each course and the number of courses enrolled. Here, we do not distinguish the difficulty level of a single course (quality) and the number of courses enrolled (quantity).

### 3.1 Difficulty and Time Choice

We derive the following propositions, in terms of the choice on difficulty and time allocation, for students at the stage of choosing courses and before revealing first-attempt grade. Proofs are in Appendix A.

**Proposition 1.** Students will choose higher level of difficulty under the "Grade Forgiveness" policy (i.e., Forgiveness = 1), relative to the averaging policy (i.e., Forgiveness = 0).

**Proposition 2.** The time allocated among different courses is more dispersed under the "Grade Forgiveness" policy (i.e., Forgiveness = 1), relative to the averaging policy (i.e., Forgiveness = 0). When having the option to repeat any course to improve grades later, students will allocate time in favor of the course(s) that are more likely to obtain higher grades and allocate time away from the course(s) that are more likely to obtain lower grades in the current semester. We should expect to see a larger variation in a student's grades within a term under the Grade Forgiveness policy relative to the grade averaging policy.

## 3.2 Probability of Repeating and Threshold Grade

We then assume that a student has made the course enrollment decision, completed a course, and revealed the first-attempt grade,  $g$ . We derive the probability for the student to repeat this course under the two different GPA policies.

**Proposition 3.1.** The average probability of repeating under Forgiveness = 1 will be higher than the probability of repeating under Forgiveness = 0.

**Proposition 3.2.** The threshold (highest) grade to repeat under Forgiveness = 1 is higher than the threshold (highest) grade to repeat under Forgiveness = 0:  $g_1 > g_0$ , and the difference between the two threshold grades is restricted as  $g_1 - g_0 \leq E[c']$ .

## 3.3 Persistence

Assume that students' dropout decision, either dropping out of college or dropping out of (switching) a major, is made depending on the cumulative GPA. A student who repeats course A in period  $t - 1$ , and makes a decision on whether to keep taking courses in period  $t$  based on the cumulative GPA,  $GPA_{t-1}$ , expected grades,  $E_t[g']$  and expected cost,  $E_t[c']$ , in the next period. The student chooses to keep taking course in the subject/major as long as the expected utility is above the threshold  $u^*$ .

$$D_t^{Persist} = \mathbb{1} [U(GPA_{t-1}, E_t[g'], E_t[c'] \mid Forgiveness = 0, 1) > u^*]$$

**Proposition 4.** A student is more likely to persist in an enrolled subject or major under the forgiveness policy, relative to the averaging policy.

## 4 Data

We use administrative transcript data from undergraduate students at the University, who entered the university between the spring 1990 and spring 2017 semesters. The raw data include approximately 170,812 students and provides admissions in-

formation, including residency, SAT/ACT scores, and demographic characteristics. The student-section-level transcript data provide all courses enrolled in, credits attempted, credits earned, grades obtained, and information on the courses that the student has repeated.

Table 1 provides summary statistics on characteristics of interest. Of the 170,812 students in the full sample, 53.6% are female. Based on 2020–2021 figures, BSU has 73% White, 13% Hispanic or Latino, 5% Two or More Races, 3% Asian, 2% Black or African American, 0.397% Native Hawaiian or Other Pacific Islanders, and 0.389% American Indian or Alaska Native. 66% of students come from the state of Idaho, and 1% are international students.<sup>14</sup> About 45% students at BSU are transfer students. We exclude students that transferred-in coursework in some of our analysis because we do not observe their full transcripts, and courses taken from other institutions may not be comparable with those offered by BSU.

Figure 1 shows the average course repeat rate across calendar semesters. There is a clear drop in the course repeat rate around Fall 1995, when the University turned off the “Grade Forgiveness” policy, and this rate stays at a low level until Fall 2001, when the University switched the policy back on. The average course repeat rate is 0.035 during the policy-off period and 0.046 during the policy-on period.

## 5 Empirical Strategy

We aim to investigate two sets of causal effects of the “Grade Forgiveness” policy on student outcomes. The first is the total effect, which estimates the policy’s effect on all students who may or may not choose to retake a course. For the students who do not repeat, we are estimating the indirect effects that we have shown via the theoretical implications. For the students who do repeat, we are estimating both the indirect effects and the direct effects from repeating courses. This set of outcome variables of interest includes the student’s choice on course enrollment—

---

<sup>14</sup><https://www.boisestate.edu/about/facts/>



measured by the type of courses and the number of credits students choose to enroll in, choice on time allocation—proxied by grade distribution within a semester, and ultimately their performance. As illustrated in the theoretical framework, students may change these choices and behaviors when they have a chance to replace bad grades. This essentially captures the “indirect” effect of the “Grade Forgiveness” policy.

The second effect we are interested in estimating is the local average treatment effect (LATE), which estimates the policy’s effect on students who retook a course *in response to the “Grade Forgiveness” policy*. This group of students would not have retaken a course without the policy in place, in contrast to high-performing students who rarely repeat and low-performing students who repeat with or without the policy in place. We are mostly interested in students’ subsequent choices and performances after repeating. The hypothesis is that, with the previous (and presumably worse) grade forgiven, the repeaters have a larger chance to make progress in their studies; what’s more, by repeating a course, they might regain confidence in the subject by obtaining knowledge or at least familiarity with the subject.

## 5.1 Total Effects

Given the richness of our individual-section level data, we explore a series of outcomes affected by the policy using a fixed effect model with two variations of the treatment over time. The estimation can be written as follows:

$$Y_{ijat} = \beta_0 + \beta Policy_t + \gamma_i + \delta_a + \theta_j + \Lambda S'_{jt} + \varepsilon_{ijat} \quad (1)$$

where  $i$  denotes students,  $j$  denotes courses, and  $t$  denotes the calendar semester.  $a$  is the academic progress/semester for each student, representing the number of semesters elapsed since a student’s initial enrollment at the time of observation (i.e., 1st, 2nd, 3rd,..., semester).  $\gamma_i$  is a vector of individual fixed effects, and  $\delta_a$  is a vector of academic progress fixed effects. The key independent variable,  $Policy_t$ , is an indicator of the policy’s presence, which equals 1 for any course taken by any stu-

dent during the Spring 1990 (start of our data)–Summer 1995 or Fall 2001–Fall 2019 (end of our data) periods, and 0 for any course taken by any student during the Fall 1995–Summer 2001 period. It is essential to distinguish the *academic semester for each student*,  $a$ , and the *calendar semester*,  $t$ . The policy varies by the calendar semester while the treatment groups considered varies for each student at their academic progress,  $a$ . The individual fixed effects,  $\gamma_i$ , and academic progress fixed effects,  $\delta_a$ , could be considered facilitating a conventional difference-in-differences approach. This approach identifies, among all partially-treated individuals, the difference between students who entered the university at different times in the differences between outcomes from any two academic progress  $s$ ,  $a'$  and  $a$ . Assuming parallel trends in a typical difference-in-differences design means that when we compare two students, A, who has not been treated in academic semesters  $a$  and  $a'$  while B has been treated in  $a'$  but not  $a$ , the difference between the outcome in  $a$  and  $a'$  of student B would have been identical to the difference between the outcome in  $a$  and  $a'$  of student A, if student B had not been treated.

It is worth noting that our approach is not an analogue to the staggered difference-in-differences. The criticism that recent work in econometric theory has on the staggered DiD designs, where groups adopt the policy or treatment of interest at a particular point in time and the key to causal interpretation of the estimates requires both a parallel trends assumption and treatment effects that are constant over time does not apply here.

We also include course level controls.  $\theta_j$  is a vector of course fixed effects, capturing any course-specific characteristics. We include  $\theta_j$  in all of our specifications when the dependent variable is at individual-section level but not course-level.  $S_{jt}$  is a matrix of course-section characteristics, including mean GPAs and enrollments, further capturing within-course between-sections variations.  $\beta$  is the coefficient of interest here.

There are two primary advantages of taking this estimation approach. First, the student-level individual fixed effects control for any time-invariant hidden charac-

teristics associated with being in a particular cohort (and hence exposure to the policy) and account for the observed differences in student outcomes, such as ability, academic preparation, and family background. This eliminates concerns regarding selections over any individual-level unobserved heterogeneity. The second advantage is that we have the policy in place for both the old cohort and the most recent cohort, while we have the policy off for the middle cohort. This feature allows us to avoid potential confounding effects from contemporaneous institutional changes over time.

## 5.2 Effects on the Students who Repeated

Next we study how the policy affects repeaters' subsequent interest and performance in taking another course in the same subject as the repeated course. We hypothesize that repeaters might accumulate knowledge and regain interest in the subject by real learning when retake the course. If repeating improves interest or persistence in the subject, we should expect to observe that repeaters take more subsequent courses in the failed and repeated subject. If repeating promote real learning on the material more than merely familiarity on the exam, we should expect to see the repeaters perform better in the subsequent same-subject courses. To do this, we estimate the local average treatment effects under the assumption that the policy effects on the subsequent outcomes of repeaters are only through repeating and not through any indirect effects. We adopt an Instrumental Variable (IV) approach, where the probability of repeating a course is instrumented by implementing this policy. Our assumption that the policy effects on the subsequent outcomes of repeaters are not through any indirect effects is strict theoretically but reasonable empirically. First, our two main outcome variables of interest here, the number of credits subsequently taken in the same subject and the grade of the same-subject subsequent course, are not directly associated with the difficulty measures of courses in section 6.2. Second, the policy indirect effects on curriculum choices are driven by high-performing students while the repeating is driven

by low-performing students. We substantially expanded this analysis on repeaters subsequent choices and performance by focusing on low-performing students in our companion paper.<sup>15</sup>

$$\begin{aligned} \text{Subsequent}_{ij'a't'} &= \beta_0 + \beta \widehat{\text{Repeat}}_{ij'at} + \gamma_i + \delta_a + \theta_j + \varepsilon_{ij'at} \\ t' &> t \end{aligned} \tag{2}$$

$\text{Repeat}_{ij'at} = 1$  if the current course is repeated and  $\text{Repeat}_{ij'at} = 0$ , otherwise. The main outcome variable here,  $\text{Subsequent}_{ij'a't'}$  is an indicator of whether a student takes another course in a certain subject after taking the current course.  $\text{Subsequent}_{ij'a't'} = 1$ , if this student takes any course,  $j'$ , which is in the same subject as the current course  $j$  in any of their subsequent semesters,  $a' > a$ . We also look at the performance— $\text{Grade}_{ij'a't'}$ —of a same-subject subsequent course, when  $\text{Subsequent}_{ij'a't'} = 1$ . For example, student A repeated French 102 in their second semester but never took any other courses in French. Their  $\text{Interest}_{ij'a't'} = 0$  and their performance measure is missing. Student B repeated Math 108 in their third semester, then she took Math 202 in the 5th semester and got grade A, the  $\text{Interest}_{ij'a't'} = 1$  and performance,  $\text{Grade}_{ij'a't'} = A$ .  $\widehat{\text{Repeat}}_{ij'at}$  is the predicted likelihood of repeating.  $\beta$  is the coefficient of interest here, representing the treatment effect on the treated group. The first stage is straightforward and identical to what we show above in Equation (1):

$$\text{Repeat}_{ij'at} = \lambda_0 + \lambda \text{Policy}_t + \gamma_i + \delta_a + \theta_j + \varepsilon_{ij'at} \tag{3}$$

---

<sup>15</sup>See [Chen and Jiang \(2021\)](#), “Pushing the Pause Button: Can Course Repetition Foster Student Motivation and Future Success?”.

## 6 Findings

### 6.1 "First Stage" Response: Do Students Repeat More?

#### 6.1.1 Probability of Repeating

We begin by analyzing students' responses to the policy in terms of course repeating behaviors. Presumably, an increase in repeating behavior is the most direct response to the "Grade Forgiveness" policy, if students are responsive to this policy. The outcome variable of interest here is an indicator of whether a course is repeated. Table 2 shows the "Grade Forgiveness" policy's effect on the probability of repeating, where the outcome variable is an indicator of a course being a repeat:  $Y_{ijat} = 1$  indicates the course  $j$  taken by student  $i$  in academic progress  $a$  and in calendar semester  $t$  is a repeat, as opposed to a first attempt. In general, we find that, the policy significantly increases the probability of repeating.

Specifically, under the policy period, the probability of a course being a repeated course is 2.31 percentage points higher than that in the no-policy period. This is shown in column (1), a baseline model that includes the individual fixed effects and the individual's  $a - th$  term fixed effects and covers the entire population. In column (2), we include course fixed effects, which capture any unobserved heterogeneity at the course level, such as subject, course level (which may reflect difficulty,) specific course grading policies, etc. Column (3) adds a control variable of the number of credits taken in the semester for each student, counting for potential confounding effect of curriculum choice, and column (4) excludes all students who transferred from other institutions to the University. The estimates are highly robust across different specifications. On average, the repeat rate for policy-off periods is 3.34% in the analysis sample. Thus, our analysis shows the policy increases a student's probability of repeating by up to 64%.

We further investigate the dynamics of the policy variations during the periods in which the "Grade Forgiveness" policies were in place, abolished, and reintro-

duced. The results are shown in Table B1. We show that the policy’s effect on the probability of repeating is robust and consistent for both the cohort who experienced a policy turn-off and the cohorts who experienced a policy turn-on. This provides convincing evidence for our identification strategy that the policy’s effects are not subject to specific time periods nor are cohort-sensitive.

### 6.1.2 Threshold Grade to Repeat

We show above that the probability of repeating significantly increases when the “Grade Forgiveness” policy is in effect. This is consistent with Proposition 5 in the theory section, which implies that the threshold grade to repeat is either decreasing or increasing within a small range. To show empirical evidence, we estimate the policy’s effect on the probability of repeating by grades. Specifically, we run a regression of an indicator of whether a course has been repeated on an interaction term of the policy indicator and the course grade of the first attempt (either numeric or letter grade).

Table 3 shows the results. Note that the dependent variable here is an indicator of whether a course (i.e., the first attempt) is being repeated (in any subsequent semesters), which is different from the dependent variable in Table 2, an indicator of whether a course is a second attempt. The regression can be written as:

$$Repeated_{ijat} = \beta_0 + \beta Policy_{t'} \times Grade_{ijat} + \lambda Policy_{t'} + \gamma_i + \delta_a + \theta_j + \varepsilon_{ijat}, \quad t' > t \quad (4)$$

The repetition decision could be made in response to the policy in any subsequent period  $t'$ . We show results using 3 specifications of policy, current, one-semester-lagged, and two-semester-lagged policy. Columns (1)–(3) use the numerical grade (grade points) and columns (4)–(6) use the categorical grade. Estimates in columns (1)–(3) imply that, on average, the increase in repeating is driven by more repeating of lower first-attempt grades, echoing Proposition 5 in the theory that the threshold grade to repeat decreases.

Columns (4)–(6) show that the policy incentivizes more repeats for courses graded C, D, or F, relative to courses in which students received an A. Specifically, the likelihood of repeating a grade D course significantly increases by 9.2 percentage points, and the probability of repeating a grade F increases by 5.97 percentage points relative to getting a grade A, when we use the current-period policy. Not surprisingly, the effects are even stronger when we look at the one-semester-lagged and two-semester-lagged policy because most repeats happen in the next two semesters, as shown in Figure 3. It is worth noting that the probability of repeating a grade C course increases significantly under the policy, however the magnitudes are relatively small. This is interesting because it implies that students are not only repeating “failing” grades (i.e., grades below D) but also “unsatisfying” grades, which may be a particularly attractive option for high-achieving students or students that need certain GPAs to be admitted into or progress in their major.

We also find that students who withdraw from classes and thus receive grades of W are significantly less likely to repeat the course, echoing our theoretical implication and later finding that more courses are attempted and withdrawn from under the policy. Moreover, it also consistent with our later finding that students are more likely to take more challenging courses. Together, the evidence above suggests that students are more likely to challenge themselves with difficult courses and greater workloads but are also more likely to drop courses without repeating. It is possible that these courses are not important or required to the students and would not have been taken in the first place without the policy.

## 6.2 Choosing More Challenging Courses

Having shown the policy’s effect on the probability of repeating, we have justified both the mechanism and validity of exploring the policy’s indirect effect on all students. One of the most important questions we would like to ask in this paper is whether students choose to challenge themselves, given that the “Grade Forgiveness” policy serves as an insurance mechanism for potential bad outcomes. The

idea is, if the option value of this insurance outweighs the expected cost, we should observe that more students are incentivized to take more challenging courses and (or) more credits per semester. We study this question by estimating the differences in the level of each course (as a proxy for how difficult or challenging it may be) as well as the number of credits attempted and the overall level of difficulty in the course bundle that students take in each semester, between the treated and untreated semesters.

We start by estimating to what extent the policy nudges a student to take a STEM course. Literature has shown that the harsher grading practices observed in STEM subjects often deters students' participation, even though STEM jobs are among the highest-paying jobs. Thus, we use an indicator of whether a course in the transcript is a STEM subject course as the outcome variable. Table 4 shows the corresponding results: the policy's effect on the likelihood of taking a STEM course. We first use a conservative definition of STEM subjects for defining the outcome variable in Panel.<sup>16</sup> We carry this analysis in the full sample and two sub-samples. Column (1) shows that, on average, the "Grade Forgiveness" policy nudges an average student to take a STEM course by 1.91 percentage points. Column (2) shows that, for non-transfer students, the policy increases their probability of taking a STEM course by 2.37 p.p, accounting for a 9.5% increase from the sample mean. More importantly, for students who have not yet declared a major, the policy increases their likelihood of enrolling in a STEM course by 2.42 p.p., accounting for a 10.6% increase from the sample mean. This is not trivial, considering the small body of STEM majored students and the possibility that this group of students might end up majoring in STEM. We use another definition of STEM subjects, based on DHS STEM Designated Degree Program List<sup>17</sup>, which could be consider as a broader definition of

---

<sup>16</sup>Our conservative definition of STEM defines all natural sciences, engineering, and most medical sciences as STEM subjects.

<sup>17</sup>The U.S. Department of Homeland Security (DHS) STEM Designated Degree Program List is a complete list of fields of study that DHS considers to be science, technology, engineering or mathematics (STEM) fields of study for purposes of the 24-month STEM optional practical training extension described at 8 CFR 214.2(f). See: <https://www.ice.gov/sites/default/files/documents/stem-list.pdf>



STEM, in Panel 2 to further explore the curriculum choice as well as to check the robustness of Panel 1. Estimates in Panel 2 shows that the policy has a similar effect on nudging student to take STEM subject courses under the broader definition.

One could argue that using STEM to measure difficulty is not accurate or comprehensive. For instance, some students could be good at Math but feel challenged in arts or languages. To broaden the definition of difficulty, we adopt an alternative and continuous measure by observing the grading harshness directly from the actual grades. The *difficulty* is calculated by taking the mean of all grades assigned in each course ever offered in the entire sample, rather than the semester-level or section-level mean, which might be subject to instructor-specific grading behaviors, students' selection into certain sections taught by certain instructors, and the possible interaction with the policy changes. Intuitively, the higher(lower) the share of grade D/F(A/B), the more difficult the course.

Considering the usual suspect, grade inflation, broadly observed in many US colleges over the past decades, we adjust for time-specific effects from the raw grades by taking the residuals from the regression of raw grades on academic progress fixed effects. Specifically, for each course  $j$ ,  $Difficulty_j = \frac{\sum_i^N \mathbb{1}[\hat{D} \text{ or } \hat{F}]_{ij}}{N}$ , where  $i$  is an individual grade at the student-attempt level,  $j$  is a course,  $N$  is the total number of students who took this course in the entire sample,  $\hat{D}$  or  $\hat{F}$  is the regression adjusted indicator for letter grades D or F. Besides the grading harshness of a particular course, we consider the number of credits as a multiplier of the course-level difficulty, based on the GPA calculation rule and the fact that choosing the difficulty level and the credits are simultaneous decisions when choosing a course. That is,  $Total\ Difficulty_j = \frac{\sum_i^N \mathbb{1}[\hat{D} \text{ or } \hat{F}]_{ij}}{N} \times Credits_j$ .

Panel 1 in Table 5 shows the estimation results. The dependent variable in columns (1) and (2) is difficulty, measured as (1) the fraction of letter grade D of F, and (2) the fraction of letter grades D, F or W. The results indicate that under the "Grade Forgiveness" policy, students chose to enroll in courses that are moderately more difficult. Specifically, the policy increases the course difficulty level chosen by

students by about 5% from the sample mean. Taking the number of credits of each course in to account, the outcome variables in columns (3) and (4) are total difficulty, measured by (3) the fraction of letter grades D or F multiplied by the number of credits, and (4) the fraction of letter grades D, F or W multiplied by the number of credits. We see that the policy nudges students to enroll in more difficult courses by about 6%–7%. We also explore alternative difficulty measures, such as letter grades (e.g., grades A or B) and numeric grades, and show that the estimation results are qualitatively consistent with our main results (see Appendix, Table B3).

Lastly, we explore the impact of the policy on curriculum choice and variation at the term-level. Choosing to enroll in more courses or credits can be considered as a way that students attempt to challenge themselves. Panel 2 in Table 5 shows estimates of the "Grade Forgiveness" policy's effect on the number of courses, the number of total credits, and the term-total *Total Difficulty*. Column (1) shows that students take 1.6% more courses and 1.9% more total credits in a semester under the "Grade Forgiveness" policy. Together, columns (1)–(2) suggest that students are incentivized by the policy to challenge themselves by taking moderately more difficult courses and taking more credits. Columns (3) and (4), using a measure that takes both the average difficulty level and the credit load into consideration, show that under the policy, the "total difficulty" level that students choose to enroll in a semester increases by about 8% from the sample mean. We have also explored alternative measures of the course difficulty by using the share of letter grade A and numeric grade points. The results (see Table B3) are qualitatively and quantitatively similar.

### 6.2.1 Heterogeneity in Gender

Policymakers and researchers have been concerned about the under-representation of women in STEM fields, given the expected shortage of STEM workers and the likely effects of the gender gap in college major choice on the pre-existing gender wage gap. A consensus in the recent literature is that women value grades sig-

nificantly more than men, and the nature of STEM subjects' often-harsher grading practices deters women's participation in STEM (Rask and Tiefenthaler, 2008; Ost, 2010; Owen, 2010; Butcher, McEwan, and Weerapana, 2014; Ahn et al., 2019; Minaya, 2017). A related and relevant question that our study can answer is that, if given another chance, are women more likely than their male counterparts to retake a course for better grades, and thus to pursue a STEM major? Although the "Grade Forgiveness" policy does not guarantee better grades immediately (like updating syllabus or grading policies suggested in the previous literature), it merely offers students a costly insurance against the lower grade. A student must forgo time, effort, reduced ability to take other courses and possible other additional mental costs by retaking a course. If the insurance value of using grade forgiveness outweighs its cost, we would see it nudging more women to retake courses.

We estimate the policy's gender-specific effect on course repeating and course enrollment in STEM subjects. Table 6 shows the estimates. Columns (1) and (2) indicate that, on average, the policy has a significantly weaker effect on course repeating for women than for men, even after controlling for the course grade of the first attempt. Thus, it is not surprising that the policy actually incentivizes women significantly less than men to enroll in STEM courses. Indeed, we see that women are 1.64–1.76 percentage points less likely to enroll in a STEM course in columns (3) and (4).

This finding suggests that the perceived cost associated with repeating a course is higher for women than for men, off-setting the gender difference in preferences of higher grades. Moreover, since men are more likely to be nudged by this policy to initially enroll in a STEM course, even before declaring their majors, implies that the option value of the "insurance" provided by this policy is larger for men than for women. A possible explanation resides in the literature on gender differences in risk aversion (Borghans, Heckman, Golsteyn, and Meijers, 2009). An alternative explanation could be related to the gender differences in negative signal taking and overconfidence (Stinebrickner and Stinebrickner, 2012): low grades may

be perceived as low ability or major mismatch for women; thus, they are less likely to choose and persist in the subject. Though the “Grade Forgiveness” policy is not efficient in mitigating the gender gap in college STEM majors, we complement previous literature by showing a need for an effective policy in reducing take-up costs for women to reduce the gender gap in these college majors.

### **6.3 Effort Allocation and Performance**

Proposition 3 suggests that a student’s grades in a term vary more under the “Grade Forgiveness” policy era compared to the “averaging” policy era because students might change their strategy of allocating effort or time on each course so that we should observe a larger within-term grade variation after the policy.

To address this proposition, we create two primary measures of the distribution of grades obtained in a semester: the standard deviation of grades in a semester and the difference between the max and min grades. The outcome variables are measured at the individual-semester level. The results are shown in Panel 1, Table 7. Overall, we observe that students’ grade distributions are more dispersed under the “Grade Forgiveness” policy.

Column (1) shows that the standard deviation of within-term grades increases by 0.037 points by the policy, explaining 5.4% increase in the changes in the dependent variable. Column (2) shows that the min-max gap of grades in a semester increases by 6.1%. Both estimates indicate that the “Grade Forgiveness” policy increases the grade dispersion in a semester, which is consistent with the implication in Proposition 3. Columns (3)-(4) replicate the regressions presented in columns (1)-(2) but add a vector of course fixed effects to control for changes of grade distribution attributed to changes in the number of courses attempted. With the changes in the number of courses attempted taken into account, the standard deviation of grades increases by 3.2%, and the min-max gap of grades increases by 3%. Overall, the “Grade Forgiveness” policy moderately increased the grades dispersion likely due to students’ increased likelihood of re-allocating their time among courses in

response to the policy in a way that favors the more promising courses.

### 6.3.1 Performance on the First Attempt

We next dive deeper into the changes in grade distributions, focusing on first-attempt courses. There are three potential competing channels that affect students' first-attempt grades. First, students are making different enrollment decisions under different policy periods, as shown in the section above. Therefore, they may face non-uniform levels of course difficulty, and thus may receive different grades. On average, we should expect that students are getting worse grades on their first-attempt as they are taking harder courses overall. Second, students might slack in classes and get worse grades when Grade Forgiveness policy is present, assuming that the value of the first-attempt grade diminished when they can retake to improve grades later. Although, from an empirical point, it is plausible that students would not waste their first-attempt due to the high cost of retaking courses—including monetary cost, time cost, and mental cost. Last, students re-allocate their effort among courses taking in the same semester with the observation of the policy, as shown in Table 7, Panel 1, which makes it hard to predict the direction of change. To test how the policy affects students' performance on the first attempt, we compare within-course differences in performance of the first-time-takers between the pre-policy and the post-policy periods, using the same estimation strategy in equation 1.

Panel 2 in Table 7 shows the policy's effect on the first-attempt grades. Each column uses a different definition of the grade: numerical grade, an indicator of getting good grade (letter grade A or B), letter grade C, an indicator of getting failing grade (letter grade D or F or W). We find a 1.17% decrease in the numerical grade of the first-attempt, a 0.67% decrease in the likelihood of getting a good grade (letter grade A or B), a 1.99% decrease in the likelihood of getting a letter grade C, and a 4.62% increase in getting a failing grade (D, F, or W). Overall, the policy is associated with slightly worse first-attempt performance. It is worth noting that

we cannot run a multinomial logit with our multiple fixed effects design, but we do observe that the increase in the likelihood of getting D/F/W is driven by the large increase in W, which echoes our finding above regarding the raised withdraw rate. In the meantime, although it is reasonable to assume that students may “skate” classes by counting on the policy to forgive their bad grades, we cannot empirically disentangle this effect from the the effects of taking difficult courses and the effects of effort optimizing. Notwithstanding, the decrease in the numerical grades and the probability of getting A or B is small.

### 6.3.2 Earned and Withdrew

As shown above, on average, students attempted more courses and more credits under the “Grade Forgiveness” policy. However, we find that students are also more likely to withdraw from a course under the policy. Specifically, on average, the number of a course being withdrawn in a semester increases by 0.0589 under the policy, which is a 31.3% increase from the sample mean withdrawal rate per semester (0.174). The increase in the number of courses attempted and the increase in the number of courses withdrawn seemingly *offset* each other. As a result, on average, we find no significant changes in the number of courses completed/earned in response to the policy changes. See table 8. This phenomenon implies a series of interesting student behavioral changes under the policy: students will register for more courses, will register for more challenging courses, will attempt to complete their degrees earlier and/or will learn more, but will also end up withdrawing from more courses. While the number of courses earned per semester does not change under the policy, the number of credits earned per semester moderately increased under the policy (in column 4), meaning that the number of credits earned per course increased. Alongside the evidence shown in Section 6.2, this implies that students might enroll in and complete more difficult courses, which are usually associated with more credits per course.

We cannot measure the “benefit of” the courses from which students end up

withdrawing in this study. Students could arguably waste their time and energy unnecessarily for these courses that they would not have attempted in the first place without the policy. They could also gain knowledge over the short term before dropping or updating their beliefs over the course/subject. These two possible effects might cancel each other out as well.

## 6.4 Repeaters' Subsequent Course Choice and Performance

Having shown the "Grade Forgiveness" policy's total effects on course choice and performance of *all students*, we next turn our attention to the policy's direct effects through *repeating*. What are the impacts of the policy on the repeaters?<sup>18</sup> Particularly, does repeating incentivize repeaters to take another course in the same subject as the failed course? If so, how do the repeaters perform in these subsequently taken courses? Intuitively, most repeaters initially fail courses (or do not meet their own expectations,) and perhaps struggle with the material in the subject. We have reason to believe that those students will avoid taking any other courses in the same subject after the failure. However, hypothetically, the policy encourages students and induce them to retake the courses they failed in their first attempt. Thus, it increases students' learning towards the difficult materials and possibly increases subsequent interest in the same subject. We should expect that there is a positive effect on students' learning and subsequent interest in the subject.

We estimate a Two-Stage Least Squares (2SLS) in equation 5.2 and present the treatment effect of repeating in Table 9. The outcome variables of interest are the number of credits taken in another same-subject course (in a subsequent semester) and the corresponding performance, measured by grade points, an indicator of succeeding in the subsequent course (i.e., getting a letter grade A or B), and an indicator of failing in the subsequent course (i.e., getting a letter grade D or F), respectively. The upper panel shows the reduced form estimates. The lower panel shows the LATE estimates. Overall, we see that students are significantly more likely to take

---

<sup>18</sup>Note that we use "repeater" to refer to all students who have ever repeated a course.

another course in the same subject after they repeated a course under the policy and are performing better. Specifically, students who repeat a course will enroll in 1.16 more credits in the same subject than student who did not repeat. This indicates that a chance to repeat a course significantly improves students' interest, or at least persistence, in the subject that they have failed. The estimates on the performance in a subsequent course when students do take another same-subject course are consistent in signs with the reduced-form estimates, but with larger magnitudes. We find that conditional on taking another same-subject course, repeaters are 52 percentage points more likely to earn good grades (i.e., letter grades A or B) and 54 percentage points less likely to fail (i.e., obtaining letter grades D or F) in the next course.

To sum up, the "Grade Forgiveness" policy encourages repeaters to take a subsequent course in the same subject which they have failed before. Repeaters are more likely to obtain higher grades in the subsequent courses under the policy. This finding implies that the policy has a strong and positive effect on marginal students' persistence and success.

## **6.5 Graduation Rate**

We have shown above the varying effects of the "Grade Forgiveness" policy on students' choices and performance. One may wonder about the policy's overall and ultimate effect on students' success in college. A commonly used measure of students' success is graduation, which is also the primary interest of most institutional research on higher education. On average, the graduation rate during the observation period at BSU is 23.7% for the entry cohorts who first enrolled between 1990 and 2012, among which the first-time (non-transferred) students' graduation rate is 15.3% and the transferred students' graduation rate is 33.2%. It is both common to transfer in and out of BSU and we are unable to observe students who transferred out of BSU to another institution. Thus, we cannot distinguish between dropouts and transfer-outs. Among the observed graduation of first-time students, the av-



erage (and median) time to graduate is 12 semesters, the 75th percentile and 95th percentile of the time to graduate are 14 and the 19 semesters, respectively. Thus, we look at the sample including cohorts who entered the University from 1990 to 2008 to ensure that we observe the majority of the graduation events.

We examine this policy's impact on the probability of graduation by comparing the outcome between different entry cohorts who were exposed to different levels of treatment of the policy. Note that we cannot adopt the same estimation strategy in the previous sections because we now look at the person-level outcome, an indicator of graduation. Considering that the first several semesters are the key time period to "survive" college and the policy's effect will be most pronounced in the first several semesters, we estimate the differences in the likelihood of graduating from BSU between cohorts who were treated for their first X semester(s) and cohorts who were not treated for their first X semester(s).

Table 10 shows the estimation results of our preferred specifications where the independent variable of interest is an indicator of the policy treatment at a certain level. The treated group here is students who have been continuously exposed to the policy in their first X semester(s) and the control group is students who were never exposed to the policy in their first X semester(s). Each column is a separate regression with a different length of treatment, X. For example, column (1) shows the difference in the likelihood of graduating from BSU between cohorts who had the policy in their first semester and cohorts who did not have it in their first semester. We can see the sample size reduction from column (1) to (4) as we include students who were enrolled at BSU for a longer time. We look at both the overall graduation rate and the on-time graduation rate (i.e., graduate in six or fewer years). Overall, we find a zero effect of the "Grade Forgiveness" policy on the probability of ever graduating from BSU for the non-transfer (first-time) students. Again, as mentioned above, it is common for students to transfer to another college from BSU. In this sense, we are forced to treat all the unobserved outflow of students as dropouts. Also, we find no effects of the policy on the on-time graduation rates. This finding

refutes the usual concern that “Grade Forgiveness” policies might cause students to delay their time to graduation from college by repeating too many courses, and thus not graduating on time.

Among the subsample of students who *graduated* from BSU, we do find a significant increase in obtaining a STEM degree for those who were exposed to the policy in the early stages of their study relative to the cohorts who were not. We run a similar set of regressions on the sub-sample of graduates by replacing the outcome variable with an indicator of obtaining a STEM degree and keep all the same right-hand-side variables (as presented in Table 10). Table 11 shows the difference in the probability of obtaining a STEM degree between cohorts who were continuously treated by the policy in their first X semester(s) and cohorts who were never exposed to the policy in their first X semester(s). We find that the policy has a positive and significant effect on the likelihood of obtaining a STEM degree, in both the broader definition and the stricter definition of STEM majors.<sup>19</sup>

Specifically, cohorts who were exposed to the policy in their first semester are more likely to obtain a STEM degree by 4 percentage points (based on the broader definition), relative to cohorts who were not exposed to the policy in their first semester. When we compare groups who have a larger difference in treatment, we see that cohorts who were continuously exposed to the policy in their first 4 semesters are more likely to obtain a OPT STEM degree (conservative STEM degree) by 8.17 (4.12) percentage points, relative to cohorts who were not exposed to the policy in any of the semesters among their first 4 semesters. This effect accounts for a 25% increase in the STEM degree recipients.<sup>20</sup> This finding echoes our earlier finding that the policy nudges students to take difficult courses and STEM courses. We further show that the policy helps students to persist and succeed in STEM majors without a significant “price” by having them take longer to graduate.

---

<sup>19</sup>We use a similar categorization of STEM majors as our categorization of STEM subjects shown above in Section 6.2.

<sup>20</sup>The share of STEM degrees among the sample cohort (1990-2008) is 32.5%.

## 7 Conclusion

Whether college students should be given a second chance in coursework is a fundamental topic in institutional research. We study, so called "Grade Forgiveness" policies in-depth for the first time by investigating changes in course repetition, curriculum choice, effort allocation, and the consequences of repeating courses. We utilize a unique administrative dataset from Boise State University, which includes an observational period that covers two recent changes in "Grade Forgiveness" policies during our data period. We use a fixed effects model, strengthened by the on-off-on policy variations, to estimate the policy's effects on college students' decision-making and outcomes throughout different stages of college.

We find that students are highly responsive to the policy changes, evidenced by their repeating behaviors. As one of the positive effects, this policy significantly increases the likelihood of attempting challenging courses for *all students*. This finding has direct policy implications in that similar practices of providing a safety net for students' academic performance could effectively nudge students to challenge themselves and explore "more difficult" courses and even majors. Furthermore, we find that students who are incentivized by the policy to repeat poor performances for better grades are more likely to take a subsequent course in the same subject, and are more likely to perform better, which suggests a positive answer to the question of whether students deserve another chance at their college studies. Together, these findings imply that "Grade Forgiveness" is an effective way to encourage students to both pursue and persist in challenging courses, and achieve higher levels of success.

We also find that students end up withdrawing from more courses under "Grade Forgiveness" policies in the mean time of attempting more challenging courses. It is out of the scope of this paper to discuss the pros and cons of increased withdrawal, but it will be interesting for future research to explore college resource use and potential gains of students' attempting (then withdrawing) courses in terms of

information and belief updating.

Our paper is the first study on the impact of college course repeating and college GPA policies. It is also the first to use transcript data and multiple policy changes to empirically study a "Grade Forgiveness" policy's impact on students' decisions and performance. We shed light on the practice of this widely-spreading but undiscovered policy. Our findings call for more attention to be directed towards institutional policies on academic requirements and their impact on students.

## References

- Ahn, T., P. Arcidiacono, A. Hopson, and J. R. Thomas (2019). Equilibrium grade inflation with implications for female interest in stem majors. Technical report, National Bureau of Economic Research.
- Astorne-Figari, C. and J. D. Speer (2019). Are changes of major major changes? the roles of grades, gender, and preferences in college major switching. *Economics of Education Review* 70, 75–93.
- Bandiera, O., V. Larcinese, and I. Rasul (2015). Blissful ignorance? a natural experiment on the effect of feedback on students' performance. *Labour Economics* 34, 13–25.
- Bettinger, E. P. and B. T. Long (2009). Addressing the needs of underprepared students in higher education does college remediation work? *Journal of Human resources* 44(3), 736–771.
- Borghans, L., J. J. Heckman, B. H. Golsteyn, and H. Meijers (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association* 7(2-3), 649–658.
- Butcher, K. F., P. J. McEwan, and A. Weerapana (2014). The effects of an anti-grade-inflation policy at Wellesley College. *Journal of Economic Perspectives* 28(3), 189–204.
- Chen, K. and X. Jiang (2021). Pushing the pause button: Can course repetition foster student motivation and future success? *Working Paper*.
- Frisancho, V., K. Krishna, S. Lychagin, and C. Yavas (2016). Better luck next time: Learning through retaking. *Journal of Economic Behavior & Organization* 125, 120–135.
- Goodman, J., O. Gurantz, and J. Smith (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy* 12(2), 115–58.

- Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics* 98(4), 671–684.
- Marx, J. and D. Meeler (2013). Strike four! do-over policies institutionalize gpa distortion. *Quality Assurance in Education* ,(,), ,.
- Minaya, V. (2017). Do differential grading norms across fields matter for major choice? evidence from a policy change in florida. Technical report, Working Paper.
- Minaya, V. (2020). Do differential grading standards across fields matter for major choice? evidence from a policy change in florida. *Research in Higher Education* 61(8), 943–965.
- Ost, B. (2010). The role of peers and grades in determining major persistence in the sciences. *Economics of Education Review* 29(6), 923–934.
- Owen, A. L. (2010). Grades, gender, and encouragement: A regression discontinuity analysis. *The Journal of Economic Education* 41(3), 217–234.
- Rask, K. and J. Tiefenthaler (2008). The role of grade sensitivity in explaining the gender imbalance in undergraduate economics. *Economics of Education Review* 27(6), 676–687.
- Reshwan, R. (2016). Does gpa matter when applying for a job?
- Stinebrickner, R. and T. Stinebrickner (2014). Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model. *Journal of Labor Economics* 32(3), 601–644.
- Stinebrickner, R. and T. R. Stinebrickner (2013). A major in science? initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies* 81(1), 426–472.
- Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics* 30(4), 707–748.

Tafreschi, D. and P. Thiemann (2016). Doing it twice, getting it right? the effects of grade retention and course repetition in higher education. *Economics of Education Review* 55, 198–219.

Vigdor, J. L. and C. T. Clotfelter (2003). Retaking the sat. *Journal of Human resources* 38(1), 1–33.

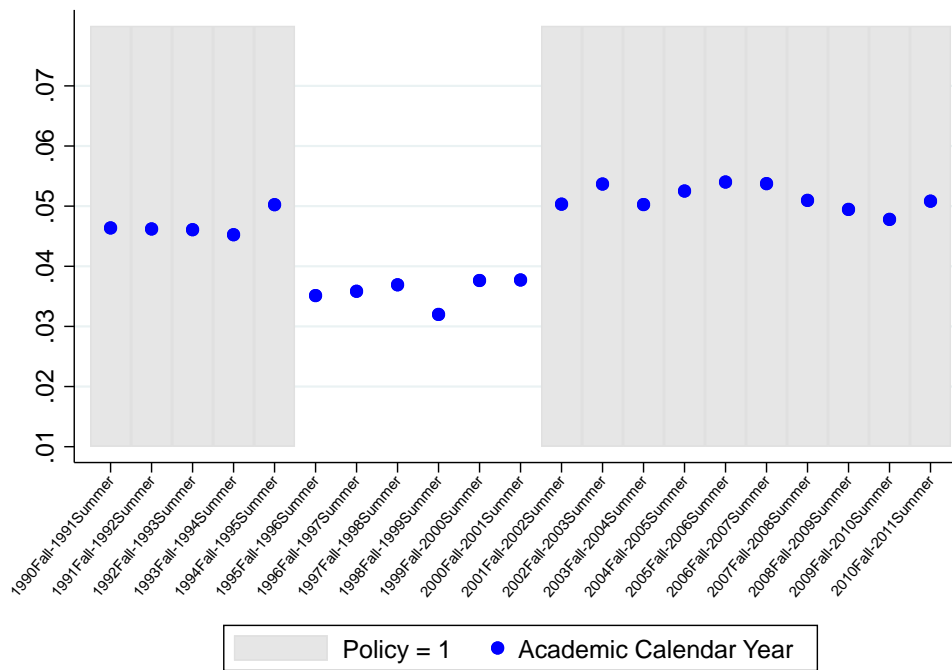


Figure 1: Probability of A Course Being Repeated

Note: This figure shows the adjusted sample mean of the probability of a course being repeated by academic calendar year. The mean is regression-adjusted by controlling course fixed effects, term season fixed effects, and individual's t-th academic semester. The policy is present in calendar semesters Spring 1990–Summer 1995 and Fall 2001–present and the policy is not in present in calendar semesters Fall 1995–Summer 2001.



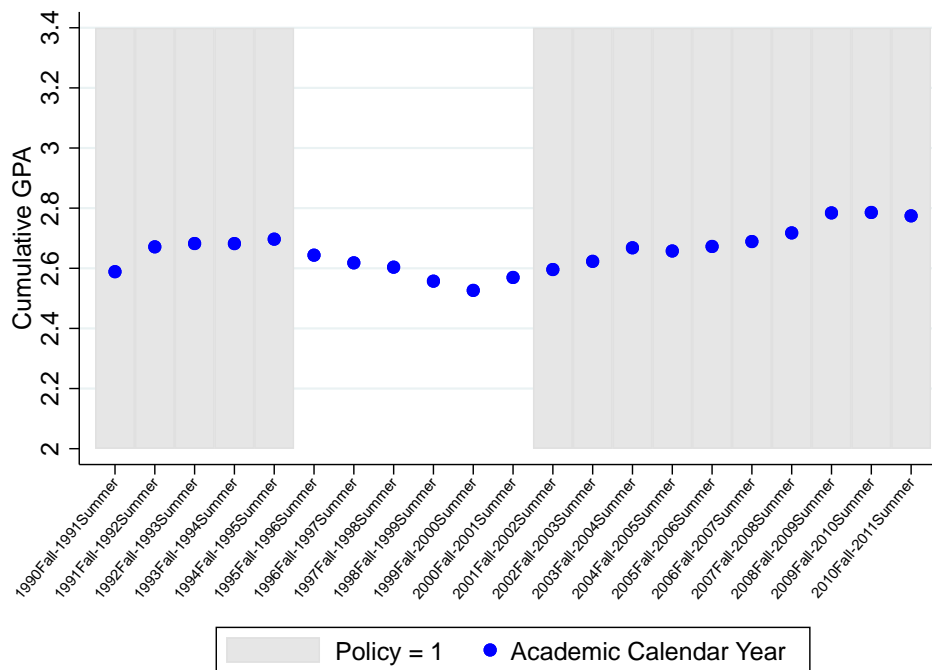
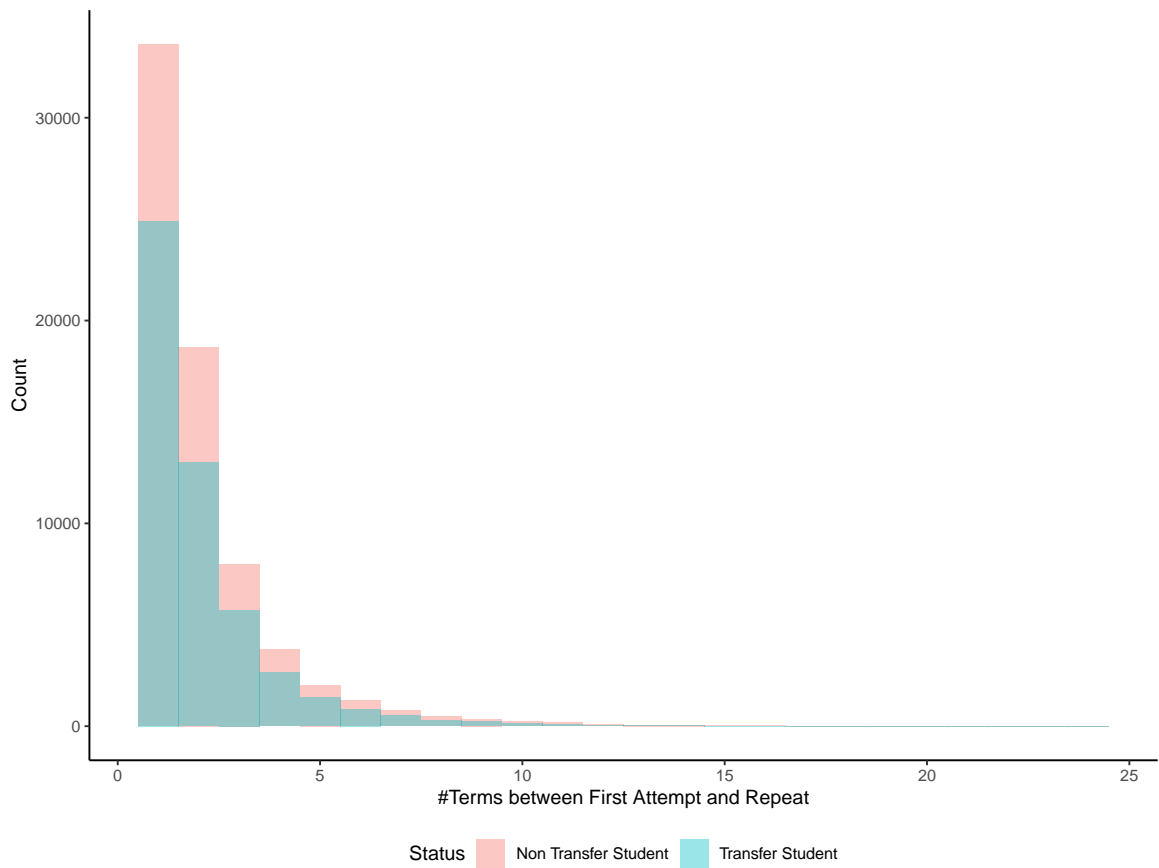


Figure 2: Evolution of Cumulative GPA at BSU

Note: This figure shows the sample mean of cumulative GPA of the non-transfer students in BSU by academic calendar year. It indicates an apparent downturn of cumulative GPA during the period when the Grade Forgiveness policy was not in present. The policy is present in calendar semesters Spring 1990–Summer 1995 and Fall 2001–present and the policy is not in present in calendar semesters Fall 1995–Summer 2001.



**Figure 3: Density of #Terms between First Attempt and Repeat**

Note: This figure shows the distribution of number of terms between a first attempt and the second attempt (repetition) by transfer students (blue) and non-transfer students (pink) separately. It indicates most of repeats happened in the next few semesters after the first-attempt.

Table 1: Summary Statistics on Key Variables

| Variable                                | Mean     | Std. Dev. | N         |
|---|----------|-----------|-----------|
| <b>Person Level</b>                     |          |           |           |
| Female                                  | .5415    | .4983     | 169,722   |
| SAT composite                           | 1478.821 | 253.0717  | 86,628    |
| Transfer student                        | .4504    | .4975     | 170,812   |
| Total #Repeat                           | .7919    | 1.9024    | 170,812   |
| Graduation Rate                         | .2790    | .4485     | 170,812   |
| Exit GPA                                | 2.62288  | 1.1136    | 170,812   |
| Graduate GPA at BSU                     | 3.1750   | .4548     | 47,614    |
| # Years to Graduate                     | 5.1804   | 3.4529    | 47,614    |
| %STEM Degrees (conservative definition) | .2435    | .4292     | 45,674    |
| %STEM Degrees (OPT definition)          | .3848    | .4865     | 45,674    |
| <b>Term Level</b>                       |          |           |           |
| #Courses Attempted                      | 3.729965 | 1.873759  | 806,577   |
| #Credits Attempted                      | 10.56368 | 4.868073  | 806,577   |
| #Courses Earned                         | 3.60792  | 1.842826  | 806,577   |
| #Credits Earned                         | 9.115406 | 5.124696  | 806,577   |
| #Courses Withdrew                       | .1220454 | .4083085  | 806,577   |
| <b>Course Level</b>                     |          |           |           |
| Repeat                                  | .0480351 | .2138405  | 2,769,204 |
| Grade Points                            | 2.760448 | 1.301253  | 2,769,204 |
| STEM (OPT)                              | .256222  | .4365459  | 2,769,204 |
| STEM (conservative)                     | .3121514 | .4633713  | 2,769,204 |

Note: This table describes the data we used in the analysis, which covers students who enter BSU between 1990–2019. The upper, middle, and bottom panels shows person-level, term-level, and course-level statistics, respectively. ACT scores are converted to SAT composite scores when only ACT scores are available.

Table 2: Policy's Effect on the Probability of Repeating

|                          | (1)                   | (2)                   | (3)                    | (4)                    |
|--------------------------|-----------------------|-----------------------|------------------------|------------------------|
|                          | Repeat=1              | Repeat=1              | Repeat=1               | Repeat=1               |
| Policy                   | 0.0231***<br>(0.0008) | 0.0183***<br>(0.0008) | 0.0190***<br>(0.0008)  | 0.0223***<br>(0.0011)  |
| #Credits in Term         |                       |                       | -0.0033***<br>(0.0001) | -0.0037***<br>(0.0001) |
| Course FE                | NO                    | YES                   | YES                    | YES                    |
| Include Transfer Student | YES                   | YES                   | YES                    | NO                     |
| Observations             | 2745415               | 2745415               | 2745415                | 1299005                |
| Sample Mean (Policy = 0) | 0.0334                | 0.0334                | 0.0334                 | 0.0351                 |

Note: This table shows the policy's effect on the probability of repeating. The regression sample include all course level observations excluding incomplete grades, audits, or courses or sessions offering Pass/No Pass grades. Each column is a separate regression specified in Equation 1. The independent variable is an indicator of whether a course shown on a student's transcript is a repeat (non-first time taking). Each regression includes individual fixed effects, academic term-t fixed effects, number of courses offered in a certain semester, number of peers in a certain semester and major, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.

Table 3: Policy's Effect on Probability of Repeating by Grades

|                | (1)                    | (2)                    | (3)                    | (4)                    | (5)                    | (6)                    |
|----------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Dep. Var.:     | Repeated=1             | Repeated=1             | Repeated=1             | Repeated=1             | Repeated=1             | Repeated=1             |
| Specification: | Policy                 | Policy_1               | Policy_2               | Policy                 | Policy_1               | Policy_2               |
| Policy         | 0.0219***<br>(0.0022)  | 0.0319***<br>(0.0022)  | 0.0402***<br>(0.0023)  | -0.0038***<br>(0.0006) | -0.0032***<br>(0.0006) | -0.0032***<br>(0.0006) |
| Policy × Grade | -0.0073***<br>(0.0006) | -0.0101***<br>(0.0007) | -0.0126***<br>(0.0007) |                        |                        |                        |
| Policy × A     |                        |                        |                        | 0.0000<br>(.)          | 0.0000<br>(.)          | 0.0000<br>(.)          |
| Policy × B     |                        |                        |                        | 0.0002<br>(0.0004)     | 0.0005<br>(0.0004)     | 0.0006<br>(0.0004)     |
| Policy × C     |                        |                        |                        | 0.0048***<br>(0.0007)  | 0.0058***<br>(0.0007)  | 0.0065***<br>(0.0008)  |
| Policy × D     |                        |                        |                        | 0.0920***<br>(0.0038)  | 0.1027***<br>(0.0039)  | 0.1116***<br>(0.0040)  |
| Policy × F     |                        |                        |                        | 0.0597***<br>(0.0034)  | 0.0740***<br>(0.0035)  | 0.0860***<br>(0.0037)  |
| Policy × W     |                        |                        |                        | -0.0652***<br>(0.0039) | -0.0639***<br>(0.0040) | -0.0623***<br>(0.0041) |
| Constant       | 0.2531***<br>(0.0020)  | 0.2656***<br>(0.0021)  | 0.2783***<br>(0.0021)  | -0.0090***<br>(0.0007) | -0.0043***<br>(0.0008) | -0.0000<br>(0.0008)    |
| Observations   | 2745415                | 2591153                | 2457520                | 2745415                | 2591153                | 2457520                |

Note: This table shows the policy's effect on grades to be repeated. The independent variable is an indicator of whether a first-time-taking course shown on a student's transcript has been repeated. Column (1)–(3) are separate regressions where the variable of interest is an interaction term of the policy and the numeric grade points; column (4)–(6) are separate regressions where the variable of interests are interaction terms of the policy and the letter grades A, B, C, D, F, or W. Each regression includes individual fixed effects, academic term t fixed effects, course fixed effects, and #credits taken in a semester by each student. Considering the repeats take places a few semesters later than the first-attempt, we use current policy, one-semester lagged policy, and two-semesters lagged policy as different specifications in columns (1), (2), and (3). Same for columns (4), (5), and (6).

Table 4: Policy's Effect on Choice of Difficulty: STEM Course

|  | (1)                   | (2)                   | (3)                   |
|--|-----------------------|-----------------------|-----------------------|
| Dep. Var.                                    | STEM Course           | STEM Course           | STEM Course           |
| Sample                                       | All                   | Non-transfer          | Undeclared Major      |
| <b>Panel 1: Conservative STEM Definition</b> |                       |                       |                       |
| Policy                                       | 0.0191***<br>(0.0017) | 0.0237***<br>(0.0022) | 0.0242***<br>(0.0049) |
| Sample Mean                                  | .2558                 | .2492                 | .2279                 |
| <b>Panel 2: OPT-based STEM Definition</b>    |                       |                       |                       |
| Policy                                       | 0.0215***<br>(0.0018) | 0.0262***<br>(0.0024) | 0.0279***<br>(0.0052) |
| Sample Mean                                  | .3117                 | .3092                 | .3023                 |
| N  | 2745415               | 1299333               | 220954                |

Note: This table shows the policy's effect on the likelihood of taking a STEM course by three sub-samples. The dependent variable of interest is an indicator of a course being a STEM course. Panel 1 and panel 2 uses two different definition of STEM: conservative and OPT-based definition. Column (1) covers the entire sample, column (2) includes only the non-transfer students, and column (3) includes only the non-transfer students who have not yet declared a major. Each regression includes individual fixed effects, academic term t fixed effects, number of courses offered in a certain semester, number of peers in a certain semester and major, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.

Table 5: Policy's Effect on Choice of Difficulty: Measures using Grades

| <b>Panel 1: Course Level</b> |                       |                       |                       |                       |
|------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                              | (1)                   | (2)                   | (3)                   | (4)                   |
|                              | %DF                   | %DFW                  | %DF×Credits           | %DFW×Credits          |
| Policy                       | 0.0061***<br>(0.0003) | 0.0073***<br>(0.0004) | 0.0227***<br>(0.0011) | 0.0288***<br>(0.0014) |
| Sample Mean                  | 0.1205                | 0.1619                | 0.3746                | 0.4998                |
| Observations                 | 2707279               | 2707279               | 2707279               | 2707279               |

**Panel 2: Term Level**

|              | (1)                   | (2)                   | (3)                   | (4)                   |
|--------------|-----------------------|-----------------------|-----------------------|-----------------------|
|              | #Credits              | #Courses              | Total %DF×Credits     | Total %DFW×Credits    |
| Policy       | 0.1989***<br>(0.0264) | 0.0585***<br>(0.0102) | 0.1031***<br>(0.0044) | 0.1281***<br>(0.0058) |
| Sample Mean  | 10.5637               | 3.7300                | 1.2686                | 1.6924                |
| Observations | 758813                | 758813                | 758813                | 758813                |

Note: This table shows the policy's effect on course-level (panel 1) and term-level (panel 2) measures of difficulty. Outcome variables in columns (1)–(4) of Panel 1 are course-level difficulty measured by fraction of letter grade D or F by course, fraction of letter grade D or F or W by course, the product of the fraction of letter grade D or F by course and the course credits, and the product of the fraction of letter grade D or F or W by course and the course credits. Outcome variables in columns (1)–(3) of Panel 2 are: number of credits attempted in each term, number of courses attempted in each term, sum of the product of the fraction of letter grade D or F (or W) by course and the course credits in each term (column 4). Each regression includes individual fixed effects, academic term  $t$  fixed effects, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.

Table 6: Heterogeneity: Gender-specific Effects in STEM Course

|                  | (1)                    | (2)                    | (3)                    | (4)                    |
|------------------|------------------------|------------------------|------------------------|------------------------|
|                  | isRepeat               | Repeated               | STEM (OPT)             | STEM (conservative)    |
| Policy           | 0.0116***<br>(0.0014)  | 0.0224***<br>(0.0014)  | 0.0377***<br>(0.0040)  | 0.0337***<br>(0.0039)  |
| Policy × Female  | -0.0050***<br>(0.0018) | -0.0053***<br>(0.0018) | -0.0176***<br>(0.0053) | -0.0164***<br>(0.0050) |
| #Credits in Term | -0.0037***<br>(0.0001) | 0.0004***<br>(0.0001)  | -0.0031***<br>(0.0002) | -0.0031***<br>(0.0001) |
| Grade Points     |                        | -0.0760***<br>(0.0004) |                        |                        |
| Observations     | 1276871                | 1276871                | 1276871                | 1276871                |

Note: This table shows the policy's gender-specific effect on probability of repeating a course and taking a STEM course. The independent variable in columns (1)–(4) are the probability of one course being a repeat, the probability of one course being repeated (later), an indicator of a course is a STEM course by OPT definition and by conservative definition. Each regression includes individual fixed effects, academic term  $t$  fixed effects, course fixed effects (in columns 1 and 2), a policy indicator and an interaction term of the policy and gender. Column (2) also controls for the course grade points. Standard errors are clustered at the individual level.

Table 7: Policy's Effect on Effort Allocation

| <b>Panel 1: Term Level</b> |                       |                       |                       |                       |
|----------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                            | (1)                   | (2)                   | (3)                   | (4)                   |
|                            | SD of Grades          | Max-Min Grade         | SD of Grades          | Max-Min               |
| Policy                     | 0.0374***<br>(0.0038) | 0.0909***<br>(0.0081) | 0.0223***<br>(0.0031) | 0.0442***<br>(0.0061) |
| Control for #Courses       | NO                    | NO                    | YES                   | YES                   |
| Sample Mean (Policy-off)   | .6942084              | 1.488054              |                       |                       |
| Observations               | 380510                | 380510                | 380509                | 380509                |

| <b>Panel 2: Course Level</b> |                        |                      |                       |                       |
|------------------------------|------------------------|----------------------|-----------------------|-----------------------|
|                              | (1)                    | (2)                  | (3)                   | (4)                   |
|                              | GradePoints            | AB                   | C                     | DFW                   |
| Policy                       | -0.0311***<br>(0.0066) | -0.0043*<br>(0.0024) | -0.0038**<br>(0.0018) | 0.0081***<br>(0.0019) |
| Sample Mean (Policy-off)     | 2.664325               | .6339883             | .1908325              | .1751792              |
| Observations                 | 651020                 | 651020               | 651020                | 651020                |

Note: This table shows empirical evidence of the policy's effect on students' effort allocation. To minimize the potential grade inflation over time, we focus on the sub-sample of Spring1990–Spring 2006. Panel 1 shows policy's effect on grade distribution in a semester, where the independent variables in Panel 1 are two measures of grade dispersion in a semester: standard deviation of grade points (columns 1 and 3) and the gap between the max and min grade points in a semester (columns 2 and 4). Each regression includes individual fixed effects, academic term  $t$  fixed effects, number of courses offered in a certain semester, number of peers in a certain semester and major, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Columns (3) and (4) add in number of courses enrolled as an additional control. Panel 2 shows the policy's effect on the first-attempt performance. Each column uses a different measure of the grade: numerical grade, an indicator of getting good grade (letter grade A or B), C, indicator of getting failing grade (letter grade D or F or W). Each regression includes individual fixed effects, academic term  $t$  fixed effects, course fixed effects (for specifications in panel 2), and a policy indicator as explained above. Standard errors are clustered at the individual level.

Table 8: Policy's Effect on Number of Courses and Credits Earned and Withdrawn in a Semester

|              | (1)                   | (2)                   | (3)                 | (4)                   |
|--------------|-----------------------|-----------------------|---------------------|-----------------------|
|              | #Courses Attempted    | #Courses Withdrew     | #Courses Earned     | #Credits Earned       |
| Policy       | 0.0556***<br>(0.0099) | 0.0589***<br>(0.0033) | -0.0032<br>(0.0098) | 0.1518***<br>(0.0275) |
| Sample Mean  | 3.659271              | .1738723              | 3.485399            | 8.821633              |
| Observations | 793585                | 793585                | 793585              | 793585                |

Note: This table shows the policy's effect on students' courses and credits earned and withdrew at semester-level. Each regression includes individual fixed effects, academic term  $t$  fixed effects, number of courses offered in a certain semester, number of peers in a certain semester and major, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.



Table 9: Probability of Taking Another Same-Subject Course and Performance

|  | (1)                   | (2)                   | (3)                   | (4)                    |
|--|-----------------------|-----------------------|-----------------------|------------------------|
|  | #Credits              | Grade Points          | A/B                   | D/F                    |
| <b>Panel 1: Reduced Form Estimates</b> |                       |                       |                       |                        |
| Policy                                 | 0.0237***<br>(0.0056) | 0.0143**<br>(0.0059)  | 0.0077***<br>(0.0022) | -0.0080***<br>(0.0015) |
| First-attempt Grade Points             | 0.1006***<br>(0.0010) | 0.0596***<br>(0.0013) | 0.0218***<br>(0.0005) | -0.0124***<br>(0.0004) |
| <b>Panel 2: IV Estimates</b>           |                       |                       |                       |                        |
| $\widehat{Repeat}_{ijtc}$              | 1.1609***<br>(0.2744) | 0.9648**<br>(0.3960)  | 0.5205***<br>(0.1483) | -0.5389***<br>(0.1029) |
| First-attempt Grade Points             | 0.1724***<br>(0.0170) | 0.1225***<br>(0.0258) | 0.0557***<br>(0.0097) | -0.0475***<br>(0.0067) |
| <i>N</i>                               | 2742040               | 1268511               | 1268511               | 1268511                |

Note: This table shows course repeating's effects on the number of credits taken on subsequent courses (after repeating) in the same subject and their performance. Panel 1 shows the reduced form estimates and panel 2 shows the IV estimates. The outcome variable in column (1) is the number of credits of a same-subject course a student enroll in after repeating a course. The dependent variables in column (2)–(4) are grade points, an indicator of succeeding in the subsequent course (i.e., getting a letter grade A or B), and an indicator of failing in the subsequent course (i.e., getting a letter grade D or F), respectively. Each regression includes individual fixed effects, academic term *t* fixed effects, course fixed effects, first-attempt grade points, number of credits attempted in the term and average course-section grade points for the subsequent course (for column (2)–(4)), and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level in the brackets.

Table 10: Probability of Graduating from BSU

|   | (1)                | (2)                 | (3)                 | (4)                 |
|---|--------------------|---------------------|---------------------|---------------------|
|   | First 1 Term       | First 2 Terms       | First 3 Terms       | First 4 Terms       |
| <b>Panel 1 Dependent Var: Graduation</b>        |                    |                     |                     |                     |
| Treated for the first X terms                   | 0.0059<br>(0.0078) | 0.0007<br>(0.0112)  | -0.0045<br>(0.0151) | -0.0115<br>(0.0218) |
| <b>Panel 2 Dependent Var: 6 Year Graduation</b> |                    |                     |                     |                     |
| Treated for the first X terms                   | 0.0045<br>(0.0053) | -0.0025<br>(0.0080) | -0.0126<br>(0.0119) | -0.0259<br>(0.0180) |
| Observations                                    | 67743              | 41971               | 28293               | 22252               |

Note: This table shows the policy's effect on students' probability of graduation. This sample include cohorts who first entered BSU between 1990 and 2008 (non-transferred). The outcome variable in Panel 1 is an indicator of graduating from BSU and the outcome variable in Panel 2 is an indicator of on-time graduation (i.e., within 6 years). The variable of interest for columns (1)-(4) are: continuously treated by the policy for the first 1 semester, for the first 2 semesters, for the first 3 semesters, and for the first 4 semesters, where the corresponding control group for each treatment group in columns (1)-(4) are *never* treated by the policy in the first 1 semester, in the first 2 semesters, in the first 3 semesters, and in the first 4 semesters. Each regression controls for individual's average SAT score, gender, a linear trend of the entry year, a squared trend of the entry year, a dummy of the entry term season (spring, summer, or fall), and entry major (program) fixed effects. Standard errors are clustered at the entry cohort level.

Table 11: Probability of Obtaining STEM Degrees from BSU, Graduates Subsample

|  | (1)                   | (2)                   | (3)                   | (4)                   |
|--|-----------------------|-----------------------|-----------------------|-----------------------|
|  | First 1 Term          | First 2 Terms         | First 3 Terms         | First 4 Terms         |
| <b>Panel 1 Dependent Var: Conservative STEM Majors</b>   |                       |                       |                       |                       |
| Treated for the first X terms                            | 0.0074<br>(0.0140)    | 0.0112<br>(0.0142)    | 0.0359**<br>(0.0151)  | 0.0412***<br>(0.0154) |
| <b>Panel 2 Dependent Var: OPT Designated STEM Majors</b> |                       |                       |                       |                       |
| Treated for the first X terms                            | 0.0406***<br>(0.0153) | 0.0425***<br>(0.0155) | 0.0745***<br>(0.0165) | 0.0817***<br>(0.0169) |
| Observations   | 8196                  | 8007                  | 7412                  | 7171                  |

Note: This table shows the policy's effect on students' probability of obtaining a STEM degree. This sample include cohorts who first entered BSU between 1990 and 2008 and graduated from BSU. The outcome variable in all columns is an indicator of obtaining a STEM degree, where the definition of a STEM degree is conservative (panel 1) and from ICE (panel 2). The variable of interest for columns (1)-(4) are: continuously treated by the policy for the first 1 semester, for the first 2 semesters, for the first 3 semesters, and for the first 4 semesters, where the corresponding control group for each treatment group in columns (1)-(4) are *never* treated by the policy in the first 1 semester, in the first 2 semesters, in the first 3 semesters, and in the first 4 semesters. Each regression controls for individual's average SAT score, gender, a linear trend of the entry year, a squared trend of the entry year, and a dummy of the entry term season (spring, summer, or fall). Standard errors are clustered at the entry cohort level.

# A Proof of the Theoretical Implication

## A.1 Difficulty Choice

To solve the optimal choice under each policy, we first expand the utility functions and then take the first order condition with respect to  $d$ , for each utility function under each policy, assuming that  $t$  is given and the expected grade of the second attempt (retake),  $E[g']$ , and the expected cost for retaking,  $E[c']$ , are constant. We denote  $g(d, t)$  as  $g$  for brevity:

Take the first order condition (FOC) with respect to (w.r.t.)  $d$ , for policy = 0:

$$\begin{aligned} \frac{\partial U}{\partial d} = & f(g) * \left\{ \frac{E[g'] + g}{2} - c'(d) + \frac{dE[g']}{2} \right\} + f'(g)g'(d) \left\{ d * \frac{E[g'] + g}{2} - c(d, t) - E[c'] \right\} \\ & - f'(g)g'(d)[d * g(d, t) - c(d, t)] + (1 - f(g))[g + d * g'(d) - c'(d)] = 0 \quad (5) \end{aligned}$$

Denote the optimal difficulty as  $d = d_0^*$ . Simplify the equation to get:

$$\begin{aligned} \frac{\partial U}{\partial d} = & \frac{1}{2} \{ f'(g)g'(d)dE[g'] - f'(g)g'(d)dg + f(g)E[g'] - f(g)g - f(g)dg'(d) \} \\ & - \{ f'(g)g'(d)E[c'] - g - dg'(d) + c'(d) \} = 0 \quad (6) \end{aligned}$$

FOC w.r.t.  $d$ , for policy = 1:

$$\begin{aligned} \frac{\partial U}{\partial d} = & f(g) * \{ E[g'] - c'(d) \} + f'(g)g'(d) \{ dE[g'] - c - E[c'] \} \\ & - f'(g)g'(d)[dg - c] + (1 - f(g)) \{ g + dg'(d) - c'(d) \} = 0 \quad (7) \end{aligned}$$

where the optimal  $d = d_1^*$ . Simplify the equation to get:

$$\begin{aligned} \frac{\partial U}{\partial d} = \{f'(g)g'(d)dE[g'] - f'(g)g'(d)dg + f(g)E[g'] - f(g)g - f(g)dg'(d)\} \\ - \{f'(g)g'(d)E[c'] - g - dg'(d) + c'(d)\} = 0 \quad (8) \end{aligned}$$

We can re-write the two FOCs above as follows:

$$\frac{1}{2}A(d_0^*) - B(d_0^*) = 0;$$

$$A(d_1^*) - B(d_1^*) = 0;$$

where  $A() = \{f'(g)g'(d)dE[g'] - f'(g)g'(d)dg + f(g)E[g'] - f(g)g - f(g)dg'(d)\}$  and  $B() = \{f'(g)g'(d)E[c'] - g - dg'(d) + c'(d)\}$ . By assumption,  $g'(d) < 0$ ,  $g''(d) < 0$ ,  $f'(g) < 0$ , and  $c'(d) > 0$ . Again, we assume the expected grade is larger than the first attempted grade, i.e.,  $E[g'] > g(d)$ , for a student to be willing to repeat a course. We can derive that function  $A()$  is strictly decreasing in  $d$  and function  $B()$  is strictly increasing in  $d$ . Finally, we obtain  $d_1^* > d_0^*$ . Therefore, student will be more likely to choose a difficult course under the Grade Forgiveness policy.

## A.2 Time Allocation

Students choose to allocate time among courses they enroll in a semester. Each student has a certain endowment of time to study,  $t$ , which does not change by the policy. To illustrate the time allocation among courses taken in the same semester, let's consider the simplest case that a student takes two courses in this period and allocates time on each course,  $t_1$  and  $t_2$ , i.e.,  $t_1 + t_2 = 1$ . The grading policy as mentioned above,  $g(d, t)$ , is concave and twice continuously differentiable. The cost of taking each course is  $c(d, t)$  is also concave and twice continuously differentiable. The utility function can be written as the summation of the utility gained from the two courses:

For policy = 0:

$$\begin{aligned}
U(g, c, d, t) = & f(g_1) * \left\{ d_1 \frac{E[g'_1] + g(d_1, t_1)}{2} - c(d_1, t_1) - E[c'] \right\} \\
& + \{1 - f(g_1)\} * \{d_1 g(d_1, t_1) - c(d_1, t_1)\} \\
& + f(g_2) * \left\{ d_2 \frac{E[g'_2] + g(d_2, t_2)}{2} - c(d_2, t_2) - E[c'] \right\} \\
& + \{1 - f(g_2)\} * \{d_2 g(d_2, t_2) - c(d_2, t_2)\} \quad (9)
\end{aligned}$$

For policy = 1:

$$\begin{aligned}
U(g, c, d, t) = & f(g_1) * \{dE[g'_1] - c(d_1, t_1) - E[c']\} + \{1 - f(g_1)\} * \{d_1 g(d_1, t_1) - c(d_1, t_1)\} \\
& + f(g_2) * \{dE[g'_2] - c(d_2, t_2) - E[c']\} + \{1 - f(g_2)\} * \{d_2 g(d_2, t_2) - c(d_2, t_2)\} \quad (10)
\end{aligned}$$

For a student who chooses the optimal time spent on course 1,  $t_1$ , we take the first order condition of the utility function under each policy with respect to  $t_1$ . By assumption,  $t_2 = t - t_1$ . We obtain the FOC as:

For policy = 0, FOC w.r.t.  $t_1$ ,

$$\begin{aligned}
\frac{\partial U}{\partial t_1} = & f(g_1) * \left\{ d_1 \frac{g'(t_1)}{2} - c'(t_1) \right\} + \{1 - f(g_1)\} * \{d_1 g'(t_1) - c'(t_1)\} \\
& + f'(g) g'(t_1) * \left\{ d_1 \frac{E[g_1] + g(d_1, t_1)}{2} - c(d_1, t_1) - E[c'] \right\} \\
& + \{-f'(g) g'(t_1)\} * \{d_1 g(d_1, t_1) - c(d_1, t_1)\} \\
& + f(g_2) * \left\{ d_2 \frac{-g'(t_1)}{2} + c'(t_1) \right\} + \{1 - f(g_2)\} * \{-d_2 g'(t_1) + c'(t_1)\} \quad (11) \\
& - f'(g) g'(t_1) * \left\{ d_2 \frac{E[g_2] + g(d_2, t_1)}{2} - c(d_2, t - t_1) - E[c'] \right\} \\
& + \{f'(g) g'(t_1)\} * \{d_2 g(d_2, t - t_1) - c(d_2, t - t_1)\} \\
= & 0
\end{aligned}$$

Simplify to be:

$$\begin{aligned}
\frac{\partial U}{\partial t_1} &= \left\{ -d_1 \frac{f(g_1)g'(t_1)}{2} + d_1 g'(t_1) \right\} \\
&\quad + \left\{ d_1/2 * f'(g)g'(t_1)E[g'_1] - d_2/2 * f'(g)g'(t_1)g(d_1, t_1) \right\} \\
&\quad \left\{ +d_2 \frac{f(g_2)g'(t_1)}{2} - d_2 g'(t_1) \right\} \\
&\quad + \left\{ -d_2/2 * f'(g)g'(t_1)E[g'_2] + d_2/2 * f'(g)g'(t_1)g(d_2, t_1) \right\} \\
&= 0
\end{aligned} \tag{12}$$

For policy = 1, FOC w.r.t.  $t_1$ ,

$$\begin{aligned}
\frac{\partial U}{\partial t_1} &= -f(g_1) * \{c'(t_1)\} + \{1 - f(g_1)\} * \{d_1 g'(t_1) - c'(t_1)\} \\
&\quad + f'(g)g'(t_1) * \{d_1 E[g'_1] - c(d_1, t_1) - E[c']\} \\
&\quad + \{-f'(g)g'(t_1)\} * \{d_1 g(d_1, t_1) - c(d_1, t_1)\} \\
&\quad + f(g_2) * \{c'(t_1)\} + \{1 - f(g_2)\} * \{-d_2 g'(t_1) + c'(t_1)\} \\
&\quad - f'(g)g'(t_1) * \{d_2 E[g'_2] - c(d_2, t - t_1) - E[c']\} \\
&\quad + \{f'(g)g'(t_1)\} * \{d_2 g(d_2, t - t_1) - c(d_2, t - t_1)\} \\
&= 0
\end{aligned} \tag{13}$$

Simplify to be:

$$\begin{aligned}
\frac{\partial U}{\partial t_1} &= \left\{ -d_1 f(g_1)g'(t_1) + d_1 g'(t_1) \right\} + \left\{ d_1 * f'(g)g'(t_1)E[g'_1] - d_2 * f'(g)g'(t_1)g(d_1, t_1) \right\} \\
&\quad \left\{ +d_2 f(g_2)g'(t_1) - d_2 g'(t_1) \right\} + \left\{ -d_2 * f'(g)g'(t_1)E[g'_2] + d_2/2 * f'(g)g'(t_1)g(d_2, t_1) \right\} = 0
\end{aligned} \tag{14}$$

Re-arranging (12) and (14) we obtain:

$$\begin{aligned}
\frac{1}{2} \left\{ -d_1 f(g_1)g'(t_1) + d_1 f'(g)g'(t_1)E[g'_1] - d_1 f'(g)g'(t_1)g(d_1, t_1) + d_2 f(g_2)g'(t_1) \right. \\
\left. - d_2 f'(g)g'(t_1)E[g'_2] + d_2 * f'(g)g'(t_1)g(d_2, t_2) \right\} = g'(t_1)[d_2 - d_1] \tag{15}
\end{aligned}$$

and

$$\{-d_1f(g_1)g'(t_1) + d_1f'(g)g'(t_1)E[g'_1] - d_1f'(g)g'(t_1)g(d_1, t_1) + d_2f(g_2)g'(t_1) - d_2f'(g)g'(t_1)E[g'_2] + d_2f'(g)g'(t_1)g(d_2, t_2)\} = g'(t_1)[d_2 - d_1] \quad (16)$$

Compare (15) and (16), we can find  $\frac{1}{2}LHS(t_1^A) = g'(t_1^A)[d_1 - d_2]$  and  $LHS(t_1^F) = g'(t_1^F)[d_1 - d_2]$ , where  $t_1^A$  and  $t_1^F$  are the optimal  $t_1$  choice under the averaging policy and the Forgiveness policy. By assumption,  $d_1 > d_2$ ,  $g' > 0$  and  $g'' < 0$ . Thus  $g'(t_1)[d_1 - d_2]$  is increasing in  $t_1$  and function  $LHS(t_1)$  is decreasing in  $t_1$ . Finally, we can conjecture  $t_1^A < t_1^F$ . The gap between the time allocated in the two courses:  $t_1 - t_2 = t_1 - (t - t_1) = 2t_1 - t$  increases as  $t_1$  increases.

### A.3 Probability of Repeating and the Threshold

We now assume that a student has already enrolled and completed a course and realized the grade,  $g$ , and the cost,  $c$ , at the first-attempt and has a constant expectation on the second-attempt grade,  $E[g']$ , and a constant expectation on the second-attempt cost,  $E[c']$ . The "Grade Forgiveness" policy basically changes the calculation of the GPA if the student chooses to repeat. That is, when repeating happens during the time that policy is off, i.e., Forgiveness = 0, the final grade is simply an average between the first-attempt grade and the second-attempt grade. When repeating happens during the time that policy is on, i.e., Forgiveness = 1, the final grade is the second-attempt grade. Here, we assume the difficulty and time allocation as given for brevity to derive the probability of repeating.

$$U(g, c) = \begin{cases} g - c, & \text{if Not Repeat} \\ \frac{E[g'] + g}{2} - c - E[c'] & \text{if Repeat under Forgiveness = 0} \\ E[g'] - c - E[c'], & \text{if Repeat under Forgiveness = 1} \end{cases} \quad (17)$$

The condition for the student chooses to repeat is the utility of repeating is larger than the utility of not repeating. Let's denote the utility of repeating as  $U^R$  and the



utility of not repeating as  $U^{NR}$ .

$$U^R(g, c) > U^{NR}(g, c) \equiv \begin{cases} \frac{E[g'] + g}{2} - c - E[c'] > g - c, & \text{if Forgiveness} = 0 \\ E[g'] - c - E[c'] > g - c, & \text{if Forgiveness} = 1 \end{cases} \quad (18)$$

We further denote the realized first-attempt grade under each policy as  $g_0$  and  $g_1$  and simplify the inequalities above as:

$$U^R(g, c) > U^{NR}(g, c) \equiv \begin{cases} E[g'] - 2E[c'] > g_0, & \text{if Forgiveness} = 0 \\ E[g'] - E[c'] > g_1, & \text{if Forgiveness} = 1 \end{cases} \quad (19)$$

Thus the probability of repeating a course can be written as:

$Pr(E[g'] - 2E[c'] > g_0)$  and  $Pr(E[g'] - 2E[c'] = g_0) = 0$  if Forgiveness = 0;

$Pr(E[g'] - E[c'] > g_1)$  and  $Pr(E[g'] - E[c'] = g_1) = 0$  if Forgiveness = 1.

By assumption, the belief on the expected grade,  $E[g']$ , and the expected cost,  $E[c']$ , is unchanged. The probability of repeating a course is determined by the inequality. We can easily obtain that:

(1) If the first-attempt grades are constant under different policies,  $g_0 = g_1 = g$ , the region of grades for one to prefer repeating over not is larger under the policy Forgiveness = 1 than under the policy Forgiveness = 0:  $E[c'] < E[g'] - g < 2E[c']$ . Thus, the average probability of repeating under Forgiveness = 1 will be higher than the probability of repeating under Forgiveness = 0.

(2) The threshold (highest) grade to repeat under Forgiveness = 1 is higher than the threshold (highest) grade to repeat under Forgiveness = 0:  $g_1 > g_0$ , and the difference between the two threshold grades is restricted as  $g_1 - g_0 \leq E[c']$ .

## A.4 Persistence

Following the same logic as previous sections, the GPA calculation formula under the "Grade Forgiveness" policy will derive higher cumulative GPA than the "av-

eraging" policy. Assuming the repeated grade is no lower than the first attempted grade, in period  $t - 1$ ,  $GPA_{t-1}|policy = F \geq GPA_{t-1}|policy = A$ . Similarly, the future expected grades, regardless of ever repeating any courses in the future, under the forgiveness policy is greater or equal to the expected grades under the averaging policy, i.e.,  $E_t^F[g] \geq E_t^A[g]$  while the expected cost should be unchanged, i.e.,  $E_t^F[c] = E_t^A[c]$ .

Based the above, comparing the LHS of the inequality,  $U(GPA_{t-1}, E_t[g], E_t[c] | policy = F) \leq U(GPA_{t-1}, E_t[g], E_t[c] | policy = A) \Leftrightarrow$

$Pr[U(GPA_{t-1}, E_t[g], E_t[c] | policy = F) > u^*] > Pr[U(GPA_{t-1}, E_t[g], E_t[c] | policy = A) > u^*]$ . The probability of persisting in the current subject or major is higher under the forgiveness policy.

## B Robustness Checks

### B.1 Policy's Effect on Repeat: turn-on versus turn-off

We exploited two variations of the grade replacement policy in our main analysis. Here, we further investigate the variation-specific effects regarding the policy's turn-on and turn-off by showing the effects on two cohort windows: Fall 1990–Summer 2001 and Fall 1995–Summer 2006. We run the same regression specification in columns (3) and (4) of Table 2 on the two sub-samples. The results in Table B1 show that the probability of repeating is significantly higher when the policy is in force and it holds true for both the cohorts who experienced a policy turn-off and the cohorts who experienced a policy turn-on. This evidence provide convincing support for our identification strategy that the policy's effects are not subject to specific time or cohort-sensitive.

Table B1: Policy's Effect on Repeat: turn-on versus turn-off

|                          | Policy on-off         |                       | Policy off-on         |                       |
|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                          | Fall 1990–Summer 2001 | Fall 1995–Summer 2006 | Fall 1995–Summer 2006 | Fall 1990–Summer 2001 |
|                          | (1)                   | (2)                   | (3)                   | (4)                   |
| Policy                   | 0.0213***<br>(0.0012) | 0.0211***<br>(0.0017) | 0.0174***<br>(0.0013) | 0.0205***<br>(0.0019) |
| Include Transfer Student | YES                   | NO                    | YES                   | NO                    |
| Observations             | 756928                | 421734                | 1018966               | 517085                |

Note: This table shows the policy's effect on the probability of repeating by two cohorts: cohorts under the first change of the policy (policy on to off period) and cohorts under the second change of the policy (policy off to on period). The regression sample in each column are sub-samples from regression sample in column (3) and (4) of Table 2. Specifically, column (1) and (2) covers Fall 1990 to Summer 2001 observations and columns (3) and (4) cover Fall 1995–Summer 2006 observations. The independent variable is an indicator of whether a course shown on a student's transcript is a repeat (second time taking). Each regression includes individual fixed effects, academic term  $t$  fixed effects, course fixed effects, number of credits attempted in the semester, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.

### B.2 Curriculum Choice, excluding the Repeats

It is debatable that whether one should also count the increased repeats on challenging courses—second or higher attempts—as students choosing more difficult

courses. Based on our findings on repeaters' subsequent choices and students persistence in STEM major in terms of increased graduation rate, we believe that it is reasonable to count the repeated courses in just as first-attempt courses. Nevertheless, we show the curriculum choice on the first-attempt only sample in the following table. These estimates imply that, the policy incentives about 8% more first-enrollment in STEM courses.

Table B2: Policy's Effect on STEM Course Enrollment, excluding Repeats

|   | (1)                   | (2)                   | (3)                   |
|---|-----------------------|-----------------------|-----------------------|
|   | STEM Course           | STEM Course           | STEM Course           |
| <b>Panel 1: Course-level STEM (conservative definition)</b> |                       |                       |                       |
| policy  | 0.0146***<br>(0.0017) | 0.0189***<br>(0.0023) | 0.0180***<br>(0.0050) |
| Sample  | All                   | Non-transfer          | Undeclared Major      |
| Sample Mean   | .2457317              | .2390266              | .2232875              |
| N   | 2612345               | 1231997               | 212045                |
| <b>Panel 2: Course-level STEM (OPT definition)</b>          |                       |                       |                       |
| policy  | 0.0169***<br>(0.0019) | 0.0209***<br>(0.0025) | 0.0202***<br>(0.0053) |
| Sample  | All                   | Non-transfer          | Undeclared Major      |
| Sample Mean   | .3007367              | .2976696              | .295541               |
| N   | 2612345               | 1231997               | 212045                |

This table is an analogue to Table 4 but the estimation samples exclude all non-first attempts (repeats) for robustness check purpose.

Table B3: Policy's Effect on Choice of Difficulty, Alternative Measures

|              | (1)        | (2)        | (3)         | (4)            |
|--------------|------------|------------|-------------|----------------|
|              | Measure 1  | Measure 2  | Measure 3   | Measure 4      |
|              | %AB        | Grades     | %AB×Credits | Grades×Credits |
| Policy       | -0.0127*** | -0.0333*** | -0.0247***  | -0.0446***     |
|              | (0.0006)   | (0.0017)   | (0.0024)    | (0.0081)       |
| Sample Mean  | 0.7715     | 2.9149     | 2.2527      | 8.5344         |
| Observations | 2707279    | 2707279    | 2707279     | 2707279        |

Note: This table shows robustness checks on Table 5 by showing alternative measures of course difficulty. The independent variable in each column are measures of difficulty described in depth in section 6.2. The outcome variables in columns (1)–(4) are course-level difficulty measured by fraction of letter grade A or B by course, average grade points by course, the product of the fraction of letter grade A or B by course and the course credits, and the product of the average grade points by course and the course credits. Each regression includes individual fixed effects, academic term  $t$  fixed effects, and a policy indicator where policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.

## C Heterogeneity: Policy's Nudge across Academic Preparation

We have also explored how students with different levels of academic preparation respond to the policy in terms of choosing more difficult courses. We conduct a similar analysis as in Table 4 by differentiating students by their first term GPA. Specifically, we regress the average course difficulty and number of credits attempted on the interaction of policy and the quartiles of the first term GPA. We use the first term GPA as a measure of students' academic preparation instead of using SAT or ACT scores because there are % student do not have a SAT or ACT score on file.

Table C1 shows the results for this exercise. Quartile 1–4 represent students with lowest–highest first term GPA. Overall, there is no significant differences in the policy's effect on taking a STEM course across the 4 groups of students. However, by including only the first 6 semesters (similar results for only including the first 4 semesters), when students are still exploring curriculum and when course repeats happen most frequently, we see that the policy's effect is driven by students who are better prepared. In specific, students in quartiles 2, 3, and 4 have statistically significant effect on choosing a STEM course than students in quartile 1. The magnitudes increases slightly when GPA increases, but not statistically significant. Additionally, students whose GPA are in quantile 4 are driving the increased number of credits per course. Overall, it is evident that the policy's nudge is especially pronounced among students who have stronger academic preparation.

Table C1: Heterogeneity: Policy's Nudge across Academic Preparation

|                     | (1)                   | (2)                   | (3)                 |
|---------------------|-----------------------|-----------------------|---------------------|
|                     | STEM Conservative     | STEM OPT              | #Credits            |
| Policy              | -0.0062<br>(0.0046)   | 0.0012<br>(0.0050)    | 0.0019<br>(0.0093)  |
| Policy × Quartile 2 | 0.0249***<br>(0.0068) | 0.0205***<br>(0.0073) | 0.0193<br>(0.0130)  |
| Policy × Quartile 3 | 0.0254***<br>(0.0085) | 0.0178**<br>(0.0091)  | 0.0139<br>(0.0158)  |
| Policy × Quartile 4 | 0.0244***<br>(0.0087) | 0.0207**<br>(0.0093)  | 0.0328*<br>(0.0172) |
| <i>N</i>            | 873210                | 873210                | 873210              |

Note: This table shows the policy's different effect on course choice by academic preparation. The sample include the first 6 semesters of all non-transfer students. The independent variable in column (1) is an indicator of STEM course (conservative definition), column (2) is an indicator of STEM course (OPT definition), column (3) is the number of credits of each course. Student's academic preparation is defined by their first term GPA, and separated into quartiles, where quartile 1 with the lowest GPA and quartile 4 with the highest GPA. Each regression includes individual fixed effects, academic term  $t$  fixed effects, number of courses offered in each semester, number of peers in each semester and each major, a policy indicator and an interaction of policy and the quartile. Policy = 1 for calendar semesters Spring 1990–Summer 1995 and Fall 2001 and after; policy = 0 elsewhere. Standard errors are clustered at the individual level.