NBER WORKING PAPER SERIES

A SECOND CHANCE AT SUCCESS:
CAN GRADE FORGIVENESS PROMOTE ACADEMIC RISK-TAKING IN COLLEGE?

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, Revised September 2023

We thank Bruce Weinberg, Audrey Light, Valerie Bostwick, and seminar and conference participants at Ohio State University, Purdue University, Jinan University, Boise State University, SOLE, AEFP, APPAM, 7th IZA Workshop on the Economics of Education, and the 2021 Symposium on Contemporary Labor Economics. All errors are our own. 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.

A Second Chance at Success: Can Grade Forgiveness Promote Academic Risk-Taking in College?
Xuan Jiang, Kelly Chen, Zeynep K. Hansen, and Scott Lowe
NBER Working Paper No. 29493
November 2021, Revised September 2023
JEL No. I21,I23

ABSTRACT

The increased popularity of college grade forgiveness policies, which allow students to substitute grades for repeated courses in their grade-point-average calculations, has been regarded as a consequence of the pressure colleges feel to ensure their "customers" are satisfied. However, this study identifies an important benefit that grade forgiveness confers on students: more risk-taking in the learning process. Using longitudinal administrative data from a four-year public institution that alternated between two grading schemes over a short period of seven years, we find that the adoption of the grade forgiveness policy, over the traditional practice of grade averaging, nudges students to pursue curriculum and/or degrees perceived as relatively more challenging and/or with harder grading standards, such as those in Science, Technology, Engineering, and Mathematics. This result holds true for first-time course-takers as well as for students who do not repeat any courses while in college. Finally, while helping students achieve ultimate mastery, we find no evidence that grade forgiveness delays graduation or elicits spending less effort by students.

Xuan Jiang
Department of Economics
Jinan University
School of Economics
Guangzhou 510000
China
jiangxuan@jnu.edu.cn

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

Zeynep K. Hansen
College of Business and Economics
Department of Economics
Boise State University
1910 University Drive
Boise, ID 83725-1620
and NBER
zeynephansen@boisestate.edu

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

1 Introduction

Academic risk-taking has been identified as a key driver of student motivation and cognitive development and is viewed as an integral part of student learning in the educational psychology literature (Clifford, 1991; Clifford and Chou, 1991; Sewell and George, 2000; Masten and Obradović, 2006; Henriksen, Henderson, Creely, Carvalho, Cernochova, Dash, Davis, and Mishra, 2021). In practice, however, very little is known about how risk-taking, generally perceived as the determination of students to strive against the difficulties they face during the learning process, can be effectively encouraged among students, especially in a post-secondary setting. An emerging body of experimental studies focuses on the pedagogical strategies that facilitate productive failures within a classroom,[1] but there is a lack of evidence on the approaches that apply to the general population of students in the real world.

This study explores a non-punitive grade-point-average (GPA) policy as a promising avenue to cultivate intellectual risk-taking for college students across different stages of study and disciplines. To our best knowledge, this is the first rigorous, quasi-experimental evidence to date of how non-punitive grading practices may affect student outcomes. By examining how student choices of curriculum and major may be influenced by the adoption of a grade forgiveness formula in the calculation of cumulative GPA, we demonstrate that a small deviation from the traditional practice of grade averaging can make a meaningful difference by incentivizing them to enroll, persist, and succeed in subjects that are perceived as challenging or have higher grading standards.

In contrast with the grade averaging scheme, which takes the average of the new and old grades, the grade forgiveness option applies the most recent grade to a student's cumulative GPA calculation, should a course be repeated. As such, it provides an opportunity for students, especially first-time course-takers, to try things out and on occasion fail without being perpetually penalized by a lower

[1] See Clifford (1991) for a comprehensive review.

overall GPA. Given the vital role of the cumulative GPA in student eligibility for graduation, honors programs, scholarships/loans, and other academic opportunities (e.g., acceptance into a graduate program after college), the ability to erase an early misstep from one's track record of academic success constitutes a powerful incentive for these students to re-evaluate the expected benefit/cost of error-making. This could be particularly helpful for students who are on the margin of pursuing courses or degrees that are perceived as difficult, but that may lead to high-paying jobs, such as those in Science, Technology, Engineering, and Mathematics (STEM). By encouraging these students to be more accepting of loss or a lower grade as part of their search/exploration process (Arcidiacono, 2004; Stinebrickner and Stinebrickner, 2013), grade forgiveness can affect these students' propensities to experiment with different subject areas and, as a result, may have a substantial impact on the distribution of desirable career outcomes.

While helping students to maintain their path, a major criticism of the grade forgiveness policy is that it may not always support a student's desired degree timeline. Existing studies contend that by diminishing the importance of course grades, non-punitive grading practices such as grade forgiveness may make students less inclined to put forth the level of effort needed to succeed in a course the first time. If the lack of student accountability and/or excessive course repetition cause significant disruptions to a student's routine progress in pursuing a degree, grade forgiveness may negatively impact student outcomes such as college completion and time-to-degree (Jewell and Tieslau, 2013; Marx and Meeler, 2013). Furthermore, given that not all students can afford to take a dramatically longer route to achieve the required skills, grade forgiveness may place an undue burden on those who face time and resource constraints and thereby exacerbate existing social and economic inequalities (Marx and Meeler, 2013). Finally, allowing students to repeat courses may generate a bottleneck in certain degrees and disrupt the flow of students into and out of a university, which could further raise equity and equal opportunity issues when there is limited access to such courses or programs (Casas and Meaghan, 1996). All of these arguments are theoretically sensible and are ultimately empirical questions that prior work has not fully answered.

To this end, we offer an initial step towards a better understanding of this debate by showing – both theoretically and empirically – that grade forgiveness does not necessarily lead to a decline in student effort or worse graduation outcomes. As illustrated in our theoretical model, a student who values both the type of and the expected grade received from a given course will pursue a more difficult curriculum when having a safety net to fall back on, despite allocating less study time to each course. Thus, after a student's choice of course difficulty is factored into the equation, it is not a priori clear whether grade forgiveness will induce students to put forward less effort on the aggregate.

Our empirical investigation, using administrative longitudinal data from Boise State University a four-year public institution, further tests these theoretical predictions. This data set encompasses the universe of first-time (i.e., non-transfer) undergraduate students who attended the institution over a 27-year period from 1990-2016. A marked advantage of the Boise State data is that the university alternated between two weighting schemes in calculating repeating students' cumulative GPAs within a short period of seven years: grade forgiveness schemes before 1995 and after 2001, and a grade averaging scheme in between. This unique institutional feature allows us to separately identify the effect of grade forgiveness from its cancellation and reinstatement, and to tease out the role of time-varying confounders in a non-parametric fashion. Simultaneously, we are able to compare the changes in outcomes across students with different propensities to be treated (i.e., students enrolled in graded versus Pass/Fail courses) and/or different treatment intensities (i.e., entry cohorts exposed to the policy for different lengths of time) in a difference-in-differences (DD) framework.

Constructing two measures for the difficulty level of a given course, one based on the objective subject matter (i.e., STEM designation) and another based on perceived grading harshness, we show that the adoption of grade forgiveness increases the likelihood of enrolling in a STEM course by 1-3 percentage points (or 4-10% from the baseline) or a course with a lower historical grading average by 0.1 points (or 1% relative to the baseline) for never-repeating students, who did not repeat any courses or experience any GPA increase associated with the grade forgiveness scheme during our observation period. Treating the former, more objective metric as the preferred measure for academic difficulty, we estimate that the increase in the initial student interest and persistence translates into an average of 8 percentage points or 23% increase in STEM graduation, while having no impact on the average time required to earn the degree. Given that our most restrictive sample (i.e., never-repeating students) omits repeating students who may also have benefited from the grade forgiveness policy, we consider the above figures as lower-bound estimates for the risk-taking effects of grade forgiveness.

Even within the most stringent sample selection criterion, we find evidence that the curriculum choice effect of grade forgiveness is driven by the students who express little initial interest in STEM subjects upon college entry (i.e., non-STEM majors), as measured by students who begin college with a major in a non-STEM discipline, as "undeclared," or "undecided." In this sample, the students of different genders (i.e., female vs male), socioeconomic backgrounds (i.e., low vs high income), and races (i.e., white vs non-white) appear to be similarly affected. Potentially related to the switching of their declared majors, we find in our mechanism investigation that grade forgiveness induces this group of students to take on a heavier semester course load while attempting more challenging courses. Correspondingly, once the student choice of study pace is introduced into our regressions, we find no evidence that grade forgiveness elicits less effort by students. In other words, if we additionally consider the ambitious nature of their academic schedules, students may put forth more effort under grade forgiveness when compared to the grade averaging scheme, which is consistent with the existing psychology theory and experimental evidence that teaching students to take risks can in fact increase the effort they spend on academics (Clifford, 1991).

The paper proceeds as follows: Section 2 provides a brief review of related literature. Section 3 proposes a theoretical framework that guides the empirical analysis that follows. Sections 4 and 5 describe the institutional context, data, methods, and main results. Section 6 explores heterogeneity in the estimated effects of grade forgiveness and Section 7 concludes.

2 Literature Review

This study is closely related to the strand of literature on the effect of instructor grading standards on student outcomes, such as test scores (Betts and Grogger, 2003; Figlio and Lucas, 2004; Bonesrønning, 2004), disciplinary problems (Figlio and Lucas, 2004), subsequent interest (Sabot and Wakeman-Linn, 1991; Fournier and Sass, 2000; Chen, Hansen, and Lowe, 2021), and future earnings in the labor market (Betts and Grogger, 2003). Instead of focusing on the grade assignment for a given level of achievement, we investigate how the measurement of achievement in grade assignment (i.e., whether a repetition penalty should be imposed) may affect student interest and success.

Furthermore, while the authority and responsibility of assigning grades primarily lie with course instructors, institutions determine which grades are included in the composite and how the average is calculated through academic policies such as course withdrawal, course repeats, and Pass/Fail grading options (Marx and Meeler, 2013). From this viewpoint, we deviate from the aforementioned studies by illustrating how the grading policies measured at the institution level may affect students. In particular, we extend earlier studies on college major choice who find that equalizing grading standards between STEM and non-STEM courses could improve participation (Ahn, Arcidiacono, Hopson, and Thomas, 2019; Minaya, 2020) by identifying an alternative solution and provide supportive evidence on grades as an important determinant of major choice and major switching (Astorne-Figari and Speer, 2019; Stinebrickner and Stinebrickner, 2013), among the many other factors established in the literature, including financial incentive, ability, preference, and peer effects (Speer, 2017; Denning and Turley, 2017; Card and Payne, 2021; Jiang, 2021; Blume-Kohout and Scott, 2022; Bostwick and Weinberg, 2022). In countries such as the U.S. where individual faculty members have a good deal of autonomy in how they grade (Dickson, 1984; Freeman, 1999) and an era when course grades may operate as a policy instrument to influence student enrollment across departments (Achen and Courant, 2009), institutional GPA policies may represent an effective option to increase student interest in relevant areas.

Given that the grade forgiveness policy directly affects students by providing a higher perceived return to course repetition, we also contribute to a body of work that causally estimates the learning gains through the various forms of repetition, including mandatory programs such as remedial education (Bettinger and Long, 2009; De Paola and Scoppa, 2014), grade retention (Tafreschi and Thiemann, 2016), and voluntary course repetition (Chen and Jiang, 2023), as well as the repetition of placement exams prior to college (Vigdor and Clotfelter, 2003; Frisancho, Krishna, Lychagin, and Yavas, 2016; Goodman, Gurantz, and Smith, 2020). For example, Bettinger and Long (2009) find that under-prepared students who are required to take below-college level courses or are placed in remediation persist longer in college in comparison to students with similar backgrounds and preparation who are not required to take the courses. Examining first-year undergraduates who are assigned to a grade retention program and are mandated to repeat all first-year courses in Switzerland, Tafreschi and Thiemann (2016) report a large and persistent positive effect of retention on student performance, along with a moderate adverse effect on student dropout. For students not in any remedial or grade retention programs, Chen and Jiang (2023), using the same data as that of the current study, argue that course repetition through voluntary decisions results in better educational outcomes for repeating students. Finally, at the pre-college level, Vigdor and Clotfelter (2003); Frisancho et al. (2016); Goodman et al. (2020) find that re-taking high-stakes college entrance exams substantially improve student scores and increases four-year college enrollment rates, particularly for low-income and underrepresented minority students. Extending these studies that focus on repeating students' behaviors, we postulate a spillover or indirect effect of repetition on non-repeating students who are exposed to the repetition policy but do not necessarily take advantage of it. Based on our knowledge, no work thus far has explored the effect of course repetition on the outcomes of non-repeaters among college students.[2]

3 Conceptual Framework

In theory, the policy alteration from grade averaging to grade forgiveness can affect student behaviors in two different ways: directly, by inducing them to repeat more courses, and indirectly by increasing their tolerance for failures and risks. For our purposes, this section lays out a conceptual framework to illustrate the latter, or the risk-taking effect of grade forgiveness. We are particularly interested in students' decisions with regard to course choices and the time allocation both to a single course and across the multiple courses attempted within a semester. Our end goal is to derive a set of testable predictions to guide our empirical analysis and also provide insight into the mechanism through which grade forgiveness may benefit/hurt students in the aforementioned dimensions.

We consider an environment where students make two choices to maximize their single-period utility, prior to enrolling in a given course: the type of the course and study time. In reality, the type of course that a student pursues is jointly determined by the difficulty level (i.e., quality) and academic load or credit hours (i.e., quantity) of the course. However, here we do not distinguish between these two. Instead, we use a one-dimensional difficulty definition, d, to represent the student's course choice for the sake of simplicity. While being left out of our theoretical model, we are able to separately identify the effects of grade forgiveness on student choices of course quality and quantity, respectively, in the empirical analysis and

[2] Hill (2014) investigates the extent to which course repeaters in high school mathematics courses exert negative externalities on their course-mates.

comment on their relationship with each other (see Section 4.2).

We assume that students value both the difficulty level d and expected initial-attempt grade g of the course while incurring a certain amount of (time and mental) cost during the learning process c.[3] Both the expected initial-attempt grade g and learning cost c are a function of course difficulty d and study time t. In particular, we assume both of them are twice continuously differentiable and follow the law of diminishing marginal returns: $g'(d) < 0$, $g''(d) < 0$, $g'(t) > 0$, $g''(t) < 0$, $c'(d) > 0$, $c''(d) < 0$, $c'(t) > 0$, and $c''(t) < 0$. In other words, the more difficult the course, the lower (higher) the grade (cost) will be received (incurred) and the more time spent on the course, the higher grade (cost) will be received (incurred).

We further assume that the probability of repeating a course is influenced by the expected initial-attempt grade, $f(g)$, and is decreasing in g: $f'(g) < 0$. If this is the case, the utility function of a typical student can be expressed as the sum of the utility from repeating and that from not repeating a given course integrated over $f(g)$:

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

Here g^* denotes the implicit threshold grade or reservation grade for the student to repeat the course. More specifically, the student will choose to repeat the course when g ranges between the lowest grade to the threshold grade, $[F, g^*]$, and will choose not to repeat when g ranges between the threshold grade to the highest grade, $[g^*, A]$. That said, upon the revelation of the initial-attempt grade, the utility for a student who chooses not to repeat a course will be simply determined by the difficulty level and initial-attempt grade of the course, along with the corresponding cost of course-taking:

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

[3]Given that students typically take courses within their expected major(s), d can be interpreted as the student's major choice, as well, even though it is not explicitly introduced in the model.

Meanwhile, the utility for a student who chooses to repeat a course will be determined by the grading scheme that prevails when the initial attempt is made:

$$U^{(\text{repeat})} = \begin{cases} d * \frac{E[G]+g(d,t)}{2} - c(d,t) - E[C], & \text{Under Grade Averaging} \\ d * E[G] - c(d,t) - E[C], & \text{Under Grade Forgiveness} \end{cases}$$

where $E[G]$ and $E[C]$ are the expected subsequent-attempt grade and cost of repeating, respectively. It is worth noting that while the student's subsequent-attempt grade assumes a zero weight (or completely replaces the student's initial-attempt grade) in his/her cumulative GPA calculation, our theoretical implications will remain the same, as long as the weight on the student's subsequent-attempt grade is higher than the initial-attempt grade.

We subsequently derive the following propositions regarding the curriculum choice and time allocation for students who make their initial attempt to a given course, prior to the revelation of the actual course grade. For the sake of completeness, we also derive propositions regarding the decision on course repetition for the students who have observed their initial-attempt grades. These proofs can be found in Appendix A.

3.1 Course Difficulty and Time Allocation to a Given Course

Proposition 1. Students will attempt a more difficult course under grade forgiveness in comparison to the grade averaging scheme.

Proposition 2. Students will allocate less study time or make less effort to a given course when grade forgiveness is enacted relative to grade averaging.

Revealed by these propositions, it is not clear whether grade forgiveness will induce less effort-making after considering their choice of course difficulty. In other words, while students may spend less time studying each course when having the option to improve their grades later, they may also pursue a more challenging curriculum when they have a safety-net to fall back on.

3.2 Time Allocation across Courses

Proposition 3. Assuming that students' total time/effort is binding, then the time allocated among different courses will be more dispersed when the grading scheme is switched from grade averaging to grade forgiveness.

Put differently, if grade forgiveness did not elicit more study time or effort, then we would observe students to allocate time in favor of the course(s) that is more likely to award higher grades and allocate time away from the course(s) that is more likely to award lower grades in a given semester.

3.3 Probability of Repetition

Proposition 4. The average probability of course repetition will be higher under grade forgiveness than grade averaging.

In essence, our theoretical model postulates that by offering students insurance against low grades, the grade forgiveness scheme increases the expected benefit of risk-taking. When the insurance value of using grade forgiveness outweighs its cost, we would expect it to promote risk-taking behavior among students.

4 Institutional Context, Data, and Method

The practice of grade forgiveness has been embraced by an increasing number of U.S. colleges, with varying requirements and qualifications across institutions.[4] Among four-year institutions with 10,000+ enrollments as of 2023 (Integrated Post-Secondary Education Data System), we estimate an over 70 percentage points increase in the likelihood of adopting a grade forgiveness policy over the past five

[4]For example, the grade forgiveness policy may differ in terms of the number of credits/courses that are allowed to be repeated, the number of times a given course can be repeated, the circumstance(s) under which a student is eligible to retake a course (e.g., a grade of D or below), whether permission must be sought in advance to register for a repeated course, and whether a student must petition for a grade to be replaced after the repeated course is successfully completed etc.

decades, from 2% in the 1960s to nearly 80% through the 2020s (Figure 1). To date, 88% of the land grant universities in the U.S. have a grade forgiveness policy in place.

Despite the popularity of grade forgiveness among four-year institutions, very little is known about its impact on the general population of students, in particular, non-repeating students who typically are not the direct beneficiaries of these initiatives. The empirical analysis of this paper sheds light on this issue by drawing evidence from a mid-sized public institution, Boise State University (BSU), which carried out a population-wide implementation and tested the effects of both grading schemes (i.e., grade forgiveness and averaging) in a real-world setting.

4.1 Institutional context

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, it had the largest undergraduate enrollment in the state of Idaho and offered nearly 80 bachelor's degrees across seven colleges: Arts & Sciences, Business & Economics, Education, Engineering, Graduate Studies, Health Sciences, and the School of Public Service.

As one of the earliest adopters of the grade forgiveness policy, prior to 1970, BSU allowed its students who received a grade of D or below to repeat a course and substitute the new grade for the previous grades in their cumulative GPA calculation. This grade limitation of D or below was removed in 1988 and remained in effect until 1995, so that all students, regardless of their programs, academic standing, or initially-attempted grades, can choose to repeat a course for better grades, provided that space was available at the time of repetition during this period. After 1995, in an attempt to "raise academic standards," the university switched to a grading-averaging scheme even though only six years later it reverted back to the grade forgiveness scheme due to perceived fairness to transfer students, given that most other colleges in the state of Idaho had implemented a grade forgiveness policy at the time.[5,6]

Students appeared to be made aware of these policy changes in a timely manner through the annually published University Catalog and the functioning of academic advising across colleges. In addition, both events were well-covered by the campus newspaper – the most popular newspaper among students at the time – the "Arbiter." On January 18, 1995, for example, an article entitled "New academic rules will greet students next fall" informed students of the cancellation of grade forgiveness that would take place in the upcoming academic year.[7] On August 30, 2001, an article entitled "Grade replacement policy takes effect this semester" announced the reinstatement of grade forgiveness and referred it as "a new tool to improve [their] all-important GPAs."[8]

During the observation period, course repetition remained free for full-time students, provided that the students did not enroll in number of credit hours in excess of a full course load per semester. There was a gradual increase in the sticker price for each credit hour for overload and/or part-time students from $61 to $297 with no difference in the fee schedule faced by in-state and out-of-state students. Importantly, with the exception of the weighting scheme used to calculate the cumulative GPA, other parameters of the course repetition policy at BSU were largely unchanged. For example, both new and old grades remained on the students' transcripts throughout the observation period. While an overall maximum of six repeats and a course maximum of 2-3 repeats were imposed at one point to limit the number of courses that a student may repeat and the number of times the student

[5] According to the meeting minutes of the Academic Standard Committee accessible to us, the university believed that the grade averaging scheme "has proven to be unfair to incoming transfer students" since these students took courses at their original institutions in good faith under the grade replacement rules. Hence, the formula would "penalize these students to a greater extent than was first proposed" and "make it difficult for them to raise their GPAs." Relevant documents are not included in the paper but are available upon request.

[6] It is also noteworthy that neither policy was retroactive and therefore both were applicable to course repetitions that occurred subsequently.

[7] See https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1440&context=student_newspapers

[8] See https://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=2194&context=student_newspapers.

may take the same course, many exceptions to the repeat count were allowed.[9,10] Even in the presence of overall and individual repeat maximums, students were granted opportunities to make additional attempts upon special request and "appeals by students of the policy were usually successful," according to BSU's Faculty Senate Meeting Minutes, December 6, 2001 (available upon request). It is also worth noting that despite the lack of regulations on who could repeat courses and under what circumstances a student was allowed to repeat in our context, excessive repetition was rare. For example, only about 7% of the students repeated more than five courses over their entire undergraduate career before the overall maximum of six was enforced in 2013. Approximately 0.05% of the students repeated a given course more than twice before the individual maximum of three was instituted in 1995 and 0.5% repeated a course more than once before the individual maximum of two in 2015.

4.2 Data

Our empirical investigation relies on the admission records and official transcripts of the 75,576 first-time/non-transfer undergraduate students who attended BSU from 1990 through 2016. We consider students who joined BSU after having begun their course of study at a different school and students who spent any time outside of BSU after their initial enrollment as transfer students. Since curriculum choices are difficult to track for this group of students, we exclude them from the analysis to obtain the cleanest estimates.[11]

For the students in our sample, we are able to observe their detailed transcripts, including all the courses/credits they attempted and completed, the grades they

[9]Examples include course-sections dropped within the first ten days of the semester, courses that could be taken multiple times for additional credit per the university catalog, courses repeated at other institutions prior to transfer, and courses taken for an additional undergraduate degree.

[10]More specifically, students were allowed to repeat as many courses as possible until 2013 when a cap of six was imposed. Besides the overall maximum, an individual maximum was also imposed to permit enrollment in the same course for three times after 1995 and then two times after 2015.

[11]A small number of cross-policy course takers, that is, students who made their initial and subsequent attempts of a given course under different grading schemes are included in the current analysis. However, in unreported results, we find that our conclusions remain robust after excluding this group of students.

received, and subsequent decisions made prior to graduation since college entry, such as course repetition and major choice, along with their gender, home address, and SAT/ACT scores reported at college entry. Table 1 shows the summary statistics of the key variables used in the empirical analysis, measured at individual-level (Panel A), semester-level (Panel B), and course-level (Panel C), respectively. During the observation period of 1990–2007, an average student had a zip code median household income of $42,829 (in 1999 dollars). The student attempted 4 courses or 10 credits per semester and received an average grade of $C+$, or 2.6 on a 4.0 scale. At the time of observation, for example, in the years 1990–1994, the semester and cumulative GPAs are 2.2 and 2.3, respectively, and 58% of the students were undecided about their major. Out of the 10 credits that the student registered, slightly over one half were at the 100 level, (i.e., introductory courses having no university-level prerequisites), 3% were a repeated attempt of a given course, and 20% were in the fields of STEM, which include all of the natural sciences, engineering, and most medical sciences. Using a more liberal definition that follows the U.S. Department of Homeland Security STEM Designated Degree Program List and considers all fields of study eligible for the 24-month optional practical training (OPT) extension as one part of the STEM,[12] 27% of the attempted credits would fall into this category. Among the cohorts for which we can track for the minimum of 9 years, 19% of them eventually completed their college education in an average of 12 semesters, with 30%–40% of them obtaining a degree in STEM.

4.3 Empirical Strategy

We identify the effect of grade forgiveness on student outcomes by exploiting the timing of the cancellation and reinstatement of grade forgiveness across students with different propensities of exposure or a different length of exposure in a DD framework. Specifically, for time-varying outcomes (i.e., repetition decision, course difficulty, semester course load, initial-attempt grade, and within-semester grade

[12]See: https://www.ice.gov/sites/default/files/documents/stem-list.pdf. Retrieved on May 22, 2023.

variation), we first implement the following student-level fixed effects model for a sample that includes graded (as opposed to Pass/Fail) courses offered at the time of observation:

$$Y_{iat} = \alpha + \beta GF_{iat} + \delta_a + \gamma_i + S'_{iat}\Lambda + \varepsilon_{iat} \quad (1)$$

where i denotes students and t denotes calendar semester. The subscript a represents the academic progress of student i and is proxied by the number of semesters elapsed since the student's college entry at the time of observation. The student-level fixed effects (γ_i) control for permanent individual traits correlated with both the likelihood of exposure (i.e., timing of college entry) and outcome, such as innate ability, academic preparation, family background, and risk aversion, whereas the academic-progress fixed effects (δ_a) hold constant of any factors related to the student's seniority or class standing that may affect the student's course choice and performance, such as knowledge about institutional policies, study skills, and mindsets/maturity. The column vector S'_{iat} includes additional covariates that vary over time, including semester fixed effects (i.e., Fall, Spring, and Summer), a proxy for required courses at the university level,[13] an indicator for whether the course is repeated for more than once (when applicable), the share of STEM courses offered at the time of observation, and the student's last-observed semester and cumulative GPAs as well as major declaration status (i.e., a binary variable for whether a major was declared) prior to the time of observation. Our key independent variable, GF_{iat}, coded as a binary variable for the adoption of the grade forgiveness scheme thus measures whether the observed change in outcome over a student's college tenure is different under grade forgiveness relative to the grade averaging scheme.

As one way to evaluate the potential impact of time-varying confounders, we replicate the above analysis using an alternative approach for a sample that includes

[13] All undergraduate students at BSU were required to take University Foundation courses as a part of their academic program. Since we do not have any effective means to identify whether a given course was a UF course at the time of observation, we include a binary variable for all 100-level courses as a rough proxy for this potentially mandatory component.

both graded and Pass/Fail courses and invokes students in the latter as a control, whenever possible:

$$Y_{iat} = \alpha + \beta(Graded \times GF)_{iat} + \zeta Graded_{iat} + \delta_a + \eta_t + S'_{iat}\Lambda + \varepsilon_{iat} \tag{2}$$

Here $Graded_{iat}$ takes a value of one if student i is enrolled in a graded (as opposed to Pass/Fail) course in semester t and η_t is a set of academic year fixed effects. While subjecting to largely the same teaching and grading practices of the instructors and other academic policies at various levels (e.g., university, college, and department), the repetition decision and curriculum choice concerning Pass/Fail courses should not be affected by the adoption of grade forgiveness, given that the old grades from Pass/Fail courses will be replaced by the new ones no matter which grading scheme is enacted, per BSU policy.[14] Importantly, during the observation period, BSU did not provide the Pass/Fail grading option for a designated graded course on an individual basis, nor did it impose any restrictions on how many Pass/Fail courses a student was allowed to take while in college. Thus, to the extent that the credits earned from these courses count toward student degree requirements and prerequisites, the distribution of the unobserved ability for Pass/Fail students should reflect that of the students enrolled in graded courses at the same time.

In the event that students in Pass/Fail courses do not serve as a satisfactory control group, as ungraded courses might tend to be delivered through a nontraditional format (e.g., seminars and workshops) and sometimes target a different group of students, we also conduct an event study analysis for graded courses only. Specifically, we treat the cancellation and reintroduction of grade forgiveness as two separate events that potentially affect outcomes in opposite directions using Equation (3):

[14] For example, a student chose to repeat a Pass/Fail course in which she initially failed and subsequently passed. The Pass grade will replace the previous Fail in the student's cumulative GPA under both grading schemes.

$$Y_{iat} = \alpha + \sum_{k=-11}^{6} \beta_k \mathbb{1}(Year_t - 2001 = k) + \delta_a + \gamma_i + S'_{iat}\Lambda + \varepsilon_{iat} \quad (3)$$

where the indicator variables $\mathbb{1}(Year_t - 2001 = k)$ measure the year relative to 2001, when grade forgiveness was reenacted.

With regard to graduation-related/individual-level outcomes (i.e., likelihood to graduate, degree type, and time-to-degree), we adopt a slightly different strategy by comparing whether the difference in outcome between entering cohorts who are never exposed to grade forgiveness (i.e., control) and those who are treated with different lengths of exposure. Assuming that a longer exposure to grade forgiveness in the early stage of a student's college career has a greater impact on their major choice and the time required to earn a degree, we would expect the estimated grade forgiveness effect to strengthen with the intensity of the treatment. Specifically, we construct a series of binary treatment variables GF_{ic} that measure a student's continuous exposure to grade forgiveness since college entrance, ranging from one to six semesters,[15] and then compare – through a series of separate regressions – whether the observed gap in outcome between the treated and control cohorts increases with time, conditional on a set of observable student characteristics P_i, including gender, an indicator for having a missing SAT score, SAT composite score, home zip-code median income, a linear trend for entry cohort, and entry-semester fixed effects (i.e., Fall, Spring, and Summer), shown in Equation (4):

$$Y_{ic} = \alpha + \beta GF_{ic} + P'_i\Gamma + \varepsilon_{ic} \quad (4)$$

In particular, we restrict attention to students who attended BSU during the narrow window of 1990-2007 to ensure a tracking period of 10 years for each student in the sample. We view it as an important advantage of this analysis, given that only 15.3% of first-time undergraduates finished their bachelor's degrees at BSU within

[15] Due to dropouts, the sample sizes for this exercise become progressively smaller as we expand our observation window. A period of six semesters is therefore the longest through which any meaningful patterns emerge.

six years during the observation period.[16] The longer observation window thus allows us to examine the behavior of full-time traditional students, but also that of the part-time and non-traditional students who stay in their programs well beyond the standard college tenure.

5 Main Results

5.1 Course Repetition

Table 2 empirically tests Proposition 4 that the average probability of course repetition will be higher under grade forgiveness than grade averaging by showing the estimated grade forgiveness effects on the likelihood of course repetition using Equation (1) for two samples: all students (columns 1-4) and the students observed immediately before/after the cancellation and reinstatement of the grade forgiveness policy, whose graduation outcomes can be accurately measured (column 5).[17] Given that repeating students are the target group of the grade forgiveness policy, we consider these results as supporting evidence for the risk-taking effect of grade forgiveness that will be investigated in Sections 5.2–5.3.

Column (1) displays the regression result without any covariates for the students observed during the broad window (1991-2016)[18] and indicates that grade forgiveness increases the probability of repeating a given course by 2.2 percentage points, or 89% from the baseline, in years 1995-2000 when the grade averaging scheme was enacted, thus providing support for Proposition 4. The additional inclusion of student-level fixed effects, academic-progress fixed effects, and additional co-

[16] Another contributing factor is the fact that BSU offered a large number of associate degrees during this time frame.

[17] Since the vast majority of students who repeat do so only once, this exercise omits any additional attempt(s) for the same course to obtain the cleanest estimates. It also excludes a repeating's student's initial attempt of a course to avoid double counting in the calculation of repetition rate. These two combined lead to a small reduction in the sample size (i.e., 4% for both broad and narrow-window observations), though additionally considering the case of multiple-time repetition essentially leaves our findings unchanged.

[18] In all analyses involving Equations 1-3, observations in 1990 are dropped to allow for the construction of lagged measures.

variates reduces the magnitude of the grade forgiveness coefficient by about one quarter but does not alter this conclusion fundamentally (columns 2-4), suggesting that the unobserved characteristics of the students such as ability or academic preparedness across cohorts do not drive the observed result. Restricting attention to the students observed in the immediate neighborhood of policy changes (1991-2007; column 5) leads to qualitatively similar results. Considering the latter as our preferred specification (column 5), Table 2 suggests an average of 1.8 percentage points, or a 72% increase in the repetition probability, as students' most direct response to the adoption of grade forgiveness over the scheme of grade averaging.

Panel A of Table B1 examines the estimated grade forgiveness effect using the group of unlikely eligibles as a control, the students who enrolled in Pass/Fail courses offered at the same time. Implementing Equation (2) for our preferred sample, the students observed during the narrow window, we find a highly similar result (i.e., a 1.7 percentage points or 74% increase in the probability of repetition), implying that the potential impact of time-varying confounders, such as student composition, instructor grading practices (e.g., grade inflation), and institutional policies that potentially affect the repetition behavior of students (e.g., tuition fee charged for overload or part-time students to take an additional course), if exists, is not extensive in our context.

To guard against the possibility that students in Pass/Fail courses do not serve as an adequate control, Figure 2 presents the results of an event study analysis using Equation (3). Focusing on the estimates around the two events, we see a clear drop, followed by a sharp surge in the likelihood of course repetition in 1995 and 2001 by approximately the same magnitude (i.e., 2 percentage points) that coincide with the cancellation and reintroduction of grade forgiveness. Importantly, all of the estimated effects of grade forgiveness during the grade-averaging period are statistically indistinguishable from zero, suggesting that the grade forgiveness policy was not enacted in response to falling (or rising) rates of course repetition.

5.2 Course Difficulty

As suggested in Proposition 1, an increased option value of the insurance against low grades will incentivize students to attempt more challenging courses. This section empirically tests this hypothesis by examining patterns in the choice of course difficulty during the same time frame. We conduct the analysis using two different measures of course difficulty, an objective difficulty based on the subject matter of a given course (i.e., STEM designation) and a perceived difficulty based on the course-level grading harshness. We treat the former as our preferred measure both because it is more independent of student individual biases/perspectives and because it covers a more representative set of courses (for more discussions see below), though as will be demonstrated in this section, both measures yield highly similar results.

5.2.1 Objective Measure of Course Difficulty

Panel A of Table 3 reports the grade forgiveness effects on the likelihood of taking a STEM course using both the classical and OPT-based definitions of STEM subjects (see Section 4.2 for more discussions) for the students observed in the broad (column 1) and narrow (column 2) windows. Across different STEM measures and data samples, we obtain highly consistent results suggesting that grade forgiveness increases the likelihood of taking a STEM course by 2 percentage points or 8% from the baseline for all students.

Replicating the analysis for non-repeating students or first-time course takers (Panel B of Table 3) and never-repeating students or students who never repeated any courses during the observation period (Panel C of Table 3), reduces the estimated magnitude of grade forgiveness by approximately one sixth and one half, respectively, though these results continue to deliver qualitatively similar conclusions. This implies that the previously observed results for all students are not solely driven by repeaters. This is an important consideration in our context, as many courses in STEM not only have relatively stringent grading standards, but also are more likely to be sequential, compared to their non-STEM counterparts. Therefore, STEM students who do poorly in a course could be more easily thrown off track of their study and forced to leave their major than non-STEM students in the absence of grade forgiveness. While our identification strategy does not allow us to separately quantify the course repetition or GPA inflation effect of grade forgiveness, the finding of a significant treatment effect for the group of never-repeaters lends credence to the notion that there is an independent effect of grade forgiveness on student curriculum choice that goes beyond course repetition and/or GPA inflation. Given that repetitions completed under the grade averaging scheme would not be associated with any GPA boost, columns 1–2 of Panel C can be viewed as lower-bound estimates of the grade forgiveness effect on student course choice: a 1 percentage point or a 4–5% increase in the likelihood of STEM course taking relative to the baseline, providing some support for Proposition 1.

Breaking the overall estimate by student type reveals that the estimated effect of grade forgiveness is concentrated on non-STEM majors or students who declared a non-STEM or no major at college entry, relative to their STEM counterparts (i.e., a 2–3 percentage points or 10–11% increase from the baseline), when both the initial and subsequent attempts of a given course are accounted for (columns 3–4 of Panel A). Even after excluding repeating students from the analysis, this pattern remains the same (i.e., a 1 percentage point or 4–5% increase from the baseline), confirming that the course choice effect of grade forgiveness is primarily driven by the students on the margin to start on the STEM trajectory. Thus, grade forgiveness appears to be more effective in shifting students into STEM than encouraging them to stay in.

Panels B–C of Table B1 uses the same specification as that in column 2 of Table 3 while invoking students in Pass/Fail courses as a control group. Once again, a highly similar pattern emerges. Across data samples and definitions of the STEM subjects, we obtain an estimated grade forgiveness effect about one-third greater than what is reported in columns 1–2 of Table 3 (i.e., a 3–4 percentage points or 13–14% increase from the baseline for all students and a 2–3 percentage points or

9–10% increase from the baseline for never-repeating students) with our findings remaining largely unchanged.

Figures 3a–3c show the event study estimates for all students, non-repeating students, and never-repeating students, respectively, using the OPT-based definition of STEM subjects. Potentially due to the low incidence of STEM course-taking in the early years, the classical definition of STEM yields a less pronounced yet qualitatively similar pattern. In two out of the three cases (i.e., all students and non-repeating students), the graphs suggest two trend breaks, a sudden drop of 2 percentage points starting in 1995 when grade forgiveness was abolished and a gradual increase by 5 percentage points starting in 2003 or two years after grade forgiveness was reinstated. Across all cases, the likelihood of STEM course-taking increased by 4 percentage points within seven years of the grade forgiveness reversal, and the effect continued to grow over the next nine years and reached 10 percentage points by 2016. None of the grade forgiveness coefficients are precisely estimated for never-repeating students in the pre-1995 years (Figure 3c), which is likely related to the smaller sample size. Additionally, the time-series pattern found in all samples appears to be a mirror image of that observed for the repetition decision (i.e., Figure 2) except for a two-year lag. This may be the case if student course choices are only indirectly affected by the grade forgiveness policy and that they tend to strategically schedule their semesters ahead of time.

5.2.2 Subjective Measure of Course Difficulty

To have a more comprehensive view of how grade forgiveness may have affected student course-taking behavior, Table B2 carries out the same analysis as Table 3 using a continuous measure for course difficulty, the grading harshness of a given course. Assuming that courses that issue relatively low grades are perceived as more difficult to succeed in by students, a negative association of this measure with grade forgiveness would indicate a higher level of risk-taking in learning. Specifically, we construct this measure based on the weighted average grade assigned by a given course across different sections and semesters or the corresponding course-specific fraction of D's and F's, where the weight is the number of course credits. We assume that the difficulty level and credit hours of a course are simultaneously chosen by a student and jointly determine the quality point of the course when calculating the student's cumulative GPA in most institutions.[19] One potential caveat of this construction is that the course-specific grade might be influenced by student risk-taking behavior when the average performance of students declines as they challenge themselves by enrolling in a course they would not have enrolled in in the absence of grade forgiveness. To circumvent this concern, we restrict our attention to courses offered during the baseline grade averaging period (i.e., 1995–2000) and use the grades awarded during this period only by each of these courses, along with associated credit hours, as the measure for grading harshness. This practice will result in an understatement of the true effect of grade forgiveness on student curriculum choice by omitting all the courses offered before 1995 and after 2000, but nevertheless sheds light on the course-taking patterns in subjects unrelated to STEM or within the STEM fields.

These results are reported in Table B2. When both initial and subsequent attempts of a given course are considered, students tend to enroll in courses that have a worse grading outcome or a lower perceived 'quality point' by 0.06-0.07 points or 1% relative to the baseline (columns 1 and 3 of Panel A) and courses that are 2 percentage points or 4–5% more likely to issue a grade of D or below relative to the baseline (columns 2 and 4 of Panel A). Excluding non-repeating students (Panel B) and never-repeating students (Panel C) diminishes the estimated magnitude but leads to qualitatively similar conclusions.

It is worth noting that the estimated magnitude of grade forgiveness using the perceived course difficulty is considerably smaller than that obtained from the ob-

[19] To be more precise, the grading harshness measure for each course j is constructed as $\frac{\sum_1^N Grade_{ij}}{N} \times Credits_j$ or $\frac{\sum_1^N \mathbb{1}(\text{Letter Grade = D or F})_{ij}}{N} \times Credits_j$, where i is an individual grade at student-attempt level and N is the total number of students who took the course across different sections and semesters over time.

jective course difficulty. This decline is likely related to the omission of pre-1995 and post-2000 courses but can also reflect the fact that the observed grade forgiveness effect is largely (if not completely) driven by students taking courses in STEM fields. This hypothesis is supported by unreported results from separate regressions for STEM and non-STEM courses, where we find an estimated effect size of grade forgiveness three times as large as that for the full sample when only STEM courses are considered (available upon request). If this is the case, while overlooking the course-taking behavior of non-STEM students, the resulting downward bias which the estimates in Table 3 are subject to likely is not extensive. At any rate, the primary message delivered by Table B2 is the same as before: across different data samples and selection criteria, we find that students take more harshly-graded courses or courses perceived as more difficult after the enactment of grade forgiveness.

In sum, we obtain consistent evidence across different measures of course difficulty that grade forgiveness nudges students into taking more challenging courses. While the nudge is, understandably, more pronounced for repeating students, we estimate a significant effect of grade forgiveness for never-repeaters and find that the effect is largely concentrated on non-STEM majors rather than existing STEM students measured at college entry.

5.3 College Major and Time-to-Degree

Having discovered that grade forgiveness incentivizes students to pursue more courses in challenging subjects, this section sets out to investigate whether the grade-forgiveness-induced course choices have any longer-term implications for the likelihood of graduation, college major, and time-to-degree. For the sake of brevity, only results using the OPT-based definition of STEM subjects are reported in the section. The use of the classical definition of STEM subjects leads to slightly weaker, but fundamentally similar results. These results are available upon request.

5.3.1 College Major

Panels A–B of Table 4 compare and contrast the results for overall degree attainment and degree attainment in STEM among graduates. While there is not any statistically distinguishable effect of grade forgiveness on the graduation rate overall (Panel A), conditional on graduation we observe an average increase in the likelihood of obtaining a STEM degree by 7.5 percentage points for students who enter college when grade forgiveness is in place relative to their observationally equivalent counterparts who start college when grade-averaging is enforced across cohorts (Panel B). Furthermore, the longer a student is exposed to grade forgiveness, the more likely he/she will be to graduate with a degree in STEM in general. For example, entering cohorts who are continuously exposed to grade forgiveness in their first three semesters are more likely to obtain a STEM degree by 7.7 percentage points relative to entering cohorts who are not exposed to the policy in any of their first three semesters. These effects translate to a 22% increase in STEM degree attainment from the baseline and the difference grows to as large as 33% when we compare cohorts who are six semesters apart in terms of the length of exposure.

Excluding students who repeated at least once reduces the sample sizes substantially (33% for all students and 62% for graduates) (Panels C–D of Table 4). Despite the diminished statistical power, we estimate an even greater impact of grade forgiveness on STEM degree attainment across cohorts (8.3 vs 7.5 percentage points or 23% vs 22%) and continue to observe a positive association of the effect size with the length of exposure. For cohorts who are six semesters apart, the enactment of grade forgiveness leads to 11.8 percentage points or a 35% increase in the likelihood of STEM degree attainment relative to the baseline.

5.3.2 Time-to-Degree

Measuring time-to-degree as the number of semesters for a student to obtain a Bachelor's degree, Panels A and D of Table 5 find no evidence that grade forgiveness delays graduation for all students or for never-repeating students, respectively. Breaking down this result by degree type (Panels B–C and Panels E–F), we find that STEM graduates who spend first semesters of their college career when grade forgiveness is enacted obtain their degrees sooner than their counterparts who are exposed to the alternative grading scheme, even though there is not a monotonic/systematic relationship between the estimated magnitude of grade forgives effect and length of exposure.

Coupling this observation with the relatively small sample sizes for this exercise (i.e., 697-1512 students), we do not attach much emphasis to these findings, though the evidence we find for the general population of students remains consistent and strong. That is, even using our most stringent sample selection criterion (i.e., never-repeating students) we find that grade forgiveness has no measurable impact on the student's overall graduation rate and/or time-to-degree across disciplines.

5.4 Mechanism Investigation

Given the findings in Sections 5.2-5.3, a natural question to ask is why the pursuit of a more challenging curriculum does not cause students to lose time toward their degrees or diminish their chances of graduation altogether. This is a valid question considering that the observed grade forgiveness effect is concentrated on the students who begin college with a major in non-STEM subjects or as undeclared. Assuming that a change in declared major typically introduces new prerequisites and renders some completed coursework irrelevant, especially for students who switch majors in later years of their studies, it is reasonable to expect grade forgiveness to have a negative impact on graduation outcomes. This section provides insight into this issue by examining the changes in students' choice of study pace and effort expenditure under grade forgiveness.

5.4.1 Semester Course Load

Besides engaging in courses with more academic rigor, Proposition 1 implies that students might also be willing to take on a heavier academic load under grade forgiveness in comparison with the grade averaging scheme. Even though our theoretical model does not differentiate between the quality and quantity aspects of curriculum choices and therefore is unable to infer their relationship with each other, we can test this hypothesis in our empirical analysis. Intuitively, if grade forgiveness encourages students to attempt a difficult course and a heavier load simultaneously, then we would expect the grade forgiveness to increase the observed semester course load. However, if students sacrifice their course load to increase their bandwidth for the more demanding coursework in the presence of resource constraints, then we would expect the average semester course load to decrease under grade forgiveness.

The evidence from the semester course load effect of grade forgiveness, as measured by the number of registered courses within a semester for the same students, supports the former. These results are presented in Table 6. The results obtained for the number of per-semester credits are highly similar, though are not reported in the paper for brevity. Between repeating and non-repeating students (Panel A), students enroll in 0.09-0.11 more courses under grade forgiveness, representing a 2-3% increase in comparison to the baseline period when the grade averaging formula was employed. Different than the results obtained for course difficulty in Table 3, the exclusion of repeating students from the sample leads to an increase in the estimated grade forgiveness coefficient by one-fourth to one-third. Based on the most stringent sample (i.e., never-repeating students), we observe an increase of 0.1 more registered courses or a 4% increase relative to the baseline load.

Interestingly, the effect of grade forgiveness is, once again, driven by students who begin college as non-STEM majors, the very group who are incentivized to attempt more STEM courses by grade forgiveness (see Table 3). Given that major switching typically implies more courses or credits to graduate, the accelerated study pace might be a natural response to the major changing activities associated with grade forgiveness. As such, the theoretical trade-off between student choices of course difficulty and study pace does not receive any empirical support in this exercise, which may be explained by a non-binding aggregate level of effort expenditure.

5.4.2 Student Effort Expenditure

As an alternative explanation to the observed results, this section provides evidence on the role of grade forgiveness in student effort expenditure. We carry out our analysis using three different methods.

Panel A of Table 7 uses students' initial-attempt performance as a proxy for their effort expenditure on a given course at the time of observation and examines its temporal patterns before and after the adoption of grade forgiveness. Implementing our original model specification along with a set of course-level fixed effects to account for unobserved heterogeneity in student course choice, we find that the enactment of grade forgiveness is associated with a small change in student grade (i.e., an increase of 0.03 points or 1% from the baseline for all students and a decrease of 0.02 points or 1% from the baseline for never-repeating students), though neither effect is precisely estimated at 5% level. Considering that STEM courses tend to have more stringent grading standards than their non-STEM counterparts and that students may receive a lower grade in a challenging or more harshly-graded course, even for the same amount of effort expended, the weak initial-performance effect of grade forgiveness reported in Panel A suggests that the decline in student effort, if there is any, is not extensive in our context.

Panel B takes a different approach by differentiating between "per-course" from "per-semester" effort for each student. In other words, we contend that the initial-attempt performance is not an accurate measure for student effort anymore if they take on a heavier course load under grade forgiveness. For example, students taking one course and receiving an A do not necessarily make more effort than those who take two courses and receive two Bs' at the same time. To this end, we use the students' semester quality points rather than the initial grade from an individual course as an alternative measure of student effort. The semester quality point is calculated as the weighted initial grade for each student, where the student's semester course load or registered credit hours serves as the weight. Using the same model specification as that in column 3 of Table 3 minus course-level covariates, we find that the per-semester effort effect of grade forgiveness to be positive, albeit remaining insignificant in our most stringent sample, never-repeating students.

Finally, assuming that grade forgiveness did not induce more effort expenditure or that students' total effort was binding upon the enactment of grade forgiveness, then Proposition 3 predicts that the time/effort allocated among the different courses attempted within a semester will be more dispersed. Panels C–D thus test this hypothesis by regressing two alternative measures of within-semester-effort-allocation, the standard deviation of and the max-min difference between the initial grades received within a semester, respectively, against the grade forgiveness indicator. In neither of the cases do we find any meaningful patterns. These results provide a counter-example that students do alter the aggregate level of effort expenditure in response to the enactment of the grade forgiveness grading scheme.

In summary, evidence revealed in this section suggests that besides more demanding coursework, non-STEM majors also take on a heavier semester load and exert similar amount of effort under grade forgiveness relative to the grade averaging scheme. Given that a lower course grade may be a result of students attempting a difficult course, there is a theoretical possibility that grade forgiveness actually induces more effort among these students, although further explorations are needed to substantiate the conjecture.[20]

[20] In addition, one implicit assumption we make in interpreting the evidence in Table 7 is that instructors do not respond to the enactment of grade forgiveness. If instructors assign a lower grade to students because of the more generous repetition policy, then it would bias our estimates downward even further.

6 Grade Forgiveness and STEM Underrepresented Groups

6.1 Female Students

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 significantly more than men and the harsher grading policies in STEM courses disproportionately discourage women's participation in STEM (Rask and Tiefenthaler, 2008; Ost, 2010; Owen, 2010; Butcher, McEwan, and Weerapana, 2014; Minaya, 2017; Ahn et al., 2019). A related and relevant question that our study can answer is whether grade forgiveness alters female students' engagement in STEM subjects, both in terms of their choices of curriculum and college major.

Panel A of Table 8 reports the estimated grade forgiveness coefficients from separate regressions for male and female students across samples. While all of the effects are positive, indicating that grade forgiveness encourages STEM participation across genders, the estimated magnitude tends to be smaller for female than male students, even though the differences are not statistically significant for our most stringent sample (see the last row of Panel A). We find consistent results for STEM-degree attainment in the longer run (Table B3). While we are unable to precisely estimate most of the grade forgiveness coefficients for never-repeating students in this case, likely due to a lack of statistical power, we observe for the full sample a greater effect size for males than for females by 1.7 percentage points after continuous exposure to grade forgiveness for the first six semesters.

The finding of a greater grade forgiveness effect for male than for female students is plausible if there exists a gender difference in overconfidence (Stinebrickner and Stinebrickner, 2012). If the level of over-optimism about completing a given degree is more substantial among male than female students, then the additional space for self-discovery through grade forgiveness might generate greater gains for the group of students who are more misinformed of their ability at college entrance or the quality of the initial match between one's ability and a chosen path. Additionally, the extent to which students are willing to bear risks and/or ambiguity may play a role when female students require a higher level of compensation for the introduction of uncertainty than male students (Borghans, Heckman, Golsteyn, and Meijers, 2009). While offering students insurance against failures, grade forgiveness does not promise higher grades. If the reservation price of exploratory learning is higher than the expected benefit of grade forgiveness more so for females than for their male counterparts, then it could exacerbate the observed gender gap.

6.2 Low-Income Students

Besides women, students from lower social-economic status are also known as underrepresented in college STEM majors. Panel B of Table 8 shows grade forgiveness effects on students from financially disadvantaged backgrounds. In particular, we proximate a measure of whether a student comes from a low-income background by linking their home address zip code to the median household income reported in the 1999 Census and consider a student as low-income if his/her home is located in a zip code below the US median.[21] Across samples, we observe a favorable impact of grade forgiveness on the course choices of both high- and low-income students with the estimated effect size being 1-2 percentage points greater for the former than the latter, though the differences are not statistically significant at 5% level for the most stringent sample. Replicating the analysis for STEM-degree attainment in Table B4 again produces standard errors that are substantially larger than those observed before, especially for the group that has smaller sample sizes (i.e., low-income students), but the observed pattern remains the same: there is a positive association of grade forgiveness with the outcome for both groups of students.

[21] The 1999 Census is currently the most up-to-date source of income information available at the zip code level.

6.3 Non-Whites

Finally, Panel C of Table 8 reports the results for white and non-white students separately as a means to assess grade forgiveness' influence on the existing racial gap. Since the information on a student's race is not available to us until later years, this exercise is restricted to the students who attended BSU after 1997. As such, we are unable to replicate this analysis for STEM-degree attainment due to a small sample size. The limited evidence we found for curriculum choice, nevertheless, suggests a greater impact of grade forgiveness for the relatively more advantageous group, though, once again, this difference is not statistically significant across samples.

Overall, results obtained in this section suggest that while grade forgiveness has a favorable impact on the participation and degree attainment of STEM for the general population of students, it is not particularly effective in closing the existing gender, income, or racial gaps in these fields.

7 Discussion and Conclusion

This study postulates that a small change in the GPA calculation formula that allows repeating students' initial-attempt grades to be exempt from as opposed to being averaged into their cumulative GPAs can make a significant impact on their learning outcomes, no matter whether the students repeat any courses while in college. By exploiting the exogenous timing of the cancellation and reintroduction of a grade forgiveness policy at a four-year public institution, we show that the adoption of the grade forgiveness scheme nudges its students into taking a more challenging course and/or degree as well as progressing at a faster pace. We also find indirect evidence that the observed results are unlikely driven by concurrent academic policies or the changes in student composition or instructor teaching/grading practices over time.

We interpret these results as the risk-taking effect of grade forgiveness that operates independently of its effects on student course repetition behavior and/or

GPA inflation. In comparison to the traditional practice of grading averaging, grade forgiveness allows students additional chances to correct a prior mistake without leaving a negative mark on their cumulative GPAs. In this way, it facilitates a safe environment where students can search, inquire, and pursue topics that engage them. As demonstrated in the paper, the extra room for trial and error is particularly critical to students who take challenging courses as a primary means to assess their aptitude in corresponding disciplines, such as non-STEM entry majors who tend to move into the STEM fields. With little existing knowledge to attach to the subject matter, the notion that they will not be docked in their academic record for the potentially poor initial performance can encourage them to step out of their comfort zone and explore something new beyond a high grade. The additional self-discovery process that otherwise would not have taken place in a punitive environment can reduce the original mismatch between students' preference/ability and their majors due to information friction, alter the type of knowledge and skills that they develop in college, and therefore has broader implications for the quality of labor force in the economy as well as the distribution of workers across sectors.

This evidence presented in this paper can be viewed as a specific instance of an unintended consequence that the traditional punitive GPA policies (e.g., increasing the hard cap on the number of course withdrawals, restraining the number and circumstances under which repeating is allowed, and tightening the eligibility criteria for pass/fail grading options) have for genuine learning.[22] These punitive practices are oftentimes equated with a higher level of expectations and academic rigor in popular opinion. However, by supporting and rewarding errorless learning, these rigid grading practices may inadvertently incentivize students to prioritize results over the process and lead to the "minimax strategy" (Kruglanski 1978; Kruglanski et al. 1977), that is, to spend the minimum amount of effort needed to obtain the

[22] A counter example would be the experimental grading policy that the Massachusetts Institute of Technology initiated in 2018, which allows up to three science core General Institute Requirements to be graded on a Pass/No Record basis for first-year students in order to "provide [them] with greater opportunity to explore fields of study at the beginning of [their] academic career." https://registrar.mit.edu/classes-grades-evaluations/grades/grading-policies/experimental-grading-policy Retrieved on May 21, 2023.

maximum benefits (i.e., grades).

The current study is certainly not the final word on the debate around the merits or demerits of grade forgiveness for students. For instance, it largely ignores the outcomes of repeating students, who are typically the target group of grade forgiveness policy (Chen and Jiang, 2023), nor do we comment on its role in grade inflation or the effectiveness in comparing student performance across institutions (Marx and Meeler, 2013). However, our finding that students take more risks upon the removal of the repetition penalty raises one interesting question as to whether the ways that colleges traditionally calculate student GPA might do more harm than good to students, by deterring them from acquiring desirable skills and pursuing a challenging but potentially rewarding career path. From this perspective, the elimination of punitive practices from college GPA policies might serve as a fresh solution to some age-old problems that higher education faces ranging from increasing student motivation and engagement and fostering independent thinking to building greater resilience in our future generations.

References

Achen, A. C. and P. N. Courant (2009). What are grades made of? *Journal of Economic Perspectives 23*(3), 77–92.

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.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of econometrics 121*(1-2), 343–375.

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.

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.

Betts, J. R. and J. Grogger (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review 22*(4), 343–375.

Blume-Kohout, M. E. and J. P. Scott (2022). Incentivizing stem participation: Evidence from the smart grant program. *Southern Economic Journal*.

Bonesrønning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics 12*(2), 151–167.

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.

Bostwick, V. K. and B. A. Weinberg (2022). Nevertheless she persisted? gender peer effects in doctoral stem programs. *Journal of Labor Economics 40*(2), 397–436.

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.

Card, D. and A. A. Payne (2021). High school choices and the gender gap in stem. *Economic Inquiry 59*(1), 9–28.

Casas, F. R. and D. E. Meaghan (1996). A study of repeated courses among secondary students in ontario. *Journal of Educational Research 90*(2), 116–127.

Chen, K., Z. Hansen, and S. Lowe (2021). Why do we inflate grades? the effect of adjunct faculty employment on instructor grading standards. *Journal of Human Resources 56*(3), 878–921.

Chen, K. and X. Jiang (2023). Learning from mistakes: The implications of course repetition for student subsequent success. *Unpublished Manuscript*.

Clifford, M. M. (1991). Risk taking: Theoretical, empirical, and educational considerations. *Educational psychologist 26*(3-4), 263–297.

Clifford, M. M. and F.-C. Chou (1991). Effects of payoff and task context on academic risk taking. *Journal of Educational Psychology 83*(4), 499.

De Paola, M. and V. Scoppa (2014). The effectiveness of remedial courses in italy: a fuzzy regression discontinuity design. *Journal of Population Economics 27*(2), 365–386.

Denning, J. T. and P. Turley (2017). Was that smart? institutional financial incentives and field of study. *Journal of Human Resources 52*(1), 152–186.

Dickson, V. A. (1984). An economic model of faculty grading practices. *Journal of Economic Education 15*(3), 197–203.

Figlio, D. N. and M. E. Lucas (2004). Do high grading standards affect student performance? *Journal of Public Economics 88*(9-10), 1815–1834.

Fournier, G. M. and T. R. Sass (2000). Take my course, please: The effects of the principles experience on student curriculum choice. *Journal of Economic Education 31*(4), 323–339.

Freeman, D. G. (1999). Grade divergence as a market outcome. *Journal of Economic Education 30*(4), 344–351.

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.

Henriksen, D., M. Henderson, E. Creely, A. A. Carvalho, M. Cernochova, D. Dash, T. Davis, and P. Mishra (2021). Creativity and risk-taking in teaching and learning settings: Insights from six international narratives. *International Journal of Educational Research Open 2*, 100024.

Hill, A. J. (2014). The costs of failure: Negative externalities in high school course repetition. *Economics of Education Review 43*, 91–105.

Jewell, R. Todd, M. A. M. and M. A. Tieslau (2013). Whose fault is it? assigning blame for grade inflation in higher education. *Applied economics 45*(9), 1185–1200.

Jiang, X. (2021). Women in stem: Ability, preference, and value. *Labour Economics 70*, 101991.

Marx, J. and D. Meeler (2013). Strike four! do-over policies institutionalize gpa distortion. *Quality Assurance in Education 21*(1), 39–53.

Masten, A. S. and J. Obradović (2006). Competence and resilience in development. *Annals of the New York Academy of Sciences 1094*(1), 13–27.

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.

Sabot, R. and J. Wakeman-Linn (1991). Grade inflation and course choice. *Journal of Economic Perspectives 5*(1), 159–170.

Sewell, A. and A. S. George (2000). Developing efficacy beliefs in the classroom. *Journal of Educational Enquiry 1*(2).

Speer, J. D. (2017). The gender gap in college major: Revisiting the role of pre-college factors. *Labour Economics 44*, 69–88.

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.

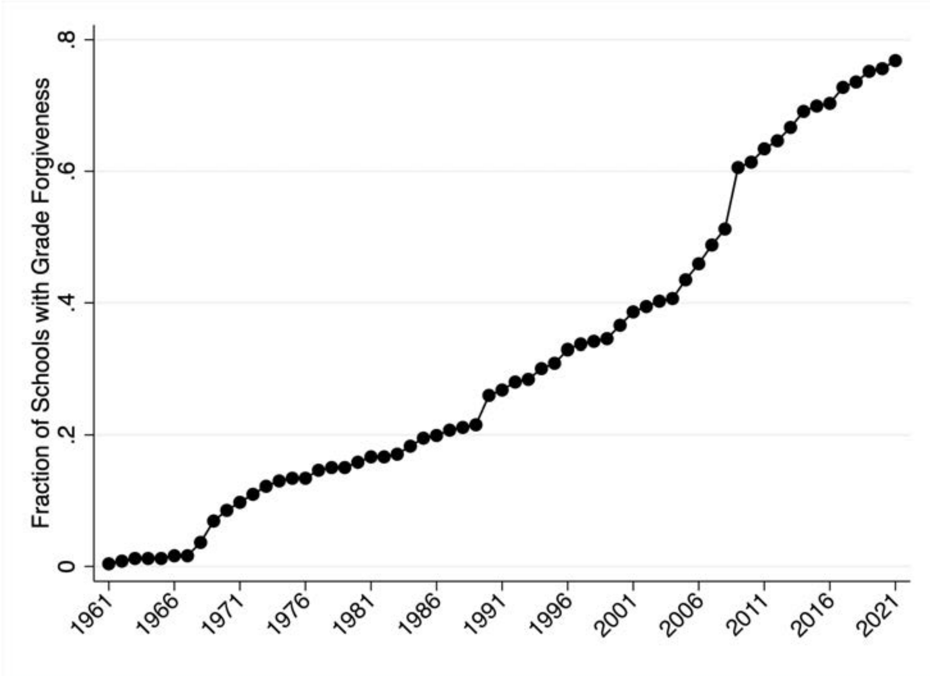


Figure 1: Fraction of Four-Year Institutions with Grade Forgiveness

Notes: This figure shows the implementation of the grade forgiveness policy across the 380 four-year institutions with above 10,000 student enrollment in the Integrated Post-secondary Education Data System as of 2021. We exclude 133 universities with missing data.

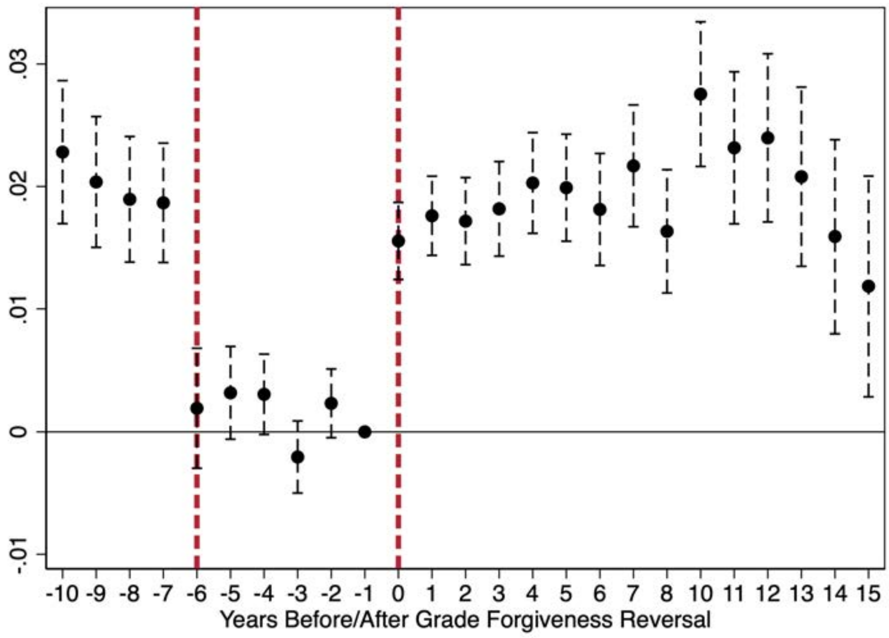
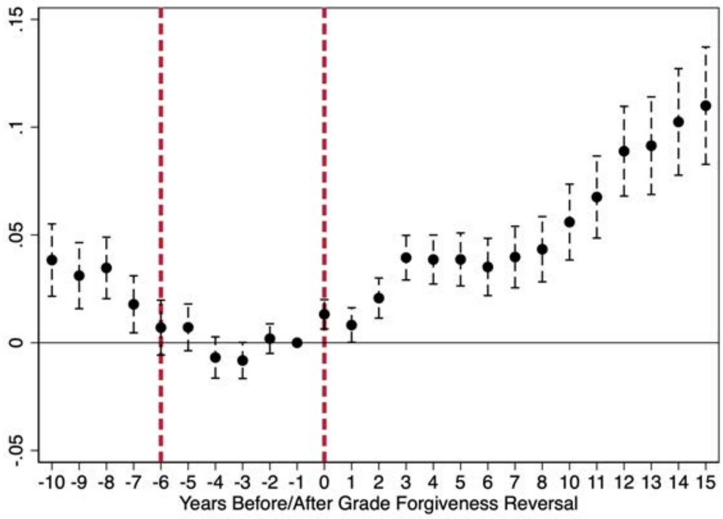
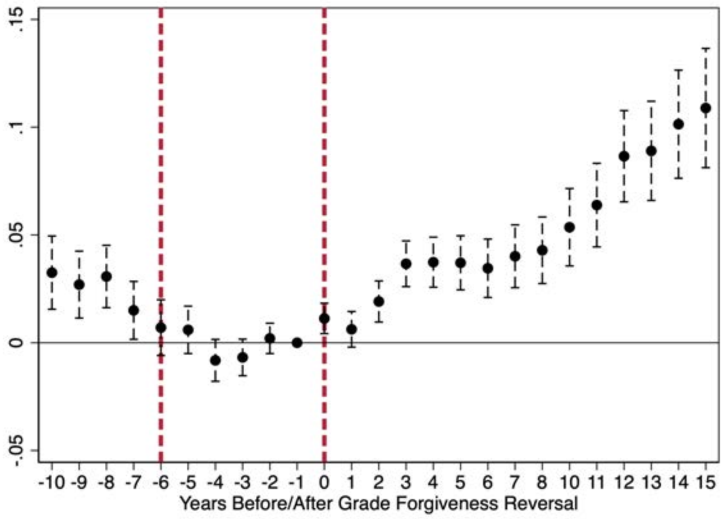


Figure 2: Probability of Course Repetition

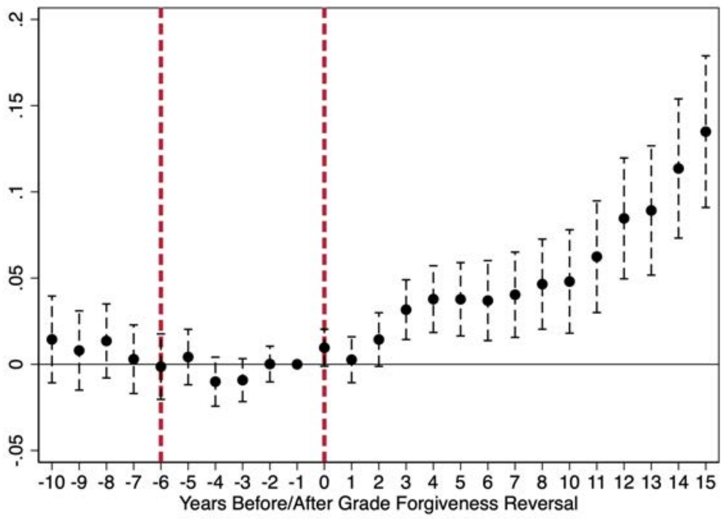
Notes: This figure shows the regression coefficients with 95% confidence intervals from Equation (2) for the likelihood of repeating a given course during the observation period. Each coefficient is estimated relative to 2001 when the grade forgiveness policy was reinstated.



(a) All Students



(b) Non-Repeating Students



(c) Never-Repeating Students

Notes: The above figures show the regression coefficients with 95% confidence intervals from Equation (2) for the probability of taking a STEM course using the OPT-based definition of STEM. Each coefficient is estimated relative to 2001 when the grade forgiveness policy was reinstated.

Table 1: Summary Statistics on Key Variables

Variable	Mean	Std. Dev.	N
Student-Level Observations (1990-2007)			
Female	0.5350	0.4988	57165
Home zipcode median household income (1999$)	42829	9947	54028
Graduation Rate	0.1874	0.3902	57426
#Semesters to Graduate	12.1533	4.1176	10787
%STEM Degrees (Classical)	0.2956	0.4563	7024
%STEM Degrees (OPT)	0.3938	0.4886	7024
#Semesters to STEM Degrees (Classical)	12.5043	3.5983	2076
#Semesters to STEM Degrees (OPT)	12.3644	3.6515	2766

By Period Statistics

Semester-Level Observations	**1990-1994**	**1995-2000**	**2001-2007**
N	72481	87432	110172
#Courses Attempted	3.5617	3.5517	3.6523
	(1.9702)	(1.8597)	(1.8020)
#Credits Attempted	10.0883	10.0183	10.5325
	(5.3041)	(4.9281)	(4.7491)
Term Grade Points	24.0252	24.4092	25.9047
	(17.1745)	(17.1915)	(17.0871)
Semester GPA	2.3447	2.2624	2.2875
	(1.4010)	(1.3969)	(1.3723)
Cumulative GPA	2.3733	2.3002	2.4328
	(1.2769)	(1.2419)	(1.2116)
Declared Major	0.4172	0.5938	0.6867
	(0.4931)	(0.4911)	(0.4639)
Share of STEM courses offered	0.1798	0.1987	0.2138
	(0.0065)	(0.0169)	(0.0097)

Course-Level Observations	**1990-1994**	**1995-2000**	**2001-2007**
N	225196	272622	342242
Repeated Attempt	0.0327	0.0246	0.0531
	(0.1777)	(0.1549)	(0.2242)
Grade Point	2.5745	2.6541	2.7152
	(1.3878)	(1.3149)	(1.3073)
STEM (Classical)	0.2017	0.2072	0.2255
	(0.4012)	(0.4053)	(0.4179)
STEM (OPT)	0.2657	0.2633	0.2865
	(0.4417)	(0.4404)	(0.4521)
100-Level Course	0.5227	0.5318	0.5350
	(0.4995)	(0.4990)	(0.4988)

Notes: This table describes the data used in the analysis by student-level (upper panel), semester-level (middle panel), and course-level (lower panel) for all first-time non-transfer undergraduate students. To reflect the regression samples, we show student-level characteristics for those who first entered BSU between 1990-2007. For semester-level and course-level characteristics, we present the mean and standard deviation (in parentheses) for the key variables by three policy periods (i.e., 1990-1994, 1995-2000, and 2001-2007).

Table 2: Probability of Course Repetition

	Broad Window (1990-2016)				Narrow Window (1990-2007)
	(1)	(2)	(3)	(4)	(5)
GF	0.0220***	0.0209***	0.0191***	0.0170***	0.0177***
	(0.0004)	(0.0004)	(0.0008)	(0.0007)	(0.0007)
Academic Progress F.E.		X	X	X	X
Individual F.E.			X	X	X
Covariates				X	X
Level of Clustering	Student-Section	Student-Section	Student	Student	Student
Sample Mean (1995-2000)	0.0246	0.0246	0.0246	0.0246	0.0246
N	1286228	1286228	1286228	1286228	839519

Notes: This table shows the estimated effect of grade forgiveness on the probability of repeating a course. Each cell represents a grade forgiveness coefficient from a separate regression based on Equation (1). Standard errors are clustered at the student-course-section and student levels, respectively, in columns 1-2 and 3-5 of the table. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Choice of Course Difficulty (STEM Designation)

	Broad Window (1990-2016)	Narrow Window (1990-2007)		
	(1)	(2)	(3)	(4)
	Full Sample	Full Sample	STEM Majors	non-STEM Majors
Panel A: All Students				
Panel A.1 Classical Definition of STEM				
GF	0.0224***	0.0218***	-0.0066	0.0245***
	(0.0022)	(0.0022)	(0.0073)	(0.0022)
Sample Mean (1995-2000)	0.2698	0.2698	0.4397	0.2344
Panel A.2 OPT-Based Definition of STEM				
GF	0.0172***	0.0164***	0.0060	0.0175***
	(0.0020)	(0.0020)	(0.0067)	(0.0021)
Sample Mean (1995-2000)	0.2120	.2120	0.3567	0.1819
N	1344704	876292	170238	706043
Panel B: Non-Repeating Students				
Panel B.1 Classical Definition of STEM				
GF	0.0199***	0.0192***	-0.0095	0.0219***
	(0.0022)	(0.0022)	(0.0074)	(0.0022)
Sample Mean (1995-2000)	0.2644	0.2644	0.4331	0.2294
Panel B.2 OPT-Based Definition of STEM				
GF	0.0153***	0.0144***	0.0032	0.0157***
	(0.0020)	(0.0020)	(0.0069)	(0.0021)
Sample Mean (1995-2000)	0.2079	0.2079	0.3502	0.1784
N	1283579	840002	162366	677624
Panel C: Never-Repeating Students				
Panel C.1 Classical Definition of STEM				
GF	0.0139***	0.0110***	-0.0052	0.0120***
	(0.0032)	(0.0032)	(0.0109)	(0.0032)
Sample Mean (1995-2000)	0.2564	0.2564	0.4315	0.2210
Panel C.2 OPT-Based Definition of STEM				
GF	0.0093***	0.0072**	0.0075	0.0070**
	(0.0029)	(0.0029)	(0.0099)	(0.0029)
Sample Mean (1995-2000)	0.1964	0.1964	0.3434	0.1667
N	624764	428743	76103	352636

Notes: This table shows the estimated grade forgiveness effect on the likelihood of taking a STEM course using two different definitions of STEM for all students (Panel A), non-repeating students (Panel B), and never-repeating students (Panel C). Standard errors are clustered at the student level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Probability of Graduation

	(1) First 1 Term	(2) First 2 Terms	(3) First 3 Terms	(4) First 4 Terms	(5) First 5 Terms	(6) First 6 Terms
Panel A: Probability of Graduation (All Students)						
GF	0.0171	0.0134	0.0135	0.0035	0.0085	0.0203
	(0.0126)	(0.0153)	(0.0234)	(0.0323)	(0.0394)	(0.0405)
Control Group Mean	.1945891	.2530396	.3466471	.4336536	.4962465	.5481639
N	45862	34443	22416	17361	13480	11333
Panel B: Probability of Graduation in STEM (Graduates Sub-Sample)						
GF	0.0407	0.0443*	0.0774***	0.0914***	0.0907***	0.1071***
	(0.0251)	(0.0256)	(0.0168)	(0.0151)	(0.0207)	(0.0159)
Control Group Mean	.369863	.3659104	.345152	.3288	.3357664	.3198529
N	6421	6269	5763	5570	5107	4869
Panel C: Probability of Graduation (Never-Repeating Students)						
GF	0.0108	0.0135	0.0259	0.0272	0.0462	0.0596
	(0.0084)	(0.0120)	(0.0230)	(0.0368)	(0.0416)	(0.0413)
Control Group Mean	.1022939	.1508324	.2501782	.361861	.4653846	.5501393
N	30911	20227	10500	7008	4870	3805
Panel D: Probability of Graduation in STEM (Never-Repeating Students; Graduates Sub-Sample)						
GF	0.0485	0.0464	0.0910***	0.0921***	0.1004**	0.1179***
	(0.0315)	(0.0334)	(0.0310)	(0.0299)	(0.0397)	(0.0362)
Control Group Mean	.3667665	.3669291	.3471698	.344898	.3582474	.3373134
N	2431	2364	2177	2104	1920	1814

Notes: This table shows the estimated effects of grade forgiveness on the probability of college completion (Panels A and C) and, conditional graduation, the probability of obtaining a STEM degree (Panels B and D) using Equation (4). Due to dropouts, the sample sizes become progressively smaller as the observation window lengthens. Standard errors are clustered at the entry-semester level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Number of Semesters to Graduation

	(1) First 1 Term	(2) First 2 Terms	(3) First 3 Terms	(4) First 4 Terms	(5) First 5 Terms	(6) First 6 Terms
Panel A: All Graduates (All Students)						
GF	-0.3430*	-0.2665	-0.2458*	-0.1083	-0.0297	0.0957
	(0.1768)	(0.1704)	(0.1409)	(0.1234)	(0.1386)	(0.1187)
Control Group Mean	12.77948	12.76337	12.90741	12.88129	12.94498	12.85798
N	9870	9437	8486	8060	7336	6898
Panel B: Graduates with a STEM Degree (All Students)						
GF	-0.6588***	-0.4780***	-0.6803***	-0.5358***	-0.5218**	-0.4961**
	(0.1621)	(0.1617)	(0.1840)	(0.1930)	(0.2155)	(0.2177)
Control Group Mean	12.84148	12.75835	12.91405	12.80535	12.81677	12.8046
N	2563	2500	2303	2215	2073	1980
Panel C: Graduates with a non-STEM Degree (All Students)						
GF	-0.1095	-0.0672	-0.0406	0.0141	0.1122	0.2628
	(0.1453)	(0.1520)	(0.1516)	(0.1565)	(0.1670)	(0.1737)
Control Group Mean	12.60522	12.59266	12.60552	12.55781	12.58085	12.47027
N	3858	3769	3460	3355	3034	2889
Panel D: All Graduates (Never-Repeating Students)						
GF	-0.4782	-0.3711	-0.2826	-0.1161	-0.1133	-0.0316
	(0.3390)	(0.3116)	(0.2580)	(0.2237)	(0.2241)	(0.1889)
Control Group Mean	10.50202	10.60424	10.90313	11.02857	11.22107	11.23797
N	3783	3574	3124	2928	2640	2440
Panel E: Graduates with a STEM Degree (Never-Repeating Students)						
GF	-0.6007**	-0.3914	-0.6309*	-0.6660**	-0.8766**	-1.2201***
	(0.2700)	(0.2539)	(0.3256)	(0.3107)	(0.3277)	(0.3328)
Control Group Mean	10.66122	10.66953	10.90761	11.05325	11.26619	11.65487
N	919	892	822	796	744	697
Panel F: Graduates with a non-STEM Degree (Never-Repeating Students)						
GF	-0.3375	-0.3670	-0.2407	-0.1184	0.0376	0.1121
	(0.2563)	(0.2677)	(0.2860)	(0.2928)	(0.2838)	(0.2841)
Control Group Mean	10.95272	11.00498	11.01156	10.96573	11.01606	11.03604
N	1512	1472	1355	1308	1176	1117

Notes: This table shows the estimated effects of grade forgiveness on time-to-degree for all graduates (Panels A and D) and graduates with different types of degrees (Panels B-C and E-F) using Equation (4). Due to dropouts, the sample sizes become progressively smaller as the observation window lengthens. Standard errors are clustered at the entry-semester level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Semester Course Load

	Broad Window: 1990-2016	Narrow Window: 1990-2007		
	(1) Full Sample	(2) Full Sample	(3) STEM Majors	(4) non-STEM Majors
Panel A: All Students				
GF	0.0867***	0.1050***	-0.0749**	0.1412***
	(0.0147)	(0.0149)	(0.0344)	(0.0165)
Sample Mean (1995-2000)	3.661425	3.662646	3.959674	3.598302
N	328664	194776	41975	152796
Panel B: Non-Repeating Students				
GF	0.1189***	0.1370***	-0.0327	0.1705***
	(0.0150)	(0.0151)	(0.0355)	(0.0167)
Sample Mean (1995-2000)	3.683127	3.684552	3.989556	3.618727
N	307796	182670	38958	143706
Panel C: Never-Repeating Students				
GF	0.1321***	0.1239***	-0.0557	0.1577***
	(0.0222)	(0.0223)	(0.0581)	(0.0240)
Sample Mean (1995-2000)	3.433563	3.435444	3.890072	3.343495
N	145253	87923	16720	71201

Notes: This table reports the estimated grade forgiveness effects on student semester course load, measured by the number of registered courses within a semester. Model specification is identical to that in column 2 of Table 3 except that course-level covariates are excluded from the regressions. Standard errors are clustered at the student level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Student Effort Expenditure (1990-2007)

	(1) All Students	(2) Never- Repeating Students
Panel A: Initial-Attempt Grade (Per-Course Effort)		
GF	0.0251*	-0.0217*
	(0.0132)	(0.0122)
Sample Mean (1995-2000)	2.6483	2.8310
N	768732	395643
Panel B: Semester Quality Point (Per-Semester Effort)		
GF	0.4725***	0.2611
	(0.1335)	(0.2026)
Sample Mean (1995-2000)	24.6660	24.8132
N	194738	87896
Panel C: Within-Semester Effort Allocation (Standard Deviation)		
GF	0.0026	-0.0034
	(0.0043)	(0.0063)
Sample Mean (1995-2000)	0.6203	0.5341
N	194738	87896
Panel D: Within-Semester Effort Allocation (Max-Min Difference)		
GF	0.0062	0.0008
	(0.0091)	(0.0131)
Sample Mean (1995-2000)	1.3034	1.1014
N	194738	87896

Notes: Based on the model specification in column 2 of Table 3, Panel A additionally controls for course fixed effects and clusters standard errors two-way by course and student. The model specification for Panels B-D is identical to that in column 2 of Table 3 except for course-level covariates. The corresponding standard errors are clustered at the student level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Heterogeneous Effects of Grade Forgiveness on Curriculum Choice

	All Students		Never-repeating Students	
	(1)	(2)	(3)	(4)
	STEM (Classical)	STEM (OPT)	STEM (Classical)	STEM (OPT)
Panel A: Gender (1990-2007)				
Male	0.0221***	0.0271***	0.0127**	0.0163***
	(0.0033)	(0.0034)	(0.0049)	(0.0052)
N	400254	400254	177088	177088
Female	0.0114***	0.0172***	0.0038	0.0079*
	(0.0025)	(0.0029)	(0.0035)	(0.0040)
N	474639	474639	250978	250978
P-value	(0.0043)	(0.0105)	(0.3894)	(0.4408)
Panel B: Income (1990-2007)				
High-income	0.0211***	0.0277***	0.0118***	0.0183***
	(0.0028)	(0.0031)	(0.0040)	(0.0044)
N	460214	460214	225047	225047
Low-income	0.0111***	0.0153***	0.0016	0.0024
	(0.0029)	(0.0032)	(0.0042)	(0.0046)
N	416072	416072	203695	203695
P-value	(0.0075)	(0.0274)	(0.2223)	(0.0557)
Panel C: Race (1997-2007)				
White	0.0121***	0.0078*	0.0085	0.0058
	(0.0037)	(0.0041)	(0.0058)	(0.0066)
N	414831	414831	174338	174338
Non-white	0.0156*	0.0026	0.0012	-0.0106
	(0.0088)	(0.0095)	(0.0147)	(0.0160)
N	84636	84636	32159	32159
P-value	(0.1835)	(0.3552)	(0.9464)	(0.5002)

Notes: The model specification is identical to that in column 2 of Table 3. Standard errors are clustered at the student level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A Proof of the Theoretical Implications

A.1 Course Difficulty

To solve for the optimal course difficulty d under each grading scheme, we first expand each respective utility function and then derive corresponding first order condition. In particular, based on the envelope theorem we take t as given (and vise versa for the derivation of optimal t), along with $E[G]$ and $E[C]$, the expected subsequent-attempt grade and learning cost, when solving for d. We also denote $g(d,t)$ and $c(d,t)$ as g and c, respectively, at times for the sake of brevity.

The expanded utility function under the grade-averaging scheme is:

$$U = f(g)\{d * \frac{E[G] + g(d,t)}{2} - c(d,t) - E[C]\} + [(1 - f(g)]\{d * g(d,t) - c(d,t)\} \tag{5}$$

Taking the partial derivative of the above equation with respect to (w.r.t.) d,

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

and simplifying the equation, we obtain the following first order condition (F.O.C) for the grade-averaging scheme:

$$\begin{aligned}\frac{\partial U}{\partial d} = \frac{1}{2}\{f'(g)g'(d)dE[G] - f'(g)g'(d)dg + f(g)E[G] - f(g)g - f(g)dg'(d)\} \\ - \{f'(g)g'(d)E[C] - g - dg'(d) + c'(d)\} = 0\end{aligned} \tag{7}$$

The expanded utility function for the grade forgiveness scheme is:

$$U = f(g)\{d * E[G] - c(d,t) - E[C]\} + (1 - f(g)\{d * g(d,t) - c(d,t)\} \tag{8}$$

which implies its partial derivative w.r.t. d to be as follows:

$$\frac{\partial U}{\partial d} = f(g) * \{E[G] - c'(d)\} + f'(g)g'(d)\{dE[G] - c - E[C]\}$$
$$- f'(g)g'(d)[dg - c] + (1 - f(g))\{g + dg'(d) - c'(d)\} = 0 \quad (9)$$

Further simplifying the equation results in the F.O.C for the grade forgiveness scheme:

$$\frac{\partial U}{\partial d} = \{f'(g)g'(d)dE[G] - f'(g)g'(d)dg + f(g)E[G] - f(g)g - f(g)dg'(d)\}$$
$$- \{f'(g)g'(d)E[C] - g - dg'(d) + c'(d)\} = 0 \quad (10)$$

Denoting the optimal course difficulty under the grade averaging and forgiveness schemes as d_0^* and d_1^*, respectively, the two F.O.Cs shown above can be expressed 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 $B() = \{f'(g)g'(d)E[C] - g - dg'(d) + c'(d)\}$.

$$\frac{\partial A}{\partial d} = \{f'(g)g'(d)g'(d)d(E[G] - g) + f'(g)g''(d)d(E[G] - g) +$$
$$2f'(g)g'(d)(E[G] - g) - 2f'(g)g'(d)g'(d)d - 2f(g)g'(d) - f(g)dg''(d)\} = 0 \quad (11)$$

$$\frac{\partial B}{\partial d} = f'(g)g'(d)g'(d)E[C] + f'(g)g''(d)E[C] - 2g'(d) - dg''(d) + c''(d) = 0 \quad (12)$$

Assuming that both the expected initial-attempt grade g and learning cost c are twice continuously differentiable, follow the law of diminishing marginal returns, $g'(d) < 0$, $g''(d) < 0$, $f'(g) < 0$, $c'(d) > 0$, $-g''(d) > g'(d)^2$, and that the expected subsequent-attempt grade is higher than the first for a student to be willing to repeat a course, i.e., $E[G] > g(d)$, we can derive that function $A()$ is strictly increasing in d and function $B()$ is strictly increasing in d. As such, we obtain $d_1^* > d_0^*$. Therefore, a typical student will be more likely to attempt a difficult course under grade forgiveness in comparison to the grade averaging scheme.

A.2 Time Allocation to a Given Course

Given that the initial-attempt grade assumes a zero weight under grade forgiveness, it is trivial to prove that the corresponding optimal study time is $t_1^* = 0$. Since the optimal study time under the grade averaging scheme is $t_0^* > 0$, we can conclude that $t_1^* < t_0^*$.

A.3 Time Allocation across Courses

For students who attempt multiple courses in a semester, a choice must be made regarding the allocation of study time among these courses. On this front, we consider the simplest case where a student takes two courses simultaneously at different difficulty levels ($d_1 > d_2$) and allocates time to each course out of a fixed endowment of study time t, i.e., $t_1 + t_2 = t$. The utility function of the student can then be written as the sum of the utility gained from the two courses, though its specific form depends on the prevailing grading policy at the time of the decision-making. Specifically, it is:

$$\begin{aligned} U(g,c,d,t) = f(g_1) * \{d_1 \frac{E[G_1] + g(d_1,t_1)}{2} - c(d_1,t_1) - E[C]\} \\ +\{1 - f(g_1)\} * \{d_1 g(d_1,t_1) - c(d_1,t_1)\} \\ +f(g_2) * \{d_2 \frac{E[G_2] + g(d_2,t_2)}{2} - c(d_2,t_2) - E[C]\} \\ +\{1 - f(g_2)\} * \{d_2 g(d_2,t_2) - c(d_2,t_2)\} \end{aligned} \tag{13}$$

under the grading averaging scheme, and

$$
\begin{aligned}
U(g,c,d,t) = f(g_1) * \{d_1 E[G_1] - c(d_1,t_1) - E[C]\} \\
+\{1 - f(g_1)\} * \{d_1 g(d_1,t_1) - c(d_1,t_1)\} \\
+f(g_2) * \{d_2 E[G_2] - c(d_2,t_2) - E[C]\} \\
+\{1 - f(g_2)\} * \{d_2 g(d_2,t_2) - c(d_2,t_2)\}
\end{aligned}
\tag{14}
$$

under the grade forgiveness scheme. Taking the partial derivative of the first utility function w.r.t. t_1, we obtain the F.O.C for the grade averaging scheme as follows:

$$
\begin{aligned}
\frac{\partial U}{\partial t_1} &= f(g_1) * \{d_1 \frac{g'(t_1)}{2} - c'(t_1)\} + \{1 - f(g_1)\} * \{d_1 g'(t_1) - c'(t_1)\} \\
&+ f'(g) g'(t_1) * \{d_1 \frac{E[G_1] + g(d_1,t_1)}{2} - c(d_1,t_1) - E[C]\} \\
&- \{f'(g) g'(t_1)\} * \{d_1 g(d_1,t_1) - c(d_1,t_1)\} \\
&+ f(g_2) * \{d_2 \frac{-g'(t_1)}{2} + c'(t_1)\} + \{1 - f(g_2)\} * \{-d_2 g'(t_1) + c'(t_1)\} \\
&- f'(g) g'(t_1) * \{d_2 \frac{E[G_2] + g(d_2,t_1)}{2} - c(d_2, t - t_1) - E[C]\} \\
&+ \{f'(g) g'(t_1)\} * \{d_2 g(d_2, t - t_1) - c(d_2, t - t_1)\} \\
&= 0
\end{aligned}
\tag{15}
$$

which can be simplified as:

$$
\begin{aligned}
\frac{\partial U}{\partial t_1} &= \{-d_1 \frac{f(g_1) g'(t_1)}{2} + d_1 g'(t_1)\} + \{d_2 \frac{f(g_2) g'(t_1)}{2} - d_2 g'(t_1)\} \\
&+ \{d_1/2 * f'(g) g'(t_1) E[G_1] - d_2/2 * f'(g) g'(t_1) g(d_1,t_1)\} \\
&+ \{-d_2/2 * f'(g) g'(t_1) E[G_2] + d_2/2 * f'(g) g'(t_1) g(d_2,t_1)\} \\
&= 0
\end{aligned}
\tag{16}
$$

Following the same logic, we obtain the F.O.C. of the grade forgiveness scheme as follows:

$$
\begin{aligned}
\frac{\partial U}{\partial t_1} &= -f(g_1) * \{c'(t_1)\} + \{1 - f(g_1)\} * \{d_1 g'(t_1) - c'(t_1)\} \\
&\quad + f'(g) g'(t_1) * \{d_1 E[G_1] - c(d_1, t_1) - E[C]\} \\
&\quad + \{-f'(g) g'(t_1)\} * \{d_1 g(d_1, t_1) - c(d_1, t_1)\} \\
&\quad + f(g_2) * \{c'(t_1)\} + \{1 - f(g_2)\} * \{-d_2 g'(t_1) + c'(t_1)\} \\
&\quad - f'(g) g'(t_1) * \{d_2 E[G_2] - c(d_2, t - t_1) - E[C]\} \\
&\quad + \{f'(g) g'(t_1)\} * \{d_2 g(d_2, t - t_1) - c(d_2, t - t_1)\} \\
&\quad = 0
\end{aligned}
\tag{17}
$$

which can be simplified as:

$$
\begin{aligned}
\frac{\partial U}{\partial t_1} &= \{-d_1 f(g_1) g'(t_1) + d_1 g'(t_1)\} + \{+d_2 f(g_2) g'(t_1) - d_2 g'(t_1)\} \\
&\quad + \{d_1 * f'(g) g'(t_1) E[G_1] - d_2 * f'(g) g'(t_1) g(d_1, t_1)\} \\
&\quad \{-d_2 * f'(g) g'(t_1) E[G_2] + d_2 * f'(g) g'(t_1) g(d_2, t_1)\} \\
&\quad = 0
\end{aligned}
\tag{18}
$$

Re-arranging (15) and (17) results in:

$$
\begin{aligned}
&\frac{1}{2}\{-d_1 f(g_1) + d_1 f'(g) E[G_1] - d_1 f'(g) g(d_1, t_1) \\
&+ d_2 f(g_2) - d_2 f'(g) E[G_2] + d_2 f'(g) g(d_2, t_2)\} \\
&= \frac{1}{2} LHS(t_1^A) = [d_2 - d_1]
\end{aligned}
\tag{19}
$$

and

$$
\begin{aligned}
&\{-d_1 f(g_1) + d_1 f'(g) E[G_1] - d_1 f'(g) g(d_1, t_1) \\
&+ d_2 f(g_2) - d_2 f'(g) E[G_2] + d_2 f'(g) g(d_2, t_2)\} \\
&= LHS(t_1^F) = [d_2 - d_1]
\end{aligned}
\tag{20}
$$

Comparing (18) and (19), we find $\frac{1}{2} LHS(t_1^A) = LHS(t_1^F)$, where t_1^A and t_1^F are the optimal time allocation to course 1 t_1 under the averaging (A) and forgiveness (F) schemes, respectively. Assuming $d_1 > d_2$, $g'(t) > 0$, $g''(t) < 0$, $E[G_1] > g(d_1, t_1)$, and $E[G_2] > g(d_2, t_2)$, it can be shown that $LHS(t_1)$ is decreasing in t_1. Hence, we can conjecture $t_1^A < t_1^F$. If we define the gap between the time allocated to each individual course as $t_1 - t_2 = t_1 - (t - t_1) = 2t_1 - t$ (provided that $t_2 = t - t_1$), we can conclude that the gap will increase as t_1 increases.

A.4 Probability of Repetition and Threshold Grade

We next derive the probability of course repetition and threshold grade for the student who have observed their initial-attempt grades g and learning cost c of a given course. Taking the difficulty level and study time for the course as given for brevity, we can formulate the utility function of a typical student as follows:

$$U(g,c) = \begin{cases} g - c, & \text{if Does Not Repeat} \\ \frac{E[G]+g}{2} - c - E[C] & \text{if Repeats under Grade Averaging} \\ E[G] - c - E[C], & \text{if Repeats under Grade Forgiveness} \end{cases} \tag{21}$$

Then the student will choose to repeat the course if the utility of repeating is greater than that of not repeating:

$$U^R(g,c) > U^{NR}(g,c) \equiv \begin{cases} \frac{E[G]+g}{2} - c - E[C] > g - c, & \text{if Forgiveness = 0} \\ E[G] - c - E[C] > g - c, & \text{if Forgiveness = 1} \end{cases} \tag{22}$$

where U^R and U^{NR} represent the utility derived from repeating and not repeating the course, respectively.

If we further denote the realized initial-attempt grades under the grade averaging and forgiveness schemes as g_0 and g_1, respectively, the above inequalities can be simplified as:

$$U^R(g,c) > U^{NR}(g,c) \equiv \begin{cases} E[G] - 2E[C] > g_0, & \text{Under Grade Averaging} \\ E[G] - E[C] > g_1, & \text{Under Grade Forgiveness} \end{cases} \tag{23}$$

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

$Pr(E[G] - 2E[C] > g_0)$ and $Pr(E[G] - 2E[C] = g_0) = 0$ under grade averaging;

$Pr(E[G] - E[C] > g_1)$ and $Pr(E[G] - E[C] = g_1) = 0$ under grade forgiveness.

Provided that the belief on the expected grade $E[G]$ and the expected cost $E[C]$ is unchanged, we can easily obtain the following propositions:

(1) If the initial-attempt grades are constant under different grading schemes, $g_0 = g_1 = g$, the region of grades for one to prefer repeating over not is greater under grade forgiveness than under grade averaging: $E[C] < E[G] - g < 2E[C]$. Thus, the average probability of repetition under grade forgiveness will be higher than that under grade averaging.

(2) The threshold (highest) grade to repeat under grade forgiveness is higher than the threshold (highest) grade to repeat under grade averaging: $g_1 > g_0$, and the difference between the two threshold grades is restricted as $g_1 - g_0 \leq E[C]$.

B Additional Tables

Table B1: Students in Pass/Fail Courses as the Control Group

	(1) All Students	(2) Non-repeating Students	(3) Never-repeating Students
Panel A: Probability of Repetition			
Graded $\times$ *GF*	0.0171***		
	(0.0007)		
Sample Mean (1995-2000)	0.0230		
N	914044		
Panel B: Probability of Taking a STEM Course (Conservative)			
Graded $\times$ *GF*	0.0259***	0.0387***	0.0158***
	(0.0017)	(0.0024)	(0.0022)
Sample Mean (1995-2000)	0.1955	0.1990	0.1806
N	986951	914516	494115
Panel C: Probability of Taking a STEM Course (OPT)			
Graded $\times$ *GF*	0.0350***	0.0512***	0.0245***
	(0.0020)	(0.0029)	(0.0027)
Sample Mean (1995-2000)	0.2505	0.2550	0.2372
N	986951	914516	494115

Notes: This table reports the estimated differential grade forgiveness effects on the likelihood of course repetition (Panel A) and attempting a STEM course (Panels B-C) between graded and Pass/Fail courses offered during the narrow window of 1991-2007 using Equation (2). The covariates included in the models are identical to those in columns 4-5 of Table 2. Standard errors are clustered at the student-course-section level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Choice of Course Difficulty (Grading Harshness)

	Broad Window: 1990-2017		Narrow Window: 1990-2008	
	(1) Grade	(2) %DF	(3) Grade	(4) %DF
Panel A: All Students				
GF	-0.0669***	0.0204***	-0.0605***	0.0174***
	(0.0088)	(0.0013)	(0.0089)	(0.0013)
Sample Mean (1995-2000)	7.6471	0.4137	7.6471	0.4137
N	1162669	1162669	830499	830499
Panel B: Non-Repeating Students				
GF	-0.0598***	0.0154***	-0.0522***	0.0123***
	(0.0090)	(0.0013)	(0.0091)	(0.0013)
Sample Mean (1995-2000)	7.6592	0.4067	7.6592	0.4067
N	1106798	1106798	795114	795114
Panel C: Never-Repeating Students				
GF	-0.0587***	0.0108***	-0.0390***	0.0078***
	(0.0136)	(0.0018)	(0.0137)	(0.0018)
Sample Mean (1995-2000)	7.7027	0.4010	7.7027	0.4010
N	547040	547040	409866	409866

Notes: This table reports the estimated grade forgiveness effects on student choice of course difficulty measured by the grading harshness of a given course. More details on the construction of this measure and relevant samples can be found in Section 5.2.2. Model specification is identical to that in column 2 of Table 3. Standard errors are clustered at the student level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B3: STEM Degree Attainment across Gender

	(1) First 1 Term	(2) First 2 Terms	(3) First 3 Terms	(4) First 4 Terms	(5) First 5 Terms	(6) First 6 Terms
Panel A: All Students (Male)						
GF	0.0249	0.0267	0.0728*	0.0918**	0.0965**	0.1137***
	(0.0444)	(0.0453)	(0.0369)	(0.0351)	(0.0400)	(0.0416)
Sample Mean (Control Group)	.3681592	.3626667	.3204047	.296846	.2874396	.2784091
Observations	2836	2759	2535	2452	2256	2161
Panel B: All Students (Female)						
GF	0.0506*	0.0546*	0.0761***	0.0858***	0.0815**	0.0966***
	(0.0279)	(0.0285)	(0.0278)	(0.0273)	(0.0365)	(0.0315)
Sample Mean (Control Group)	.2262488	.2239422	.2116603	.2011252	.2055046	.200431
Observations	3585	3510	3228	3118	2851	2708
Panel C: Never-Repeating Students (Male)						
GF	0.0108	0.0023	0.0414	0.0429	0.0790	0.0989
	(0.0542)	(0.0563)	(0.0616)	(0.0616)	(0.0677)	(0.0712)
Sample Mean (Control Group)	.3671875	.3640167	.3398058	.328125	.3096774	.2941176
Observations	916	882	819	792	725	690
Panel D: Never-Repeating Students (Famale)						
GF	0.0677*	0.0663*	0.1146***	0.1152***	0.1078*	0.1232**
	(0.0384)	(0.0390)	(0.0383)	(0.0381)	(0.0540)	(0.0527)
Sample Mean (Control Group)	.1917476	.1919192	.1666667	.1711409	.1802575	.1708543
Observations	1515	1482	1358	1312	1195	1124

Notes: This table shows the STEM degree attainment effects of grade forgiveness for male (Panels A and C) and female (Panels B and D) students using Equation (4). Due to dropouts, the sample sizes become progressively smaller as the observation window lengthens. Standard errors are clustered at the entry-semester level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B4: STEM Degree Attainment across Socioeconomic Background

	(1) First 1 Term	(2) First 2 Terms	(3) First 3 Terms	(4) First 4 Terms	(5) First 5 Terms	(6) First 6 Terms
Panel A: All Students (High Income)						
GF	0.0666**	0.0670**	0.0985***	0.1173***	0.1195***	0.1400***
	(0.0283)	(0.0297)	(0.0232)	(0.0207)	(0.0209)	(0.0224)
Sample Mean (Control Group)	.2804348	.278481	.2560113	.2362205	.2323651	.2289157
Observations	3530	3454	3207	3099	2860	2744
Panel B: All Students (Low Income)						
GF	0.0116	0.0182	0.0516**	0.0613***	0.0533*	0.0633**
	(0.0280)	(0.0270)	(0.0212)	(0.0210)	(0.0312)	(0.0255)
Sample Mean (Control Group)	.2972376	.2905882	.2607407	.2487805	.2494759	.2394015
Observations	2891	2815	2556	2471	2247	2125
Panel C: Never-Repeating Students (High Income)						
GF	0.0701*	0.0665	0.1174***	0.1308***	0.1351***	0.1475***
	(0.0407)	(0.0418)	(0.0337)	(0.0317)	(0.0357)	(0.0384)
Sample Mean (Control Group)	.2307692	.2291667	.2055749	.1969697	.2047619	.2032086
Observations	1350	1317	1227	1181	1082	1031
Panel D: Never-Repeating Students (Low Income)						
GF	0.0270	0.0259	0.0619	0.0502	0.0616	0.0825
	(0.0402)	(0.0423)	(0.0495)	(0.0466)	(0.0636)	(0.0585)
Sample Mean (Control Group)	.2902208	.2876254	.2674897	.2743363	.2640449	.2432432
Observations	1081	1047	950	923	838	783

Notes: This table shows the STEM degree attainment effects of grade forgiveness for high (Panels A and C) and low-income students (Panels B and D) using Equation (4). Due to dropouts, the sample sizes become progressively smaller as the observation window lengthens. Standard errors are clustered at the entry-semester level.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$