I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16690

Learning from Mistakes: The Implications of Course Repetition for Student Subsequent Success

Kelly Chen
Xuan Jiang

DECEMBER 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16690

Learning from Mistakes: The Implications of Course Repetition for Student Subsequent Success

Kelly Chen
Boise State University

Xuan Jiang
Jinan University and IZA

DECEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Learning from Mistakes: The Implications of Course Repetition for Student Subsequent Success*

Most colleges allow low-performing students to make a repeated attempt for the same course, but little is known about its implications for the academic success of these students. Using the variations in repetition induced by the cancellation and reversal of a university GPA policy to correct for student selection, we quantify the effects of course repetition on below-average students' subsequent outcomes. We find that students develop greater interest, persist longer, and perform better in a given subject upon repetition in comparison to their non-repeating classmates who receive the same initial-attempt grade. The observed repetition effects are particularly pronounced for the students who are exposed to the college environment and/or a subject matter for the first time and are entirely explained by the gains in learning. Importantly, while boosting graduation rates, a moderate number of repetitions during a student's undergraduate career is not found to cause any disruptions to the student's routine progress in pursuing a degree.

JEL Classification: I23, I24, J24

Keywords: course repetition, grade replacement, college student success, graduation, time-to-degree

Corresponding author:
Kelly Chen
Boise State University
1910 W University Dr, Boise
ID 83725
USA

E-mail: kellychen@boisestate.edu

* All opinions expressed, and all errors and omissions, are entirely the authors'. This paper uses confidential data requested from several offices of Boise State University. The data can be obtained by signing a nondisclosure agreement with the university and the authors are willing to assist (Kelly Chen, kelly-chen@boisestate.edu). The authors declare no conflict of interest.

1. Introduction

Course repetition is a common practice in higher education to help primarily low-performing students raise their grade-point averages (GPA), fulfill a major requirement, strengthen their knowledge of a particular area, and prepare for future classes, though the extent to which repeat options are available to students vary widely across institutions. For example, recent survey data (Advancing Global Higher Education, 2015) suggest that while 80% of the undergraduate institutions around the world did not restrict the number of courses/credits that can be repeated, over 70% of them imposed an upper limit of the original grade earned for a course to be repeatable. Out of this total, 21% of the institutions allowed students to repeat a course if an initial grade of D or below (including pluses and minuses) is awarded, 60% permitted repetition for an initial grade of C or below, and 16% B or below. Furthermore, upon successful completion of the course, it was common for the most recent grade to be factored into student cumulative GPAs (54%), despite the fact that a considerable number of institutions also opted for alternative weighting schemes by allowing the highest grade earned to apply to the student's record (40%) or taking both initial and subsequent grade(s) into account (11%).

The varied qualifications and requirements of course repetition policies raise the question of how much is gained by permitting students to make a repeated attempt for the same course, especially in the context where budgetary constrictions may potentially impact course repeatability for students across public institutions.[1] On the one hand, repeating courses can be an effective strategy to address the academic deficiencies of under-prepared students, if the repeated exposure to the same material allows them to rediscover the meaning of the course topic, connect what they learn to their own values at a deeper level, and feel more competent about the subject matter. The multiple opportunities and ways to interact with information and learning may be particularly valuable for students who lack self-discipline and/or the study

[1] For example, amid a financial crisis California Community Colleges sought approval from the Board of Governors in 2012 to prohibit students within their 112-college system from retaking some credit-bearing courses they have already passed (`https://www.insidehighered.com/news/2012/05/08/california-community-colleges-cut-back-course-repeating`). Community colleges in Michigan and New York also considered or started to put limitations on repeated courses as a means to save taxpayer dollars (`https://www.communitycollegereview.com/blog/the-attack-on-repeating-classes-heed-these-warnings.`). Accessed on June 7, 2023.

habits required to match the demands of college work (i.e., first-year students) and for students who are exposed to a subject matter for the first time, where there is little or no existing knowledge to attach to the new course content (i.e., students in introductory-level courses). On the other hand, a generous repeat option can disincentivize students from working hard the first time and negatively impact outcomes such as time-to-degree and college completion. Considering that individual students have different resources and time constraints, the additional time that repeating courses adds to a degree program can reinforce existing social and economic inequalities (Jewell and Tieslau 2013; Marx and Meeler 2013). Finally, allowing students to repeat may backlog seats for courses and disrupt the flow of students in and out of a university, which could raise additional concerns over equity and equal opportunity when there is limited access to such courses or programs (Casas and Meaghan; 1996).

This theoretical ambiguity has not been resolved in empirical literature. Most existing studies on course repetition at the post-secondary level focus on mandatory programs such as remedial education (Bettinger and Long 2008; Martorell and McFarlin 2011; De Paola and Scoppa 2014; Scott-Clayton et al. 2014) and grade retention (Tafreschi and Thiemann 2016)[2] and therefore are unable to isolate the effect of course repetition from that of the program itself. It is conceivable that repeating even the same courses may bear distinct implications for students when the student chooses to repeat, instead of being required to repeat. First, there may be less socio-emotional and adjustment difficulties if the timing and/or course-topic of the repetition are determined through students' own perceived merits. Second, repeating one rather than a set of courses at once (e.g., first-year courses) may be less time consuming and less likely to impede repeating students' progress and their completion of their degree.[3] Third, relative to students in remedial education, those who take college-level courses tend to have a higher level of scholarly ability. Hence, whether or not course repetition can help students outside of the remedial or grade retention programs to achieve

[2]The former places students who fall short on placement tests prior to college entry such as ACT and SAT to below-college-level courses in achieving expected competencies in core academic skills such as English, reading, and math, whereas the latter requires students with subpar performance during college to repeat a semester or a year of study before continuing with their college work.

[3]For example, dropout rate is found to be negatively affected by grade retention in Tafreschi and Thiemann (2016).

academic success remains largely unknown.

From this perspective, this study offers the first quasi-experimental evidence on whether and how a voluntary repetition of a college-level course may affect the future success of repeating students across various stages of study, disciplines, and socioeconomic backgrounds. The two existing studies that examine repeaters' subsequent performance focus on students in a specific type of courses (i.e., introductory finance, see Biktimirov and Armstrong 2015) or students of a certain socioeconomic group (i.e., low-income/Pell-eligible, see Sovero and Griffith 2023), so it is unclear how their findings extend to the general population of students and whether the shorter-term effects of repetition persist in the longer term.

Granted, a straightforward comparison between repeating and non-repeating students could lead to biased estimates of the repetition effect. The bias in part stems from the fact that students who choose to retake a course tend to have lower achievement and perceptions of competence, and therefore, are also less likely find meaning and value in their courses, develop greater interest, perform better, persist longer, and complete their degree programs relative to their non-repeating peers. To address the potential role of student selection, we exploit the plausibly exogenous variations in the repetition decisions of students that result from the cancellation and reintroduction of a GPA policy at a four-year public institution, Boise State University (BSU). Before academic years 1995 and after 2001,[4] BSU adopted a grade replacement formula to calculate repeating students' cumulative GPAs, whereas between 1995 and 2001, a grade averaging formula was employed instead. The difference between the two formulas is whether the initial grades of the repeating students are excluded from, as opposed to being averaged into, the students' overall GPAs. We argue that the alternation between these two grading schemes creates a unique opportunity to identify the effect of course repetition, as grade replacement (versus grade averaging) generates a differential incentive for students to repeat a course based on their initial performance: all else equal, the lower the grade a student receives on his/her initial attempt, the more

[4] During the observation period, BSU utilized a semester system, where each course is 15 weeks in length. While summer courses are typically shorter, often just three, five, or seven weeks long, we include all summer sessions in this analysis to provide an accurate characterization of student course-taking behavior in college.

incentivized the student will be to repeat a given course, since the expected increase in the final quality point of the course is inversely proportional to the student's initial grade.[5]

Figure 1 illustrates this idea through the raw differences in the course repetition rate between C and DF students over time. The former refers to students who received a C (including C- and C+ and thereafter) on their initial attempt and the latter refers to students who received an F or D as their initial grade. As shown, the repetition rate stayed consistently low for C students over the observation period, as indicated by the hollow dot line at around 0.01. The comparable figure for DF students over the same period, however, displays two distinct structural breaks while hovering in the range of 0.20 to 0.35, a sudden drop of approximately 15 ppt in 1995 and a sudden increase of approximately 5 ppt in 2001. Importantly, the timing of these structural breaks coincides with the timing of the switch in GPA formula: the first one corresponds to the year when the replacement formula was abolished, and the second one when grade replacement was re-introduced. If there is no reason to believe that the trajectory of the repetition rate to be any different between these two groups of students, then this figure suggests that the grade replacement formula may have provided a powerful and sustainable incentive for low-performing students to repeat a course during our observation period.

Our main analysis exploits the differential effects of grade replacement across the students who receive different first-attempt grades as a source of exogeneity and tests whether the changes in the outcome gap between repeating students and their non-repeating classmates co-vary with the enactment of grade replacement.[6] We do so through two alternative methods: an event study analysis and an instrumental variable (IV) approach. To ensure a sufficiently strong first-stage estimation for the latter,

[5]For example, suppose that there are two students who received a C (or a grade point of 2) and a F (or a grade point of 0), respectively, from the same course-section upon initial enrollment and that they both received a B (or a grade point of 3) on their subsequent attempt. The replacement formula would provide the F student with a greater boost in the final quality point than the C student (1.5 vs 0.5) relative to the averaging formula. Following the same logic, it can be demonstrated that the lower the initial grade, the greater gain a repeating student would anticipate under grade replacement compared to grade averaging.

[6]Since very few students repeated a course more than once during the observation period (i.e., 1%; see Table 1), our estimates mingle the first-time with multiple-time repetition effects on student outcomes to maximize sample size, though excluding the latter produces highly similar results.

we focus on the likely beneficiaries of the replacement formula, students who receive below-average grades on their first attempt (i.e., a grade of C or below) and constitute the vast majority of the repeaters in our context (96%; see Table 1). By omitting any favorable impact of course repetition on the relatively high-performing students, our identification strategy will likely underestimate the true effect of course repetition, though it allows us to avoid the danger that our results will be biased upward, given that the differences in outcome can be a result of any concurrent policy or time-varying hidden characteristics of the students, the courses they enroll in, and the instructors who teach these courses that jointly affect the students' likelihood to repeat and subsequent decisions.

In addition, since the identifying assumption of our empirical strategy – that any pre-treatment differences between students of different first-attempt performance would have continued on the same trends absent grade replacement – may be violated when students of different initial performance sort into different academic paths without grade replacement, we conduct a falsification test by examining students who were enrolled in Pass/Fail courses offered at the same time. While subjected to the same teaching and grading practices of the instructors and other academic policies at various levels (e.g., university, college, and department), the outcomes of these students should not be affected by the adoption of grade replacement, as repeating students' new grades will replace the old ones no matter which formula is enacted[7] and thus can serve as a potentially valid counterfactual for those obtained from the graded courses offered simultaneously.

In the short run, we find that a one-time repetition of a college-level course fosters the subsequent interest and success of repeating students in the field – as measured by the number of courses they enroll in and complete in a given subject, the difficulty level of the follow-on courses they attempt and pass, and their performance in the next courses taken. These favorable effects are entirely explained by learning gains made through repetition and are driven, to a large extent, by students who are ex-

[7] For example, a student who chose to repeat a Pass/Fail course in which she initially failed and subsequently passed, the Pass grade would replace the previous Fail in the student's cumulative GPA under both formulas.

posed to the college environment and/or a subject matter for the first time, that is, first-year students and those taking introductory-level courses. In the long run, the improved persistence and performance further contributes to the successful completion of a degree program, even though the gains from repetition diminish with the number of repeated courses and can cause delays in graduation when the repetition becomes excessive (i.e., five or more courses during a student's degree program).

Through this, we make the following principal contributions to existing literature: First, by considering the outcomes observed immediately upon repetition and those that emerge up to 10 years after, this study offers the first analysis of the role of course repetition in both the retention and graduation among low-performing students across disciplines and socioeconomic groups. Second, we demonstrate, for the first time in the literature, how tabulating the subsequent-attempt grade into cumulative GPA can have a profound effect on repeating students' academic success.[8] Third, our mechanism investigation contributes to a broader economics literature on the gains in deep learning through repetition among high school students who take college entrance exams (Vigdor and Clotfelter 2003; Frisancho et al. 2016; Goodman et al. 2020) and first-year undergraduates in grade retention programs (Tafreschi and Thiemann 2016). Our estimates of the effects of repetition on first-year students square well with the previously published estimates for students in grade retention programs (Tafreschi and Thiemann, 2016) and corroborate others that find experiences during one's freshmen years and/or introductory courses to be an important determinant of a student's subsequent interest and achievement (e.g., Hoffmann and Oreopoulos 2009; Carrell and West 2010; Fournier and Sass 2000; Figlio et al. 2015).

The remainder of this paper is organized as follows. Sections 2 and 3 describe the data used in the analyses and outline our main research methodology. Section 4 presents the estimated effects of course repetition on course-related outcomes such as subsequent student interest and performance. Section 5 focuses on graduation-related outcomes, such as study pace, time-to-degree, and college completion. Section 6 explores underlying mechanisms and Section 7 concludes.

[8]Using the same data but different identification strategies, Jiang et al. (2023) explores the impact of the grade replacement formula on non-repeating students' risk-taking behavior.

2. Background and Data

2.1 Institutional Context

BSU is a four-year public university located in the northwest United States with an undergraduate population of approximately 22,000. It currently has the largest undergraduate enrollment in the state of Idaho and offers nearly 80 bachelor's degrees across seven colleges: Arts & Sciences, Business & Economics, Education, Engineering, Health Sciences, and the School of Public Service.

BSU was one of the earliest adopters of the grade replacement formula – which is also commonly referred to as the grade forgiveness policy – among four-year institutions (for more details see Jiang et al. 2023). This relatively lenient repetition policy was first introduced prior to the 1970s and remained effective until 1995, when the institution switched to a grading-averaging formula in an attempt to "raise academic standards," even though BSU reverted back to the grade replacement scheme only six years later primarily out of peer pressure, given that most other colleges in the state of Idaho had implemented the grade replacement formula at the time.[9]

During the observation period, course repetition was free at BSU for full-time students, provided that the students did not enroll in courses for a total number of credit hours in excess of a full course load. There was a gradual increase in the sticker price for each credit hour for overload and/or part-time students (from \$61 to \$297), but no difference in the tuition charged of in-state versus out-of-state residents. Importantly, with the exception of the weighting formula used for GPA calculation, most other parameters of the repetition policy stayed unchanged. For example, both high and low-performing students could choose to repeat a course at any stage of study, provided that the student received a grade upon initial enrollment and that space was available at the time of repetition. After a given course is repeated, both new and old grades would appear on the students' transcripts. While an overall maximum of six and a per-course maximum of two to three times were imposed at one point to limit

[9]According to the meeting minutes of the Academic Standard Committee accessible to us, it was believed that the grade averaging scheme "[had] 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."

the number of courses that a student could repeat and the number of times the student could take the same course, many exceptions to the repeat count were allowed.[10],[11] Students were also 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. The lack of restrictions on course repeatability results in considerable heterogeneity in the observed repetition behavior over time, constituting the essential identifying variations for current analysis.

2.2 Data

In carrying out our analyses, we utilize the admission records and transcripts of 196,812 students who enrolled in the undergraduate courses at BSU between 1990 and 2016. We exclude entering cohorts after 2015 to ensure that each repeating student in the sample will be tracked for at least two years after course repetition. Given that course repetition always occurs after one's initial attempt, this implies a longer observation window for repeating students' non-repeating peers, though the length of this depends on the timing of the repetition. While maximizing sample size and time horizon for our event study, this practice may potentially bias our estimates if students in later cohorts do not take subsequent courses immediately after our observation. In a robustness check, we thus restrict our analyses to pre-2007 entering cohorts, so that each repeating student will be tracked for a minimum of 10 years (see Section 4.3).[12]

In addition, given that the GPA computation at BSU excludes credits for Pass/Fail courses and courses registered for but later dropped, unless otherwise stated (e.g., Section 4.3) our analysis sample is restricted to students who received a letter grade

[10]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.

[11]More specifically, students were allowed to repeat as many courses as possible until 2013, when a cap of six was imposed. Besides overall maximum, an individual maximum was also imposed to limit enrollment in the same course to three times after 1995 and then two times after 2015. Despite the lenient nature of the policy, excessive repetition was rare (also see Table 1). 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.

[12]This threshold is chosen based on the fact that only 15.3% of first-time undergraduates finished their bachelors' degrees within six years during our observation period. Replicating the analysis for pre-2011 entering cohorts where each repeating student is tracked for a minimum of six years yields highly similar results (available upon request).

(i.e., A-F) in their initial enrollment and students who attended a regular course that utilized a graded system, though students who were awarded grades of P or F from a Pass/Fail course or W (or its equivalent such as CW) in their initial attempt are included in our calculation of subsequent outcomes, when possible.[13] To ensure maximum compatibility, nontraditional courses such as exercise sections and tutorials, labs, studies abroad and on satellite campuses, concurrent enrollment, professional education, and zero-credit courses are also omitted, as are the observations from the first year of observation (i.e., 1990) for the construction of the historical GPA measures used in the analyses.[14]

In the end, our main analysis sample contains 2,843 courses offered across 240 academic subjects at BSU from 1991 through 2016. For all students, we observe the courses they enrolled in, the grades they received, all subsequent enrollment decisions they made within the university, along with their repetition attempt(s), outcomes of the repeated attempt(s), and some basic demographic characteristics (i.e., gender and transfer status). Beginning in 1998, additional information also becomes available when BSU transitioned from paper recording to a centralized digital archiving system, including the specific course-section that the students attended and some other demographic characteristics of the students that were not previously accessible to us including ethnicity, age at college entry, and in-state status. Thus, for each student in our post-1998 sample, a relatively rich set of characteristics can be identified not only for the student and his/her peers who attended the same course-section, but also for the instructor who taught a given course-section.

2.3 Summary Statistics

Summary statistics for all students (columns 1-2), students who received a grade of C, D, or F on their first attempt (columns 3-4), and the CDF students observed after 1998 (columns 5-6) are reported in Table 1. In terms of student-course-section-level characteristics (Panel A of Table 1), we see a highly similar pattern across samples,

[13] In other words, whether a student attempts and persists to the end of a Pass/Fail course upon repetition is considered in our analysis of course choice but is not included in our analysis of subsequent performance, as grades received from Pass/Fail courses are not assigned any quality points and therefore become indistinguishable on 0 to 4 scale.

[14] Utilizing the first five years of observations for this purpose results in a limited time horizon for our event analysis but generates similar results (available upon request).

implying that the vast majority of repeaters were CDF students and that this trend changed little over time. This conjecture is confirmed by the repetition rates across students of different first-attempt performance reported in the first five rows of column 1: 96% of repeating students had an initial grade of C or below and out of this total, 92% earned a grade of D or below. In addition, repeating students consistently had poorer academic performance than their non-repeating peers, as measured by their GPAs both from the most recent semester and from all coursework they had attempted/cumulative prior to retaking a given course. Students during the first year of study were less likely to repeat a course than their more senior counterparts, though introductory level or lower-division courses were the more repeated ones relative to upper-division level courses. Noticeably, the vast majority of the repeating students under study were first-time repeaters (99%) who chose to repeat less than two courses (70%) during their entire undergraduate program(s). About one in five repeated three to four courses (21%) and one in ten repeated more than four courses (10%), despite the fact that additional attempt(s) were accommodated by the university's policy (also see footnote 10).

Switching attention to student-level characteristics (Panel B of Table 1) suggests that repeating students were more likely to be male, non-white, in-state, young students who entered college under the age of 21, and students who had declared a major at the time of course repetition, though the pattern is somewhat mixed in terms of their transfer status.

Looking at measures for subsequent outcomes, we see a clear gap between repeating and non-repeating (Panel C): repeating students were much less likely to attempt a course (i.e., receiving a grade of W or equivalent in their initial enrollment) and complete an extra credit (i.e., receiving a passing grade of D- or above). Even though there are slight differences in the difficulty level of the next course they attempted and passed in a given subject, conditional on enrollment they performed worse compared to their non-repeating peers. These observed gaps in future outcomes, again, are highly consistent across samples.

3. Empirical Framework

3.1 Difference-in-Differences Approach

We begin our formal analysis for student-course-section-level outcomes (i.e., non-graduation-related outcomes) by implementing a standard DD event study that exploits the adoption of the grade replacement formula as a source of exogeneity.[15] Specifically, we treat the abolition and reintroduction of grade replacement as two separate events that potentially alter repetition behavior in opposite directions. We use the grade a student receives on his or her first attempt to approximate the intensity of treatment, while restricting attention to below average or CDF students. By mapping the timing of repetition to the years when grade replacement was cancelled and reversed, respectively, we estimate the following model for the first-stage outcome (i.e., probability of course repetition) and some second-stage outcomes, including measures for student persistence/enrollment and subsequent course choice:

$$\begin{aligned} Y_{jsa,t\ or\ (t+1)} = & \sum_{k=-7}^{-10+} \beta_k GP_{jsa,t-1} \mathbb{1}(Year_t - 2001 = k) \\ & + \sum_{k=-6}^{-2} \gamma_k GP_{jsa,t-1} \mathbb{1}(Year_t - 2001 = k) \\ & + \sum_{k=0}^{14+} \delta_k GP_{jsa,t-1} \mathbb{1}(Year_t - 2001 = k) + \mathbb{1}(GP_{jsa,t-1})' T \\ & + \kappa_{s,t-1} + \lambda_{a,t-1(t)} + \Gamma'_{jsa,t(t-1)} S + \epsilon_{jsa,t(t-1)} \end{aligned} \tag{1}$$

Here $Y_{jsa,t}$ and $Y_{jsa,t+1}$ represent the probability of course repetition $Repeat_{jsa,t}$ and subsequent outcomes related to enrollment and course choice $Enrollment_{jsa,t+1}$ for student j observed at semester-year t and $(t+1)$, respectively, upon the initial enrollment in a given course-section s at $(t-1)$, respectively. The academic progress or class standing at the time of observation is denoted by a, which implies the time of repetition (i.e., t) for a repeating student and the time of first attempt (i.e., $t-1$) for a non-repeating student. Hence, $\kappa_{s,t-1}$ and $\lambda_{a,t-1(t)}$ represent the course-section fixed effects and academic-progress fixed effects, respectively. The former $\kappa_{s,t-1}$ holds

[15] A variant of the empirical strategy described in Section 3 is used to analyze student-level outcomes of the study, including college completion and time-to-degree. The details are provided in Section 5.

constant the potential impact of any time-varying factors that are specific to a course-section, such as the number of students enrolled in the section, the composition of student quality, teacher effectiveness, number of meeting frequency, and time-of-day and day-of-week of the course-section. Since the digitized data prior to 1998 does not permit us to identify individual course-sections (for more details see Section 2.2), we use a set of course-by-semester-year dummies as proxies in the analyses involving pre-1998 data. This is undoubtedly an imperfect solution for courses with multiple sections offered in one semester. We thus replicate our analyses with post-1998 samples whenever possible to cross-check our results. As demonstrated, the choice of course-section identifiers have a negligible impact on our main findings. The latter $\lambda_{a,t-1(t)}$ is approximated by the number of semesters elapsed since a student's initial enrollment at the time of observation. In other words, this measure captures the amount of physical time spent at BSU when a repeating student repeats a course or when his or her same-course-section non-repeating peers make their first attempt. The inclusion of $\lambda_{a,t-1(t)}$ is important for our purposes as repeating students tend to be more academically advanced and biologically more mature at the time of repetition relative to their non-repeating peers, that if unaccounted for, may contaminate /upward bias our estimates.[16]

$GP_{jsa,t-1}$ is continuous measure for a student's initial-attempt performance (i.e., grade on a 0 to 4 scale).[17] For easy interpretation, we reverse code this variable so that a higher value corresponds to a worse performance. Correspondingly, the term $\mathbb{1}(GP_{jsa,t-1})'T$ serves as our initial-attempt-performance fixed effects, which capture the differences in outcome common to students who receive the same letter grade from their initial attempt. The key variables of interest are the coefficients on the interaction of $GP_{jsa,t-1}$ and an indicator function $\mathbb{1}(Year_t - 2001 = k)$, which is equal to one when the years of observations is $k = (-10+), -9, ..., 13, (14+)$ years from 2001, when

[16] Ideally, we would like to adopt a measure more reflective of the student's achievement at the time of observation, such as the number of cumulative credit hours earned or the officially designated class standing status (i.e., freshmen, sophomore, junior, and senior), but these measures can be more heavily influenced by the repetition decision than the one currently used.

[17] The grade-point conversion is as follows: A+ = 4; A = 4; A- = 3.7; B+ = 3.3; B = 3; B- = 2.7; C+ = 2.3; C = 2; C- = 1.7; D+ = 1.3; D = 1; D- = 0.7; F = 0.

grade replacement was reenacted.[18] Given that the indicator for $k = -1$ is omitted, coefficients β_k, γ_k, and δ_k combined provide a complete picture of the differential trend in outcome across students of varying treatment intensity relative to 2000. In particular, we estimate a fully saturated model but only focus on the dynamic effect of grade replacement policy for a limited event window to ensure model parameters are well estimated and also to avoid potential bias arising from sample composition changes due to the exclusion of post-2015 entering cohorts (more details see Section 2.2).

Additional covariates included in model Γ are student gender, transfer status (a dummy to indicate if the student ever transferred across institutions), declared major at the time of observation (i.e., repetition for repeaters and first attempt for non-repeaters), last-observed semester, and cumulative GPAs prior to the time of observation. Finally, $\epsilon_{jsa,t/t-1}$ is the error term allowed to be clustered at the course section level.

One advantage of event study specification is that it does not impose any ex ante restrictions on when the structural breaks will occur and therefore relaxes the standard assumption of DD that treatment is associated with one-time level shift in outcomes. Given that many courses were offered once a year in our context, it may take some time for students to complete the repeated course and register for extra credits. We thus anticipate a smaller effect of grade replacement in early years of its enactment compared to later, and an even smaller initial effect for second-stage outcomes than for the first-stage outcome. By comparing the time patterns of the estimated β_k's, γ_k's, and δ_k's across models, we gain a first impression about the effect of grade forgiveness on student repetition decision and subsequent outcomes. The estimated lead effects also provide us with an important falsification test about whether any differential, pre-existing trends among students of different treatment intensity, when they are anticipated, may confound our estimates.

Besides using equation (1), we additionally include a set of next-course-section fixed

[18] We model $GP_{jsa,t-1}$ as a continuous variable to move beyond an arbitrary threshold and increase identifying variations. Coding $GP_{jsa,t-1}$ as a dummy variable (e.g., C vs DF) produces highly similar results.

effects $\rho_{s,t+1}$ into the model when investigating how well students do in the next-course attempted:

$$\begin{aligned}Y_{jsa,t\ or\ (t+1)} = &\sum_{k=-7}^{-10+} \beta_k GP_{jsa,t-1}\mathbb{1}(Year_t - 2001 = k) \\ &+ \sum_{k=-6}^{-2} \gamma_k GP_{jsa,t-1}\mathbb{1}(Year_t - 2001 = k) \\ &+ \sum_{k=0}^{14+} \delta_k GP_{jsa,t-1}\mathbb{1}(Year_t - 2001 = k) + \mathbb{1}(GP_{jsa,t-1})'T \\ &+ \rho_{s,t+1} + \kappa_{s,t-1} + \lambda_{a,t-1(t)} + \Gamma'_{jsa,t(t-1)}S + \epsilon_{jsa,t(t-1)}\end{aligned} \tag{2}$$

This slightly different specification allows us to compare the subsequent performance across students attending the same course-section after the initial course repetition, thereby controlling for the possibility that course repetition may influence a student's choice of the next course and preference for diverse types of instructors in the next course.

3.2 Instrumental Variables Approach

The estimates obtained through the event study analysis above can be viewed as the intent-to-treat or net effect of grade replacement on repeating students' outcomes. Since not every student utilized the available repeat options during our observation period, we instrument the student's actual decision to repeat with the implementation of grade replacement and estimate a parsimonious version of equations (1)-(2) to efficiently identify relevant repetition effects:

$$\begin{aligned}Enrollment_{jsa,t+1} = &\beta GP_{jsa,t-1}\widehat{Repeat}_{jsa,t} + \mathbb{1}(GP_{jsa,t-1})'T \\ &+ \kappa_{s,t-1} + \lambda_{a,t-1(t)} + \Gamma'_{jsa,t(t-1)}S + \epsilon_{jsa,t(t-1)}\end{aligned} \tag{3}$$

$$\begin{aligned}Performance_{jsa,t+1} = &\beta GP_{jsa,t-1}\widehat{Repeat}_{jsa,t} + \mathbb{1}(GP_{jsa,t-1})'T + \rho_{s,t+1} \\ &+ \kappa_{s,t-1} + \lambda_{a,t-1(t)} + \Gamma'_{jsa,t(t-1)}S + \epsilon_{jsa,t(t-1)}\end{aligned} \tag{4}$$

where $\widehat{Repeat}_{jsa,t-1}$ is the predicted probability that student j of academic progress a would repeat the course-section s at semester-year $t-1$. In addition to addressing the issue of imperfect compliance and thus providing an estimate of the repetition effect on actual repeaters, this more parsimonious specification of the IV models also enables us to conduct a systematic investigation into the biases arising from the omission of some potentially relevant and important variables from our analysis (see Section 4).

3.3 A Placebo Analysis and Other Robustness Checks

Estimates of equations (3) and (4) will produce unbiased estimates of the course repetition effects, provided that the differences in future course-taking behavior between high and low-performing CDF students are determined by exogenous shocks and are independent of unobservable determinants of students' subsequent outcomes. To check the validity of our identifying assumption, we conduct a formal test for the common trend assumption and reproduce our results for students who were enrolled in Pass/Fail courses, whose subsequent-attempt grades were not affected by the differential weight assigned by the two alternative grading schemes. This is feasible because, during the observation period, BSU did not permit students to register for a graded course as Pass/Fail on an individual basis, nor did it impose any restrictions on how many Pass/Fail courses a student was allowed to take. Therefore, to the extent that the credits earned from these non-graded courses count toward student degree requirements and prerequisites for future courses, the distribution of the unobserved ability/effort across these students should reflect that of the students enrolled in graded courses at the same time. This is certainly not a perfect placebo test, as ungraded courses with Pass/Fail options tend to be delivered through a non-traditional format (e.g., seminars and workshops) and sometimes to different groups of students, but a failure to find any similar trend in outcome for students in these courses would serve as an indication that concurrent policies or the unobserved characteristics of the students did not cause the observed effect.

Finally, additional robustness checks are also performed with regard to students are presumably more alike to each other in terms of unobservable characteristics – similar-grade recipients, students who never transferred across institutions during the observation period, and students whose outcomes are tracked for a longer time period (see Section 4.3).

4. Repetition and Course-Related Outcomes

4.1 Event Study Results

Focusing on coefficients within the event window, the upper left panel of Figure 2 shows the estimated grade replacement effects, along with their 95% confidence intervals on the likelihood of course repetition, our first-stage outcome, using equations (1). Given that grade replacement was implemented prior to 1995 and after 2001, we expect a visible trend in the estimated coefficients during these time periods relative to the baseline year of 2001. Data results are consistent with this expectation. In year -9 or 1992, the repetition rate of low-performing CDF students was significantly higher than their high-performing counterparts and the difference stayed significant until year -6 or 1995 when grade replacement was abolished. Upon the cancellation of the policy, this gap quickly dissipated and started to widen, again, after year 0 or 2002 when grade replacement was reinstated. Noticeably, all the estimated γ's are small and are individually insignificant from zero, suggesting that the pre-existing difference in the repetition rate across students of differential initial-attempt performance is unlikely a serious concern in our context.

These point estimates and their joint significance are also summarized in a DD model where the individual indicators are substituted by dummies for 2-year categories (when possible): -9 to -7, -6 to -5, -4 to -3, and -2 years before and 0 to 1, 2 to 3, 4 to 5, 6 to 7, 8 to 9, 10 to 11, and 12 to 13 years after the baseline (column (1) of Table 2). These two-year average estimates suggest a 4.4 ppt jump in the repetition rate immediately upon the inception of grade replacement in 2001 from the previous years. By years 6 to 7, the repetition effect of grade replacement had increased to 12.4 ppt and continued to rise to 18.8 ppt 13 years after the enactment of the formula, nearly two times the repetition rate in the baseline.

The scenario is quite different, however, when we examine the repetition rate among the students in Pass/Fail courses at the same time using the same model specification (equation 1; lower left panel of Figure 2). We find no meaningful trend overall in spite of some individually significant coefficients. To the extent that the time-varying confounders affect the repetition rates between the graded and Pass/Fail courses similarly, this observation suggests a rather limited role, if any, played by factors such as concurrent policies or unobserved characteristics of the students/instructors in the observed results.

Switching attention to one selected second-stage outcome, the number of subsequent credits attempted (upper right panel of Figure 2 and column (2) of Table 2) reveals a highly similar time-series pattern. The event study graph (upper right panel of Figure 2) suggests two significant trend breaks, a sudden drop of 0.2 credits, or 6%, from the baseline starting one year after the abolition of grade replacement (year -7 or 1996), and a sudden increase in credits by 0.4 points, or 11%, from the baseline starting one year after the reinstatement of grade replacement (year 2002). Notably, relative to repetition behavior (upper left panel of Figure 2) a time lag of 1-2 years is apparent in this case. This is plausible if subsequent course-taking takes place after a successful repetition of a given course. Correspondingly, it should not be surprising to see that most estimated effects of grade replacement during the pre-treatment period are statistically indistinguishable from zero, except for the year immediately after the cancellation of the policy (i.e., year -6 or 1995). Finally, just as in the case of repetition rate, a drastically different picture emerges for the students taking Pass/Fail courses at the same time. None of the estimated coefficients are individually significant, providing suggestive evidence that the parallel assumption is unlikely to be violated in this case.

Moving to the remaining measures for student interest/persistence (Figure 3 and columns (3)-(4) of Table 2), we obtain highly comparable results. In all cases, there is a positive and upward-trending gap over time through 1996, one year after the abolition of grade replacement, indicating that prior to 1996 low-performing students were more likely to retake a course and that for the most part, the likelihood of repetition increased at a faster rate than that for high-performing students. However, there is a dramatic reversal in this pattern after 1996, at which point the likelihood to retake a course among low-performing students suddenly slowed down and eventually reached the same level as that of high-performing students, until 2002 when the upward trend becomes apparent again.

Likely due to a smaller sample size and/or less identifying variations, the estimates for student course choice and subsequent performance are noisier than those for student persistence, when we restrict attention to students who took at least one follow-on course and additionally control for the next-course-section fixed effects in the regression model (Figure 4 and columns (6)-(9) of Table 2). While we no longer observe any significant trends before 1996, the overall patterns remain the same across outcomes. Applying the same specifications to Pass/Fail courses, Appendix Figure A1, once again, finds no meaningful patterns in the level of trend for any outcome under consideration.

Finally, while all regressions control for course-level fixed effects, it is likely that our results are driven by imbalances in the lead or lag variables in the event study since different courses were offered during our observation period. Appendix Table A1 thus reports the 2-year average DD results for the courses that were offered every single year from 1998 through 2011.[19] While the limited time horizon does not permit us to assess any preexisting trends prior to 1998, this practice allows us to maximize our sample size by including 87% of the courses examined before/in the full sample in this exercise.[20] As shown, a remarkably similar pattern emerges across outcomes after 1998, 4 years prior to the reintroduction of the replacement policy.

Overall, the event study analysis provides consistent evidence that during the observation period, grade replacement creates a strong incentive for low-performing CDF students – its likely beneficiaries – to retake a graded course, although it has no impact on the same students' decisions to repeat with regard to ungraded courses offered at the same time. The significant and sustained improvement in student subsequent interest and success, almost immediately following the increase in course repetition rate, thus provides suggestive evidence of a potential causal link between these two.

4.2 Instrumental Variable Regression Results

The regression results obtained from the more efficient specifications through the IV approach can be found in Tables 3 and 4. For easier interpretation, we first focus on a selected outcome, same-subject credits attempted (Table 3) to illustrate our model

[19]The event study plots are not presented for the sake of brevity but are available upon request.

[20]For example, all the technical programs in the College of Applied Technology at BSU were transitioned to a newly created community college, the College of Western Idaho, in 2008, resulting in a variety of new courses being created after that.

selection process and then present (in Table 4) the results for all other outcomes using only the preferred specification.

Columns (1) and (2) in Table 3 show the reduced form and IV estimates of the repetition effect for the full sample. In column (2), we see that the adoption of grade replacement is associated with a 34.6 ppt increase in repetition rate, and that course repetition leads to 0.52 additional subsequent credits attempted. Combining these two produces a net effect of grade replacement of 0.18, nearly identical to the reduced form estimates reported in column (1).

Given that the administrative records before 1998 do not contain course-section identifiers and some demographic characteristics of the students, we replicate the analysis for the sub-sample of students who attended courses after 1998 in columns (3)-(6) using identical model specifications. Consistent with what can be seen in the event study plots, we find the relationship of grade replacement with repetition behavior to be stronger for the students enrolled after 1998, as evidenced by a jump in the explanatory power of the instrument or the first-stage F-statistics from 566 to 1222 and an increase in the magnitude of the estimated first-stage coefficient by 51% (0.35 vs 0.52). The switch from proxy to authentic course-section fixed effects, on the other hand, has little impact on the estimated effect of grade replacement (0.52 vs. 0.50), suggesting that the multi-section courses did not drive our observed results, though the ability to identify each course-section enhances our first-stage strength (F=1222 vs 3627) and improves estimation precision. In a similar vein, the inclusion of other demographic characteristics of the students does not have any substantive impact on the main findings, except for a marginal improvement in the first-stage F-statistics (3627 vs 3633).[21]

Using the first-stage F statistics as guidance, estimates obtained for our preferred specification (column 6) suggest that grade replacement induces the repetition rate to increase by 50.4 ppt among CDF students, or six times relative to that in the baseline, 8.4%.[22] Putting this result together with that from the second stage, 0.81 follow-on

[21]The same pattern is also observed for other subsequent outcomes under study, though the results are not reported in the paper for the sake of brevity (available upon request).

[22]An alternative interpretation of the results is that each additional grade point decline on the first attempt, say from a D to F, increases a student's likelihood of retaking the course by 50 ppt

credits attempted upon repetition (the second stage estimate of column 6) leads to a net effect of grade replacement on subsequent course-taking of 0.41 credits, which, once again, closely matches the reduced form estimates reported in column (3).

Panel A of Table 4 displays the IV results through the preferred specifications (columns (2) and (6) in Table 3) for the remaining outcomes related to student interest/persistence. As shown, course repetition is estimated to have a similarly positive effect across data samples. For example, the preferred estimate of course repetition on the likelihood of attempting an additional credit in a given field (column 6 of Panel A) suggests a marginal effect of 4 ppt, or a 10% increase relative to the baseline. Combining this effect with the overall rise of 50.4 ppt in the repetition rate yields a net effect of grade replacement at 2 ppt, or a 5% increase relative to the baseline. In addition, both observed effects on the probability of attempting an extra credit and the total number of credits attempted are driven by students who successfully earned these credits (columns 3-4 and 7-8 of Panel A).

Switching attention to the students who enrolled in at least one follow-on course (Panel B of Table 4), repetition increases the average level of the next-course attempted and passed within the subject by 0.1 and 0.09 points, respectively, equivalent to a 5% and 4% increase from the baseline (columns 1-4 of Panel B).[23] Conditional on the student curriculum choice, a repeating student is 6 ppt, or 9% , more likely to pass the next course they attempted and receive a course grade 0.2 points, or 13% higher – which corresponds to nearly a plus or minus difference – than if she had not chosen to repeat the course (columns 5-8 of Panel B).

Our placebo analysis through the IV approach (columns 1-2 of Table 5) yields qualitatively similar results to those obtained through the event study regressions. While approximately half of the estimated coefficients are positive, suggesting that repeating students tend to have better outcomes than their non-repeating peers, in none of the cases is the coefficient significant.[24] Importantly, across outcomes and

under the grade replacement policy.

[23]It is not a priori clear how to define course difficulty. In addition to course level, we also replicate our analysis for two alternative measures, the average pass rate and course grade of a course section, and find highly consistent results. These results are not reported for the sake of brevity but are available upon request.

[24]We are unable to obtain the performance effect of course repetition through the multi-way fixed-

data samples, there is a dramatic reduction in the magnitude of first-stage F statistics (i.e., approximately 98% for the full and 93% for the post-1998 samples). In a few cases, the first-stage F does not exceed 10, implying that failing-grade recipients in ungraded courses were less likely to respond to the changes in the replacement policy relative to their peers who enrolled in graded courses at the same time.

In summary, results from our IV regressions suggest that repeating students benefit from course repetition by attempting and completing more follow-on credit hours in the field of the repeated course. Among the students who take extra similar-subject courses, they are also more likely to advance to a higher level and perform better in these courses, relative to their non-repeating peers.

4.3 Robustness

As an additional step to check the validity of our identifying assumption, we replicate the original analysis for the initial-attempt grade recipients of the narrowest range possible and compare outcomes among CD and DF students (columns 3-6 of Table 5). Presumably, the closer the range of the letter grade, the more similar are the students to each other in terms of their unobservable characteristics. We thus expect to see a similar trend across these samples. Columns 3-6 of Table 5 show this is exactly the case. Across outcomes, although we tend to see a decline in the effect size and an increase in the standard errors estimated, potentially attributable to the less identifying variations and smaller sample sizes, our main conclusions are nevertheless fundamentally unchanged. Thus, even with a likely downward bias, we observe a positive association of course repetition with all outcomes under study.

An alternative source of bias can arise if repeating students who enter the university in later years do not take follow-on courses immediately/within two years upon course repetition. While this concern is alleviated for their non-repeating peers, who have comparatively more time to take additional courses after the initial attempt, it is still plausible for them to have a significant time gap between initial and subsequent enrollment. To explore the possible discrepancies in our results associated with the length of the observation window, Panel B of Appendix Table 2 replicates the analysis

effect estimator for the post-1998 sample likely due to the lack of variation.

for the students who can be tracked for a minimum of 10 years and finds little difference in results between the pre-2007-cohort and all-student samples (i.e., Panel A of Appendix Table A2). This is in line with the observation that out of the 47% of the students who attempted an additional course in the same field, 75% of them took the next course after their initial attempt within two years after initial enrollment.

Finally, our estimate can also be potentially misleading because of the presence of transfer students in our sample, which poses unique econometrics challenges since (1) the courses repeated outside of BSU are not reflected in the student's transcript, and (2) the potential experience with a different institution(s), such as the extra time spent with peers and other courses taken prior to the initial enrollment, can make the transfer student more mature as well as academically advanced than her counterpart who never transfer across institutions. Panel C of Appendix Table A2 therefore re-runs our preferred model for the samples that exclude students with any outside experience, before and after their first attempt. As shown, both the statistical significance and magnitude of the repetition effects are highly similar, as before.

Overall, Table 5 and Appendix Table A2 present several specification and robustness tests that investigate potential threats to the internal validity of our results. The results provide little support for the hypothesis that our estimated repetition effects are driven by concurrent policies, unobserved characteristics between repeating and non-repeating students, our choice of the observation window, or missing data on the students who transferred across institutions during our sample period.

4.4 Differential Effects

Thus far, we have estimated the course repetition effect for students in different stages of study and courses that fall into a broad spectrum. Is the repetition effect the same for all students? In theory, the learning gains from repetition may be greater for first-year students who are in the process of adapting to the new college environment than for their more senior counterparts. Retaking a course thus may provide them with the additional time necessary to acquire knowledge of the resources the college provides, the skills their courses require, and the attitudes needed to be academically successful. Results reported for our preferred data sample in columns 1-2 of Table 6 support this hypothesis, by finding a more pronounced enrollment effect of repetition for first-year students than for their more senior counterparts (columns 1-2 of Panel A). All else equal, a course repeated by a student during her first year of study leads to 1.1 additional follow-on credit hours attempted and passed, a 34-41% increase relative to the pretreatment level, and over 2 times as large as the relative effect sizes estimated for other students (17%). Similarly, the likelihood of attempting and passing an extra credit increases by 7 ppt, or 19-24% upon repetition, nearly 3 times the relative effect sizes of non-first-year students (9%). Conditional on persistence, we see that the observed performance effect of repetition is concentrated on first-year students (Panel C of Table 6). Repetition boosts an average student's chance of passing the next course in the subject by 12 ppt or 20% and the course grade she receives by 0.3 points or 21%, when no such effect is observed for other students. While the absolute magnitude of the course choice coefficients is smaller for first year than that for other students, the relative size is nevertheless comparable: a 4-6% increase in both cases.

In terms of the type of courses that are repeated, we anticipate introductory/lower-division courses to matter more than upper-division courses, as introductory courses are often large, impersonal, and populated by diverse groups of students with varied levels of knowledge and motivation. Yet, they are often critical gateways to majors and careers, requiring high grades to continue in the field (Harackiewicz et al. 2016). Columns (3)-(4) of Table 6 provide evidence consistent with this conjecture. A student who repeats an introductory level course is approximately 4 ppt or 10-12% more likely to attempt and earn an extra credit in the field of the repeated course than her non-repeating peers, compared to a 3-4 ppt or 7-9% gain for students who repeat an upper-division course. Therefore, in relative terms, the estimated enrollment effect of course repetition is in fact greater for the former than for the latter. Similarly, repeating an introductory course leads to 21-25% more similar-subject credits attempted and passed, and an increase in the average level of courses attempted and passed by 5%, whereas the comparable figures for students who repeat an upper-division course are 19-23% and 2%. Just like the case of first-year students, we find the estimated performance effect of course repetition to be entirely driven by students repeating introductory-level courses. Repetition is associated with an increase in the pass rate of the next course attempted in the field by 6 ppt or 9% and the average course grade received by 0.2 points or 13% for students repeating introductory courses, but it makes no statistically distinguishable difference for those repeating upper division ones. It is worth noting that, in our sample, only 40% of the students enrolled in introductory courses are first-year students. These results thus point to an independent effect of repeating introductory-level courses - separate from that of repeating as a first-year student - on the student's future course-taking behavior and learning outcomes.[25]

Overall, the evidence presented in this section suggests that course repetition is most beneficial for students who are exposed to the college environment or a subject matter for the first time. If having the additional time and resources allows these first-timers to engage with course material at a deeper level (which serve as a foundation for future study within the degree program), then learning gains made while repeating a course could have reinforced their academic interest and subsequent success.

5. Repetition and Graduation-Related Outcomes

Section 4 reports that low-performing students who repeat a course tend to fare better than their comparable non-repeating classmates, but are these findings evidence of just a temporary gain? This section explores if course repetition has any longer-term impact on the academic success of the students in our preferred sample who can be tracked for a minimum of 10 years (i.e., 1998-2006 entry cohorts). In particular, we assess whether repetition behavior inadvertently extends a student's length of time to earn a degree (in semesters) and/or diminishes their chance of college completion. Since the relevant GPA policy changes over time, this exercise exploits an additional source of variation in the data by instrumenting the total number of courses repeated by the student during the entire undergraduate career with their average exposure to the grade replacement policy, while controlling for the time-invariant covariates in the original model specification, namely, the student's gender, ethnicity, transfer and

[25] Separately examining the repetition effect by student demographic characteristics such as gender, race, and age at college entry (i.e., below 21 versus 22 and above) yields mixed patterns across outcomes. These results are not reported in the paper but are available upon request.

in-state status, and age at college entry (in 4 categories).[26] The student's average exposure to the policy is proxied by the fraction of time (in semesters) during which the grade replacement policy was enacted.

Panel A of Table 7 indicates that, conditional on a sufficiently strong first-stage result (i.e., First-Stage F = 294) where a 10% increase in the exposure to the grade replacement policy induces a 0.04 increase in the total number of repeated courses, repeating one additional course as a result of the grade replacement policy adds 0.4 semesters or 3% more time from the baseline to an average graduate's time-to-degree (column 1). This adverse effect, however, is concentrated on the graduates who repeat more than four courses during their degree programs (columns 2-3). To determine if transfer students may skew our results, Panel B replicates the analysis for the graduates who did not have any outside experience during the observation period using the same selection criterion as in Panel C of Appendix Table A2, and it finds the results remain largely unchanged, despite a substantial reduction of nearly 70% in the sample size.

Whether or not a student graduates, course repetition increases a student's chance of graduation, regardless of the frequency of repetition. Counting both transfer and non-transfer students (Appendix Table A2, columns 1-3 of Panel B), repeating one course boosts the likelihood of attaining a degree by an average of 16 ppt or 167% from the baseline, and the additional gain decreases with the number of courses repeated. Students who engaged in less than four repetitions during their college tenure are 4 ppt more likely to graduate than those who had more than four repetitions – for whom a longer time is required for degree requirements (see column 3 of Panel A) – and this pattern is further enhanced when we restrict attention to the 52% of the first-time students (6 ppt; columns 5-6 of Panel B).

To provide more rationale for the results observed above, Panels C and D investigate student study pace, or the number of semesters elapsed before attempting and completing an extra credit within the field for the same students in Panels A and B, respectively, using the model specification identical to column 6 of Table 3. Across

[26]Repeating a course twice is counted as two repeats according to our calculation, though excluding the one percent of repeating students who repeated a course twice from the analysis does not change our findings qualitatively. These results are not reported for brevity but are available upon request.

transfer status, we find that repeating students attempt and complete a similar-subject credit 1-2 semesters sooner than their non-repeating peers, equivalent to a 18-56% increase from the pretreatment level of repeaters and that the estimated effects are driven by the actual completion rather than incomplete attempt of a given course. This finding is in line with the result reported in Jiang et al. (2013), where students are observed to take on a heavier semester course load under grade replacement than when the grade averaging formula was employed, and it is plausible if 1) the extra time and resources that repeating students gain from course repetition enable them to catch up with their non-repeating classmates academically and behaviorally, and/or 2) the safety net effect of grade replacement policy allows these lower-performing students to be more receptive to loss/risk-taking and thereby become less likely to trade a full course load for higher grades in comparison to their high-performing, non-repeating peers.

6. Mechanisms

Tables 8-10 provide a more formal investigation into the underlying mechanisms through which course repetition may influence student motivation and subsequent success. In principle, a course do-over under grade replacement may affect the observed outcomes in two different ways: 1) through its effect on deep learning, if the improved experience, judgment, and familiarity with the course material helps students prepare for future challenges, and induces them to update beliefs regarding their own ability and the relevance of the course topic; and 2) through a GPA effect, if the disproportional increases in the cumulative GPAs of repeating students incentivize them to take more follow-on courses. For example, the minimum requirements to declare a major are often set in accordance with a student's earned credit hours and cumulative GPA. The inflated GPA due to grade replacement thus may mechanically allow repeating students to pursue an academic path of their choice that otherwise would not have been feasible. Additionally, there might be an independent effect of an elevated GPA on student perception of their own ability for the same scholarly achievement. For example, Owen (2010) finds that receiving an A for a final grade in the first economics class is associated with a meaningful increase in the probability of majoring in economics, even after controlling for the numerical grade earned in the class, though the effect is concentrated on female students. Echoing this result, Chen et al. (2021) also observe that across students in different disciplines, the feedback embedded in the course letter grade incentivizes subsequent enrollment.

Are our observed results driven by one of the two proposed channels or a combination of both? While we do not have credible variations to causally estimate the effects of deep learning and/or GPA, we can provide preliminary evidence by gauging how different repeating students perform between their initial and subsequent attempts and whether the differences in performance over time help explain the observed pattern in outcome for *the same students*. Assuming that learning gains through repetition are crystallized in subsequent-attempt outcomes, we attribute any unexplained differences in outcome to the effect of GPA/grade inflation.

Panel A of Table 8 examines whether there is evidence that course repetition enhances deep learning for students in all courses by comparing the subsequent-attempt performance of repeaters with the initial-attempt performance of their non-repeating peers in the same course with respect to: 1) the likelihood of passing a course, 2) the likelihood of earning a B or better, and 3) grade point on a 0 to 4 grading scale. Our preferred estimates based on post-1998 data (columns 4-6 of Panel A), which largely mirror those obtained for the full sample (columns 1-3 of Panel A), indicate that repeating students are 15 ppt (or 17%) more likely to pass a course (column 4 of Panel A), 15 ppt (or 27%) less likely to receive a CDF (column 5 of Panel A), and score 0.60 points (or 33%) higher the second time around than their non-repeating peers in the same course. Considering that the average grade of CDF repeaters in the baseline or 2001 was 1.739, this improvement thus is equivalent to raising one's grade from approximately a C- to a C+.

One concern with the previous analysis is that a repeating student may self-select into a particular course-section on the subsequent attempt of the same course. In other words, differences in instructor grading leniency, teaching effectiveness, or other course-section specific characteristics (e.g., the time and location the section is offered) could drive the observed results and lead to misleading conclusions. To tackle this issue, Panel B of Table 8 focuses on repeaters who make both attempts with the same instructor.[27] If the grading standards adopted by an instructor for the same course do not change abruptly over time, then a higher grade awarded to the same student on the subsequent attempt would largely reflect the student's improved understanding of course materials. As shown, while about 76% of repeaters are excluded from these exercises as a result, applying this sample selection criterion leaves our findings fundamentally unchanged (Panel B of Table 7). There are two potential explanations for this result. One is that when making their subsequent attempt, repeating students do not necessarily shop for instructors based on their grading leniency. The other is that additional learning occurs no matter whether the student repeats the course with the original instructor or not. Regardless, results presented in this panel provide evidence that differences in course-section-specific factors, such as instructor grading standards, are unlikely to account for the observed effects.

As an alternative approach, Panel C of Table 8 examines students who enrolled in courses offered through one single section during the entire observation period.[28] Since these solo-section courses leave no room for students to choose which section they enroll in, consistent results obtained from this sample would provide additional comfort in our findings. As demonstrated, despite a substantial reduction in sample size (by 83%), we once again observe little change in the estimated deep learning effect of course repetition.

Taken together, the evidence presented in this section suggests that repeating students gain more knowledge about the subject matter through their repetition as measured by their subsequent-attempt performance and that their nonrandom sorting into course-sections on their subsequent attempt is not the driving force for the observed relationship. Thus, the enhanced learning through repetition might be the cause why repeating students develop greater interest, persist longer, and achieve better future learning outcomes than their non-repeating peers.

[27] Including instructor fixed effects in our original models produces highly similar results.

[28] One limitation of this approach is that it forces us to focus on a small subset of courses that have lower average enrollment and potentially other characteristics that are different from a typical course offered on campus at the time (e.g., upper vs lower division courses). If the subsequent course-taking behavior of the students enrolled in these courses diverges, our findings might be misleading.

To further test the plausibility of this hypothesis, we re-estimate our original models for a selected outcome, the number of subsequent credits attempted, for the full (columns 1-2) and post-1998 samples (columns 3-4) in Table 9, while adding repeating students' grade differences between initial and subsequent attempts and their interactions with the students' initial-attempt grades (Table 8). The grade difference variable would take a value of zero for their non-repeating counterparts. If the repetition effect operates through this channel, we would expect the estimated interaction term GP *times* Course Repetition to decline under this specification. While informative, this analysis should be interpreted cautiously, because the course grades of a repeating student are likely correlated with important determinants of the outcome not included in the regression and hence potentially endogenous in the equation.

Estimates indicate that the set of grade-difference-related variables are jointly significant on their own in most cases (see the p-values of F-statistics), consistent with the results found in Table 8. Importantly, after the inclusion of these variables, the previously observed repetition effects are greatly reduced, particularly after the repeaters' self-selection is accounted for (columns 6 and 8), which suggests that the observed difference in subsequent enrollment is largely driven by the learning gains made through repetition.

Table 9 presents the results for the remaining outcomes, and a similar pattern emerges. When we focus on the samples that account for repeaters' self-selection (columns 5-8), all the previously estimated results disappear, and in some cases the trend is even reversed. For example, supposing that repetition did not take place, or, even if it did, no learning gains occurred through the second attempt, repeating students would have in fact been less likely to pass the next course in the subject by 33.5 ppt (or 41%), and would receive a lower average grade by 1.1 points or 54% than their non-repeating peers (see column (8) in Table 10).

In conclusion, evidence found in this section suggests that the observed course repetition effect is entirely explained by the enhanced learning through subsequent attempts. Controlling for subsequent-attempt course outcomes, repeating students are just as likely to pursue/earn extra credits or challenge a more difficult course in a field as those who do not repeat any courses. Conditional on enrollment in the next course, the performance of non-repeaters could be even worse.

7. Discussion and Conclusion

College education leads to increased social mobility, better employment prospects, and a lifetime of higher wages. Yet college enrollments are declining despite many federal and state efforts to encourage more college-going in recent decades (National Center for Education Statistics 2023). This study explores a policy option that might help low-performing students—especially first-year students and other students taking the introductory-level courses of a given field (e.g., economics majors in a computer science course)—capitalize on their existing interest, persist in degree programs, and eventually graduate on time.

We quantify the short-term effects of course repetition on repeating students by comparing within-course-section student responses to a grade replacement formula that was implemented, abolished, and then re-introduced at a four-year public institution over a 27-year period. By examining the changes in the subsequent enrollment and performance patterns among below-average students who receive different course grades on their initial attempt and therefore are differentially affected by the changes in the GPA policy, we find course repetition to have a favorable impact on repeating students' future academic interest and learning outcomes in the subject of the repeated course. Even in the presence of a likely downward bias, we estimate that course repetition significantly boosts the likelihood of enrolling in and successfully completing an additional course by 5% and increases the number of similar-subject credits attempted and earned by 21% relative to their pretreatment levels. Given that these results can be driven by the different curriculum choices of the students, our follow-up investigation further reveals that repeating students in fact tend to challenge more difficult courses upon repetition than their non-repeating peers, as measured by a 4-5% increase in the difficulty level of the next course attempted and passed. Conditional on next-course choice, repeaters are 9% more likely to pass a course-section and receive a grade 13% higher, or nearly an additional "plus" on their letter grade (e.g., from C to C+), than their non-repeating classmates.

There is considerable heterogeneity in the estimated repetition effects depending on the timing of repetition and the type of courses being repeated. Retaking a course during one's first year of study boosts the student's likelihood of attempting and completing a follow-on course in the field by as much as 24% and increases the number of follow-on credits the student attempts and completes by 41%, over two times as large as the estimated effect size for other students. Repeating an introductory-level course generates slightly weaker but qualitatively similar effects on outcomes under study relative to repeating an upper-division course. In both cases, a student's performance – conditional on the student's course choice – is substantially improved after the course repetition, whereas no such a performance effect is observed for any upper-division courses being repeated or among the students who repeat any courses at any alternative stage of study. The estimated next-course grade effect of repetition for first-year students falls roughly within the range of previous estimates by Tafreschi and Thiemann (2016) for retained first-year undergraduates in Germany (0.2 vs 0.5 standard deviations). A smaller magnitude in our case is plausible if the cumulative gains in deep learning through repeating all first-year courses are greater than that from repeating an individual one.

Although the effect size estimated in this study is relatively modest, it can make a meaningful difference in reducing the outcome gap between high and low-performing students. Based on the descriptive statistics reported in Table 1 (columns 2 and 4), the estimated effect of course repetition is 28-40% of the subsequent enrollment gap and 26-60% of the next-course performance gap observed between AB and CDF students, which implies that an easier access to course repetition can substantially reduce these education inequalities at the school starting gate, given that first-year freshmen and students in introductory courses stand to reap the most benefit. To the extent that the introductory-level courses of a given field serve as critical gateways to majors and careers, a more lenient repetition policy also has the potential for encouraging low-performing students to pursue a degree in the fields that are difficult but nevertheless lead to high-paying jobs (i.e., STEM), and for reducing income inequalities in the longer run.

While popular beliefs speculate that less stringent oversight of course repetitions may pose a detriment to on-time degree completion and even college completion, we find no evidence to support these conjectures. Upon repetition, repeating students catch up with their non-repeating peers by attempting and completing a similar-subject extra credit at a faster pace by 1-2 semesters or 18-56% sooner than the baseline speed. Correspondingly, students who repeat a moderate number of courses (i.e., four or less) during their college tenure are in fact 22-24 ppt (or 300% relative to the baseline rate) more likely to attain a four-year degree in a timely fashion than their non-repeating classmates. Assuming that individuals with a bachelor's degree earn an average of $2.8 million during their careers, $1.2 million more than the median for workers with a high school diploma (Carnevale et al. 2021), this result implies an increase of $264,000-$288,000 in lifetime earnings for each affected student, far exceeding the monetary cost of course repetition using the published sticker price for tuition (before scholarships/financial aids and after state appropriations) as a rough indicator (e.g., $756 for a three-credit course in 2022). In this view, a natural question to ask is whether institutional decisions to restrict course repeatability for college students should be made solely on the grounds of budgetary control, since allowing additional opportunities for the students to learn from their mistakes – as deemed necessary by themselves – clearly brings additional benefits that are not only important for improving the employment outcomes of individual students, but also for reducing the economic inequality of the community at large.

8. References

Advancing Global Higher Education (2015). Course Repeat Practices.https://www.aacrao.org/docs/default-source/research-docs/aacrao-september-2015-60-second-survey-course-repeat-practices.pdf?sfvrsn=51fd4b36_4. Accessed on June 9, 2023.

Bettinger, Eric P., and Bridget Terry Long. "Addressing the needs of underprepared students in higher education does college remediation work?." Journal of Human resources 44, no. 3 (2009): 736-771.

Biktimirov, Ernest N., and Michael J. Armstrong. "Is the second time the charm for students repeating introductory finance?." Journal of Financial Education (2015): 32-49.

Carnevale, Anthony P., Ban Cheah, and Emma Wenzinger. "The College Payoff: More Education Doesn't Always Mean More Earnings." Georgetown University Center on Education and the Workforce (2021).

Carrell, Scott E., and James E. West. "Does professor quality matter? Evidence from random assignment of students to professors." Journal of Political Economy 118, no. 3 (2010): 409-432.

Casas, François R., and Diane E. Meaghan. "A study of repeated courses among secondary students in Ontario." The Journal of Educational Research 90, no. 2 (1996): 116-127.

Chen, Kelly, Zeynep Hansen, and Scott Lowe. "Why do we inflate grades? The effect of adjunct faculty employment on instructor grading standards." Journal of Human Resources 56, no. 3 (2021): 878-921.

Jiang, Xuan, Kelly Chen, Zeynep K. Hansen, and Scott Lowe. "A Second Chance at Success? Effects of College Grade Forgiveness Policies on Student Outcomes." No. w29493. National Bureau of Economic Research, September 2023.

De Paola, Maria, and Vincenzo Scoppa. "Procrastination, academic success and the effectiveness of a remedial program." Journal of Economic Behavior & Organization 115 (2015): 217-236.

Figlio, David N., Morton O. Schapiro, and Kevin B. Soter. "Are tenure track professors better teachers?." Review of Economics and Statistics 97, no. 4 (2015): 715-724.

Fournier, Gary M., and Tim R. Sass. "Take my course, please: The effects of the principles experience on student curriculum choice." The Journal of Economic Education 31, no. 4 (2000): 323-339.

Frisancho, Veronica, Kala Krishna, Sergey Lychagin, and Cemile Yavas. "Better luck next time: Learning through retaking." Journal of Economic Behavior & Organization 125 (2016): 120-135.

Goodman, Joshua, Oded Gurantz, and Jonathan Smith. "Take two! SAT retaking and college enrollment gaps." American Economic Journal: Economic Policy 12, no. 2 (2020): 115-158.

Harackiewicz, Judith M., Jessi L. Smith, and Stacy J. Priniski. "Interest matters: The importance of promoting interest in education." Policy insights from the behavioral and brain sciences 3, no. 2 (2016): 220-227.

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

Hoffmann, Florian, and Philip Oreopoulos. "Professor qualities and student achievement." The Review of Economics and Statistics 91, no. 1 (2009): 83-92.

Jewell, R. Todd, Michael A. McPherson, and Margie A. Tieslau. "Whose fault is it? Assigning blame for grade inflation in higher education." Applied economics 45, no. 9 (2013): 1185-1200.

Martorell, Paco, and Isaac McFarlin Jr. "Help or hindrance? The effects of college remediation on academic and labor market outcomes." The Review of Economics and Statistics 93, no. 2 (2011): 436-454.

Marx, Jonathan, and David Meeler. "Strike four! Do-over policies institutionalize GPA distortion." Quality Assurance in Education 21, no. 1 (2013): 39-53.

National Center for Education Statistics. (2023). "College Enrollment Rates." Condition of Education. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/cpb. Retrieved on June 19, 2023.

Owen, Ann L. "Grades, gender, and encouragement: A regression discontinuity analysis." The Journal of Economic Education 41, no. 3 (2010): 217-234.

Tafreschi, Darjusch, and Petra Thiemann. "Doing it twice, getting it right? The effects of grade retention and course repetition in higher education." Economics of Education Review 55 (2016): 198-219.

Scott-Clayton, Judith, Peter M. Crosta, and Clive R. Belfield. "Improving the targeting of treatment: Evidence from college remediation." Educational Evaluation and Policy Analysis 36, no. 3 (2014): 371-393.

Sovero, Veronica and Amanda Griffith (2023). "Second Try's a Charm: The Impact of Financial Aid Policy on Course Retaking Behavior for Low Income Students." Manuscript.

Vigdor, Jacob L., and Charles T. Clotfelter. "Retaking the SAT." Journal of Human Resources 38, no. 1 (2003): 1-33.

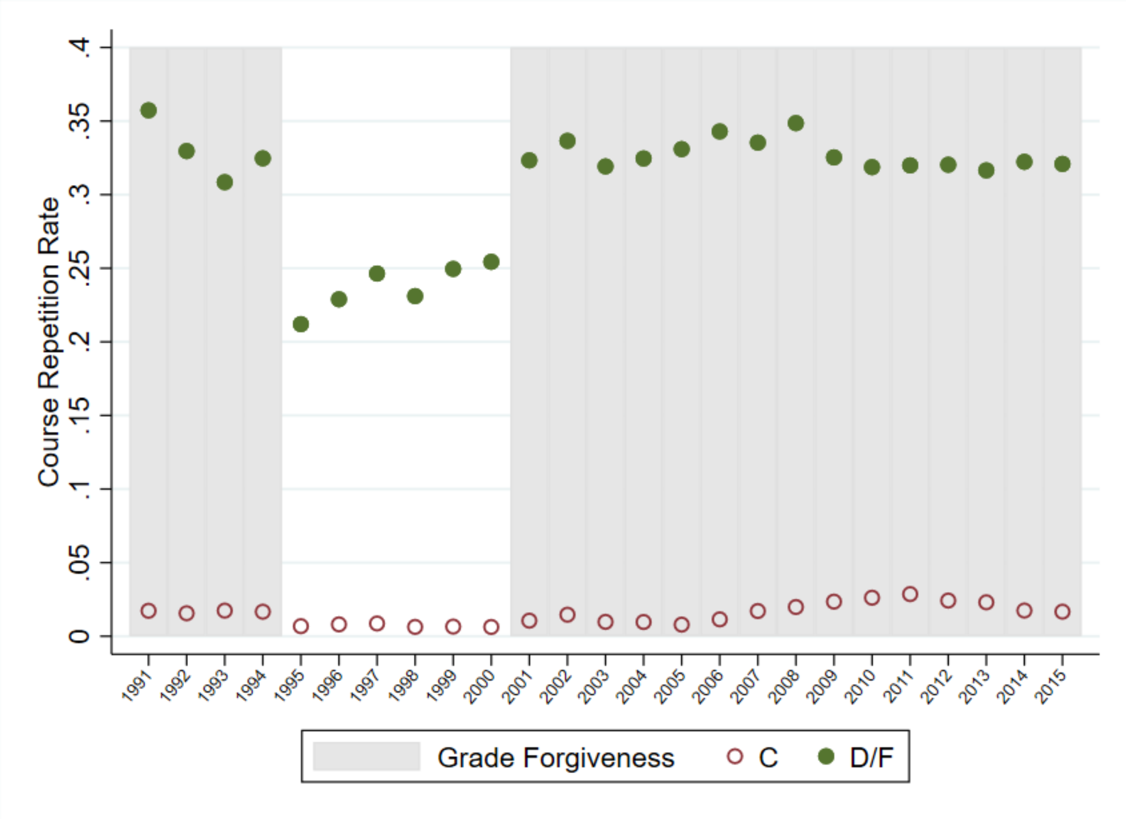
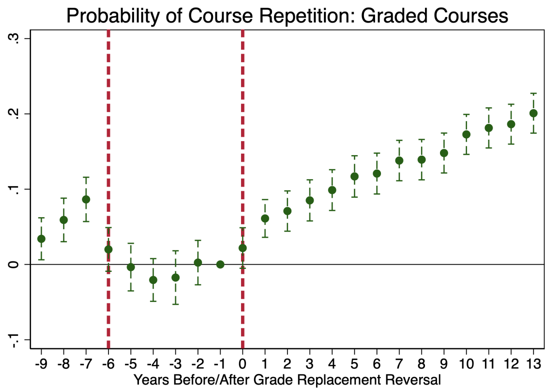
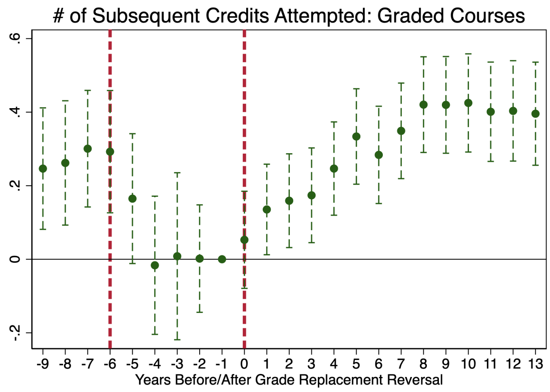


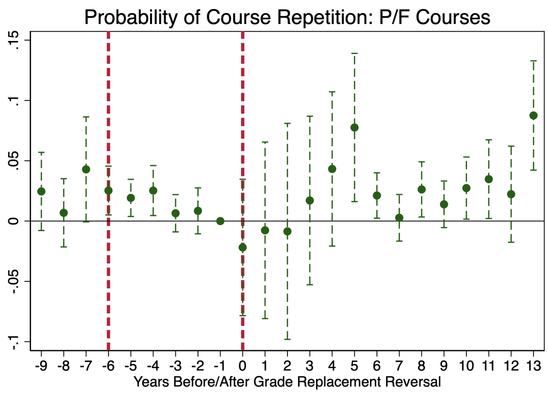
Figure 1: Raw Difference in Course Repetition Rate between C and DF Students



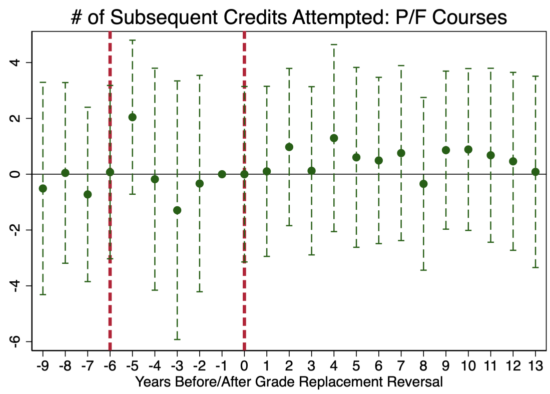
(a)



(b)



(c)



(d)

Figure 2: Repetition Rate and Same-Subject Credits Attempted (Graded vs P/F Courses)

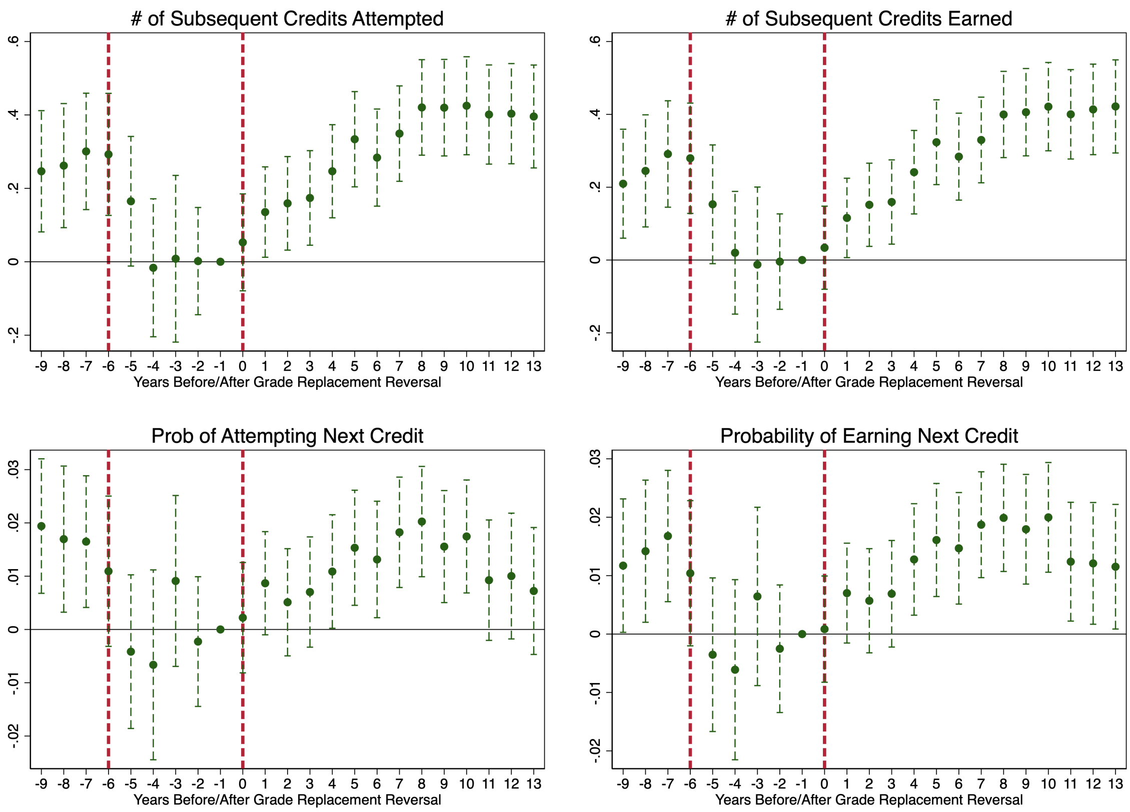


Figure 3: Student Interest/Persistence (Graded Courses)

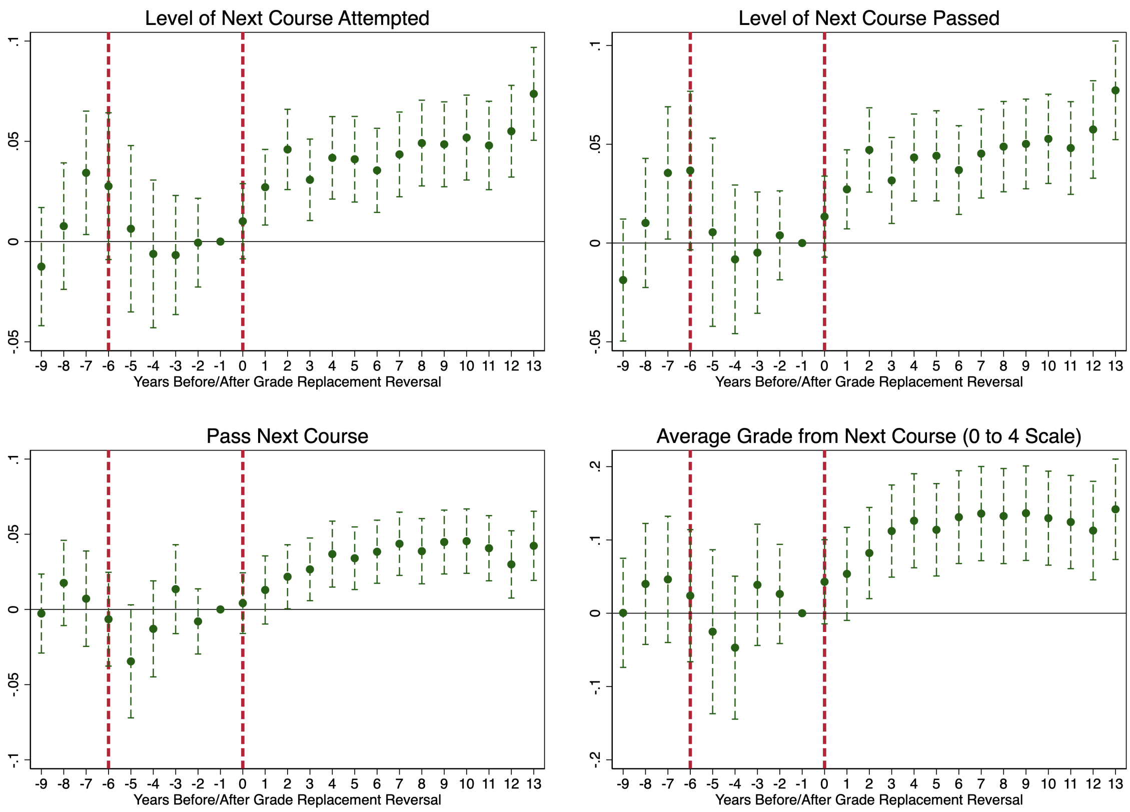


Figure 4: Student Course Choice and Subsequent Performance (Graded Courses)

Table 1: Summary Statistics of Repeating and Non-Repeating Students

	All Students		CDF Students			
	1991-2016		1991-2016		1998-2016	
	R (1)	NR (2)	R (3)	NR (4)	R (5)	NR (6)
Panel A: Student-level characteristics						
Female	0.46	0.54***	0.46	0.48***	0.46	0.48***
Transfer student	0.49	0.52***	0.48	0.46***	0.5	0.51
Major declaration	0.95	0.90***	0.95	0.89***	0.98	0.96***
White	–	–	–	–	0.77	0.78***
Hispanics	–	–	–	–	0.09	0.08***
Age at entry	–	–	–	–	20.93	21.41***
Idaho residence	–	–	–	–	0.87	0.85***
# of courses repeated: 1-4	0.90	–	–	–	–	–
N	41,379	155,433	40,940	111,638	34,407	83,666
Panel B: Student-course-section-level characteristics						
First attempt grade: A	0	0.40***	–	–	–	–
First attempt grade: B	0.03	0.33***	–	–	–	–
First attempt grade: C	0.07	0.18***	0.07	0.68***	0.07	0.67***
First attempt grade: D	0.37	0.03***	0.38	0.10***	0.35	0.10***
First attempt grade: F	0.52	0.06***	0.53	0.21***	0.54	0.22***
Prior term GPA	1.93	2.49***	1.92	1.92***	1.92	1.97***
Prior cumulative GPA	2.31	2.50***	2.3	2.04***	2.31	2.11***
First-year students	0.11	0.33***	0.11	0.39***	0.1	0.35***
Introductory courses	0.83	0.65***	0.83	0.77***	0.83	0.75***
First-time repeater	0.99	–	0.99	–	0.99	–
N	87,015	1,982,771	85,402	517,928	73,316	402,529
Panel C: Student-course-section-level outcomes						
Panel C.1: Persistence						
# of credits attempted	3.72	6.71***	3.69	4.13***	3.55	3.98***
# of credits earned	3.09	6.32***	3.05	3.59***	2.95	3.49***
Likelihood of attempting extra credit	0.39	0.47***	0.38	0.37***	0.38	0.36***
Likelihood of passing extra credit	0.33	0.45***	0.33	0.33***	0.33	0.32***
N	87,015	1,982,771	85,402	517,928	73,316	402,529
Panel C.2: Course choice						
Level of next-course attempted	2.82	2.70*	2.81	2.71	2.8	2.75
Number of observations	32,559	917,392	31,581	179,310	26,548	137,885
Level of next-course passed	2.63	2.69	2.61	2.63	2.59	2.66
N	28,069	877,478	27,097	158,149	22,518	121,438
Panel C.3: Performance						
Pass next course	0.76	0.91***	0.75	0.81***	0.75	0.81***
Next-Course Grade	1.96	2.91***	1.94	2.11***	1.95	2.15***
N	18,334	402,470	17,189	79,071	13,837	54,177

Notes: Stars are the p-value of a t-test for the standardized difference (i.e., the difference in standard deviations) between non-repeating and repeating students. R and NR represent repeating and non-repeating students, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Probability of Course Repetition and Course-Related Outcomes

	First-Stage	Interest/Persistence				Course Choice		Subsequent Performance	
	Prob of Course Repetition	# of Credits Attempted	# of Credits Earned	Prob of Attempting Next Credit	Prob of Earning Next Credit	Level of Next-Course Attempted	Level of Next-Course Passed	Pass Next Course	Average Grade from Next Course
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean DV in 2001	0.1	3.804	3.026	0.409	0.351	2.245	2.312	0.691	1.673
(SD)	(0.301)	(7.74)	(6.821)	(0.492)	(0.478)	(1.084)	(1.089)	(0.463)	(1.345)
Years -9 to -7	0.0546***	0.257***	0.233***	0.0167***	0.0133**	0.00687	0.00599	0.00362	0.0201
	(0.012)	(0.074)	(0.067)	(0.006)	(0.005)	(0.013)	(0.014)	(0.012)	(0.036)
Years -5 to -6	0.00303	0.222***	0.208***	0.00354	0.00328	0.0152	0.0205	-0.0165	0.00332
	(0.014)	(0.075)	(0.069)	(0.006)	(0.006)	(0.016)	(0.018)	(0.014)	(0.042)
Years -4 to -3	-0.0225	-0.0107	0.00102	0.0000243	-0.000864	-0.00746	-0.00647	0.000806	-0.000914
	(0.014)	(0.086)	(0.078)	(0.007)	(0.006)	(0.014)	(0.014)	(0.013)	(0.038)
Year -2 (to -1)	-0.0000551	-0.00349	-0.00826	-0.00293	-0.0031	-0.00146	0.00322	-0.00967	0.0227
	(0.015)	(0.074)	(0.066)	(0.006)	(0.006)	(0.011)	(0.011)	(0.011)	(0.034)
Years 0 to 1	0.0435***	0.0978*	0.0776	0.00571	0.00417	0.0185**	0.0202**	0.00834	0.0492*
	(0.012)	(0.058)	(0.051)	(0.005)	(0.004)	(0.009)	(0.009)	(0.01)	(0.028)
Years 2 to 3	0.0745***	0.161***	0.150***	0.00545	0.00572	0.0363***	0.0381***	0.0221**	0.0937***
	(0.012)	(0.06)	(0.054)	(0.005)	(0.004)	(0.01)	(0.01)	(0.01)	(0.03)
Years 4 to 5	0.103***	0.281***	0.274***	0.0123**	0.0138***	0.0418***	0.0447***	0.0335***	0.115***
	(0.012)	(0.061)	(0.055)	(0.005)	(0.005)	(0.01)	(0.011)	(0.01)	(0.031)
Years 6 to 7	0.124***	0.299***	0.291***	0.0140***	0.0155***	0.0388***	0.0409***	0.0394***	0.131***
	(0.012)	(0.063)	(0.057)	(0.005)	(0.004)	(0.01)	(0.011)	(0.01)	(0.03)
Years 8 to 9	0.138***	0.401***	0.386***	0.0159***	0.0173***	0.0474***	0.0484***	0.0394***	0.131***
	(0.012)	(0.063)	(0.057)	(0.005)	(0.005)	(0.01)	(0.011)	(0.01)	(0.031)
Years 10 to 11	0.172***	0.399***	0.397***	0.0127**	0.0156***	0.0502***	0.0510***	0.0414***	0.125***
	(0.012)	(0.065)	(0.059)	(0.005)	(0.005)	(0.01)	(0.011)	(0.011)	(0.031)
Year 12 to 13	0.188***	0.386***	0.404***	0.0079	0.0112**	0.0625***	0.0657***	0.0340***	0.124***
	(0.012)	(0.067)	(0.06)	(0.006)	(0.005)	(0.011)	(0.012)	(0.011)	(0.032)
# of Courses	2,843	2,843	2,843	2,843	2,843	2,283	1,521	1,521	1,521
# of Course Sections	28,343	28,343	28,343	28,343	28,343	20,013	8,860	8,860	8,860
# of Observations	603,330	603,330	603,330	603,330	603,330	210,891	185,246	96,260	96,260

Notes: All models control proxy-course-section fixed effects, academic-progress fixed effects, and initial-attempt-performance fixed effects. Covariates include student gender, transfer status, declared major, the student's last-observed semester and cumulative GPAs prior to the time of observation, and their interactions with the student's initial-attempt grade. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Reduced Form and IV Results for the Number of Subsequent Credits Attempted

	Full Sample		1998-2016			
	OLS	IV/2SLS	OLS	IV/2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Reduced Form	0.181***		0.406***			
	(0.035)		(0.048)			
First Stage		0.346***		0.521***	0.504***	0.504***
		(0.015)		(0.015)	(0.008)	(0.008)
Second Stage		0.523***		0.823***	0.820***	0.805***
		(0.103)		(0.104)	(0.094)	(0.094)
F-Test of Excluded Instruments		566		1222	3627	3633
Course-Section Fixed Effects	No	No	No	No	Yes	Yes
Basic Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Additional Covariates	No	No	No	No	No	Yes
Mean DV in 2001	3.804	3.804	3.946	3.946	3.946	3.946
(SD)	(7.74)	(7.74)	(8.669)	(8.669)	(8.669)	(8.669)
Mean Repetition Rate in 2001	0.1	0.1	0.084	0.084	0.084	0.084
(SD)	(0.301)	(0.301)	(0.278)	(0.278)	(0.278)	(0.278)
N	603,330	603,330	475,837	475,837	475,837	475,837

Notes: All models control for academic-progress fixed effects and initial-attempt-performance fixed effects. Covariates controlled in regressions using the full sample are identical to those in Table 1 (i.e., basic covariates). Regressions using data from the post-1998 sample additionally include the student's ethnicity, in-state status, and age at college entry (i.e., additional covariates). Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: IV Results for Course-Related Outcomes

Panel A: Interest/Persistence

	# of Credits Attempted		# of Credits Earned		Prob of Attempting Next Credit		Prob of Earning Next Credit	
	Full Sample	98-17	Full Sample	98-17	Full Sample	98-17	Full Sample	98-17
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1st Stage	0.346***	0.504***	0.345***	0.504***	0.345***	0.504***	0.345***	0.504***
	(0.015)	(0.008)	(0.015)	(0.008)	(0.015)	(0.008)	(0.015)	(0.008)
2nd Stage	0.523***	0.805***	0.494***	0.762***	0.030***	0.040***	0.030***	0.042***
	(0.103)	(0.094)	(0.096)	(0.089)	(0.008)	(0.005)	(0.008)	(0.005)
1st-Stage F	566	3633	566	3633	566	3633	566	3633
Mean DV	3.804	3.946	3.026	3.077	0.409	0.415	0.351	0.355
(SD)	(7.74)	(8.669)	(6.821)	(6.87)	(0.492)	(0.493)	(0.478)	(0.479)
Mean Repetition Rate	0.1	0.084	0.1	0.084	0.1	0.084	0.1	0.084
(SD)	(0.301)	(0.278)	(0.301)	(0.278)	(0.301)	(0.278)	(0.301)	(0.278)
N	603,330	475,837	603,330	475,837	603,330	475,837	603,330	475,837

Panel B: Course Choice and Subsequent Performance

	Level of Next-Course Attempted		Level of Next-Course Passed		Pass Next Course		Average Grade from Next Course	
	Full Sample	98-17	Full Sample	98-17	Full Sample	98-17	Full Sample	98-17
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1st Stage	0.375***	0.569***	0.373***	0.579***	0.344***	0.590***	0.344***	0.590***
	(0.02)	(0.014)	(0.02)	(0.016)	(0.029)	(0.033)	(0.029)	(0.033)
2nd Stage	0.071***	0.090***	0.069***	0.086***	0.076***	0.059***	0.209***	0.206***
	(0.016)	(0.013)	(0.017)	(0.014)	(0.017)	(0.018)	(0.048)	(0.049)
1st-Stage F	363	1561	338	1315	138	324	138	324
Mean DV	2.245	1.993	2.312	2.062	0.691	0.693	1.673	1.628
(SD)	(1.084)	(1.066)	(1.089)	(1.094)	(0.463)	(0.463)	(1.345)	(1.319)
Mean Repetition Rate	0.115	0.096	0.112	0.092	0.138	0.108	0.138	0.108
(SD)	(0.319)	(0.295)	(0.315)	(0.29)	(0.345)	(0.31)	(0.345)	(0.31)
N	210,891	164,371	185,246	143,918	96,260	67,989	96,260	67,989

Notes: Odd and even columns use the identical model specifications as those in columns 2 and 6 of Table 3, respectively. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Robustness Tests

	P/F Courses		Graded Courses (CD Students)		Graded Courses (DF Students)	
	Full	98-17	Full	98-17	Full	98-17
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Persistence						
# of Credits Attempted	1.003	0.427	0.471***	0.628***	0.783***	0.665***
	(9.535)	(0.729)	(0.156)	(0.141)	(0.107)	(0.097)
1st-Stage F	13	268	349	2327	281	2784
N	153,899	116,332	442,846	342,555	244,524	192,162
# of Credits Earned	-0.569	0.494	0.545***	0.702***	0.702***	0.616***
	(9.174)	(0.72)	(0.148)	(0.132)	(0.099)	(0.09)
1st-Stage F	13	268	349	2327	281	2784
N	153,899	116,332	442,846	342,555	244,524	192,162
Prob of Attempting Next Credit	0.624	0.0434	0.00865	0.0205***	0.0573***	0.0395***
	(0.438)	(0.042)	(0.009)	(0.008)	(0.007)	(0.005)
1st-Stage F	13	268	349	2327	281	2784
N	153,899	116,332	442,846	342,555	244,524	192,162
Prob of Earning Next Credit	0.512	0.0563	0.0213**	0.0330***	0.0557***	0.0392***
	(0.401)	(0.04)	(0.009)	(0.007)	(0.007)	(0.005)
1st-Stage F	13	268	349	2327	281	2784
N	153,899	116,332	442,846	342,555	244,524	192,162
Panel B: Course Choice						
Level of Next-Course Attempted	0.259	-0.0856	0.0725***	0.104***	0.0747***	0.0978***
	(0.707)	(0.087)	(0.026)	(0.021)	(0.026)	(0.021)
1st-Stage F	9	910	240	1093	122	559
N	67,082	50,955	176,156	135,071	57,397	41,936
Level of Next-Course Passed	0.0614	-0.0911	0.0799***	0.0146	0.0609**	0.0939***
	(0.73)	(0.089)	(0.028)	(0.034)	(0.03)	(0.024)
1st-Stage F	7	790	221	722	95	400
N	64,813	49,430	158,930	367,574	44,789	32,307
Panel C: Next-Course Performance						
Pass Next Course	-0.484	–	0.0788***	0.0614**	0.305*	0.162
	(0.485)	–	(0.022)	(0.025)	(0.167)	(0.196)
1st-Stage F	19	–	114	201	17	21
N	8,422	–	77,658	53,212	25,650	15,352
Average Grade from Next Course	-1.564	–	0.216***	0.180***	0.0529	-0.0515
	(2.043)	–	(0.07)	(0.07)	(0.058)	(0.076)
1st-Stage F	19	–	114	201	17	21
N	8,422	–	77,658	53,212	25,650	15,352

Notes: Odd and even columns use the identical model specifications as those in columns 2 and 6 of Table 3, respectively. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Heterogeneous Effects of Course Repetition

	Student Seniority		Course Type	
	First-Year (1)	Non-First-Year (2)	Introductory (3)	Upper-Division (4)
Panel A: Persistence				
# of Credits Attempted	1.139***	0.791***	0.795***	0.806***
	(-0.262)	(-0.143)	(-0.102)	(-0.23)
Mean DV	3.371	4.264	3.878	4.23
(SD)	(-8.153)	(-8.321)	(-7.877)	(-7.449)
N	185,600	238,128	366,400	109,441
# of Credits Earned	1.081***	0.777***	0.753***	0.803***
	(-0.248)	(-0.134)	(-0.096)	(-0.213)
Mean DV	2.626	3.319	3.029	3.421
(SD)	(-7.293)	(-7.071)	(-6.954)	(-6.17)
N	185,600	238,128	366,400	109,441
Prob of Attempting Next Credit	0.069**	0.040***	0.042***	0.033**
	(-0.013)	(-0.008)	(-0.005)	(-0.013)
Mean DV	0.367	0.402	0.406	0.476
(SD)	(-0.483)	(-0.491)	(-0.491)	(-0.501)
N	185,600	238,128	366,400	109,441
Prob of Earning Next Credit	0.071***	0.037***	0.043***	0.040**
	(-0.013)	(-0.008)	(-0.005)	(-0.012)
Mean DV	0.302	0.356	0.345	0.429
(SD)	(-0.46)	(-0.48)	(-0.476)	(-0.497)
N	185,600	238,128	366,400	109,441
Panel B: Course Choice				
Level of Next-Course Attempted	0.085***	0.092***	0.092***	0.065**
	(-0.031)	(-0.029)	(-0.014)	(-0.03)
Mean DV	1.527	1.975	1.811	3.296
(SD)	(-0.829)	(-1.021)	(-0.958)	(-0.882)
N	49,919	89,411	109,364	55,039
Level of Next-Course Passed	0.066*	0.122***	0.088***	0.066**
	(0.034)	(0.033)	(0.016)	(0.032)
Mean DV	1.611	2.02	1.865	3.294
(SD)	(0.881)	(1.034)	(0.989)	(0.901)
N	40,561	80,385	91,884	52,041
Panel C: Subsequent Performance				
Pass Next Course	0.116*	-0.09	0.061***	-0.021
	(0.067)	(0.053)	(0.021)	(0.055)
Mean DV	0.588	0.792	0.674	0.75
(SD)	(0.507)	(0.415)	(0.470)	(0.447)
N	17,593	33,532	49,111	15,919
Average Grade from Next Course	0.290*	-0.0635	0.201***	0.018
	(0.155)	(0.164)	(0.056)	(0.156)
Mean DV	1.412	1.833	1.498	2.25
(SD)	(1.460)	(1.239)	(1.261)	(1.528)
N	17,593	33,532	49,111	15,919

Notes: To save space, results reported in this table are based on our preferred model specification as in column (6) of Table 3 for the post-1998 sample. Results for the full sample are qualitatively similar and available upon request. First-stage F statistics range from 36 to 606. Standard errors (in parentheses) are clustered at the course section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Repetition Effects on Time-to-Degree, Graduation, and Study Pace

	All Students			First-Time Students		
	All	Repeats <=4	Repeats >4	All	Repeats <=4	Repeats >4
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Time to Gradation (Graduates)						
# of Semesters	0.385**	0.286	0.564**	0.241	0.123	0.549***
until Degree Completion	(0.155)	(0.260)	(0.221)	(0.179)	(0.325)	(0.201)
First-Stage F	294	186	161	134	81	87
Mean DV	11.944	10.231	16.4	12.333	10.857	17.5
(SD)	(4.478)	(2.803)	(5.225)	(4.444)	(2.854)	(6.364)
Number of Observations	24,389	23,126	15,835	7,341	6,745	4,102
Panel B: Graduation (All Students)						
Likelihood of Completing	0.159***	0.238***	0.198***	0.142***	0.215***	0.158***
an Undergraduate Degree	(0.006)	(0.009)	(0.010)	(0.005)	(0.009)	(0.008)
First-Stage F	3445	2745	1132	2040	1608	692
Mean DV	0.095	0.078	0.162	0.067	0.058	0.143
(SD)	(0.295)	(0.269)	(0.295)	(0.250)	(0.235)	(0.378)
Number of Observations	39,886	38,373	29,189	20,641	19,767	15,068
Panel C: Study Pace (Graduates)						
# of Semesters Elapsed	-1.238***	–	–	-1.619**	–	–
before Attempting Next Credit	(0.472)	–	–	(0.779)	–	–
First-Stage F	230	–	–	114	–	–
Mean DV	3.4	–	–	2.97	–	–
(SD)	(6.198)	–	–	(4.579)	–	–
Number of Observations	52,770	–	–	17,044	–	–
# of Semesters Elapsed	-1.312***	–	–	-1.675**	–	–
before Completing Next Credit	(0.512)	–	–	(0.837)	–	–
First-Stage F	201	–	–	105	–	–
Mean DV	3.294	–	–	3	–	–
(SD)	(6.436)	–	–	(4.649)	–	–
Number of Observations	50,377	–	–	16,203	–	–
Panel D: Study Pace (All Students)						
# of Semesters Elapsed	-1.078***	–	–	-0.600**	–	–
before Attempting Next Credit	(0.191)	–	–	(0.259)	–	–
First-Stage F	988	–	–	453		–
Mean DV	3.393	–	–	3.326	–	–
(SD)	(5.260)	–	–	(4.728)		–
Number of Observations	148,840	–	–	62,731	–	–
# of Semesters Elapsed	-1.437***	–	–	-1.187***	–	–
before Completing Next Credit	(0.215)	–	–	(0.318)	–	–
First-Stage F	790	–	–	342	–	–
Mean DV	3.459	–	–	3.398	–	–
(SD)	(5.616)	–	–	(5.014)	–	–
Number of Observations	129,859	–	–	53,203	–	–

Notes: 1) Results in Panels A and B are obtained for student-level data where a student's average exposure to the grade replacement policy during the observation period is used as an instrument for his/her total number of repeated courses. Covariates included in the regressions include the student's gender, ethnicity, transfer and in-state status, and age at college entry (in 4 categories). Standard errors (in parentheses) are clustered at the student level. 2) Results in Panels C and D are obtained for student-course-section level data using the model specification in columns 2 and 6 of Table 3. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Repetition Effects on Subsequent-Attempt Performance

	Full Sample			Post 1998		
	B or above	Pass	Grade Point (0 to 4 scale)	B or above	Pass	Grade Point (0 to 4 scale)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Courses						
GP × Course Repetition	0.155***	0.174***	0.611***	0.153***	0.151***	0.582***
	(0.005)	(0.004)	(0.012)	(0.004)	(0.003)	(0.009)
Mean DV	0.468	0.882	1.82	0.434	0.881	1.739
(SD)	(0.499)	(0.322)	(1.426)	(0.496)	(0.324)	(1.107)
N	603,330	603,330	603,330	475,845	475,845	475,845
Panel B: Same-Instructor Courses						
GP × Course Repetition				0.109***	0.147***	0.490***
				(0.005)	(0.005)	(0.012)
Mean DV				0.418	0.889	1.672
(SD)				(0.495)	(0.315)	(1.191)
N				418,973	418,973	418,973
Panel C: Solo-Section Courses						
GP × Course Repetition				0.131***	0.156***	0.540***
				(0.009)	(0.008)	(0.02)
Mean DV				0.397	0.905	1.724
(SD)				(0.491)	(0.294)	(1.026)
N				82,779	82,779	82,779

Notes: The full sample and post-1998 results are based on the model specifications in columns 2 and 6 of Table 3, respectively. First-stage F statistics range from 229 to 3678. Standard errors (in parentheses) are clustered at course section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Mechanism Investigation: Number of Subsequent Same-Subject Credits Attempted

	Full Sample		Post 1998					
	All Courses		All Courses		Same-Instructor Sections		Solo-Section Courses	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GP × Course Repetition	0.523***	-0.193	0.805***	0.611**	0.882***	1.162	0.914**	1.06
	(0.103)	(0.310)	(0.094)	(0.248)	(0.274)	(0.791)	(0.456)	(1.264)
Grade Difference		1.172***		0.295		-0.623		-0.216
		(0.347)		(0.276)		(1.061)		(1.520)
Grade F as the Base								
Grade C+	3.136***	3.100***	3.297***	3.260***	3.240***	3.229***	6.152***	6.184***
	(0.124)	(0.124)	(0.120)	(0.119)	(0.120)	(0.119)	(0.680)	(0.679)
Grade C	2.500***	2.488***	2.796***	2.762***	2.746***	2.734***	4.663***	4.696***
	(0.062)	(0.061)	(0.059)	(0.058)	(0.058)	(0.058)	(0.314)	(0.313)
Grade C-	2.202***	2.162***	2.316***	2.279***	2.244***	2.232***	4.058***	4.083***
	(0.113)	(0.113)	(0.112)	(0.112)	(0.112)	(0.112)	(0.744)	(0.743)
Grade D+	1.009***	1.124***	1.114***	1.084***	1.047***	0.982***	1.833***	1.844***
	(0.071)	(0.082)	(0.074)	(0.080)	(0.084)	(0.102)	(0.295)	(0.338)
Grade D	0.883***	0.940***	0.880***	0.791***	0.778***	0.720***	1.420***	1.509***
	(0.034)	(0.043)	(0.037)	(0.047)	(0.041)	(0.058)	(0.138)	(0.182)
Grade D-	0.614***	0.724***	0.680***	0.715***	0.671***	0.675***	0.509	0.627
	(0.083)	(0.093)	(0.089)	(0.096)	(0.102)	(0.106)	(0.381)	(0.446)
Grade Diff. x C+		-0.626**		-0.450*		0.125		-2.854***
		(0.256)		(0.247)		(0.619)		(0.902)
Grade Diff. x C		-0.372***		-0.273**		0.133		-1.231**
		(0.137)		(0.132)		(0.438)		(0.508)
Grade Diff. x C-		-0.824***		-0.640***		-0.298		-1.956***
		(0.186)		(0.193)		(0.469)		(0.608)
Grade Diff. x D+		0.0297		0.148		0.267		-0.0837
		(0.106)		(0.101)		(0.209)		(0.379)
Grade Diff. x D		0.0569		0.216***		0.257***		-0.271*
		(0.050)		(0.054)		(0.096)		(0.161)
Grade Diff. x D-		-0.0762		-0.0317		-0.0953		-0.303
		(0.099)		(0.104)		(0.239)		(0.376)
F Statistics and P-Values		8.61		9.61		1.33		5.35
on Joint Hypotheses		[0.000]		[0.000]		[0.229]		[0.000]
First-Stage F	566	203	3633	890	872	302	597	155
Mean DV	3.804	3.804	3.946	3.946	5.423	5.423	8.231	8.231
(SD)	(7.740)	(7.740)	(8.669)	(8.669)	(10.748)	(10.748)	(12.132)	(12.132)
N	603,330	603,330	475,837	475,837	418,982	418,982	84,357	84,357

Notes: The full sample and post-1998 results are based on the model specifications in columns 2 and 6 of Table 3, respectively. Standard errors (in parentheses) are clustered at course section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Mechanism Investigation: All Outcomes

	Full Sample		Post 1998					
	All Courses		All Courses		Same-Instructor		Solo-Section	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# of Credits Attempted	0.523***	-0.193	0.805***	0.611**	0.882***	1.162	0.914**	1.06
	(0.103)	(0.310)	(0.094)	(0.248)	(0.274)	(0.791)	(0.456)	(1.264)
P-Values of F-Tests		[0.000]		[0.000]		[0.229]		[0.000]
N	603,330	603,330	475,837	475,837	418,982	418,982	84,357	84,357
# of Credits Earned	0.494***	-0.138	0.762***	0.628***	0.833***	1.117	0.778*	0.648
	(0.096)	(0.288)	(0.089)	(0.233)	(0.259)	(0.748)	(0.433)	(1.198)
P-Values of F-Tests		[0.000]		[0.000]		[0.075]		[0.000]
N	603,330	603,330	475,837	475,837	418,982	418,982	84,357	84,357
Prob of Attempting Next Credit	0.030***	-0.042*	0.040***	-0.005	0.027**	-0.02	0.027	-0.004
	(0.008)	(0.025)	(0.005)	(0.013)	(0.014)	(0.040)	(0.017)	(0.046)
P-Values of F-Tests		[0.000]		[0.000]		[0.031]		[0.002]
N	603,330	603,330	475,837	475,837	418,982	418,982	84,357	84,357
Prob of Earning Next Credit	0.030***	-0.042*	0.042***	0.001	0.034**	-0.004	0.034**	-0.003
	(0.008)	(0.023)	(0.005)	(0.012)	(0.013)	(0.038)	(0.017)	(0.046)
P-Values of F-Tests		[0.000]		[0.000]		[0.235]		[0.002]
N	603,330	603,330	475,837	475,837	418,982	418,982	84,357	84,357
Level of Next-Course Attempted	0.071***	0.003	0.090***	0.095**	0.116***	0.2	0.123***	0.192
	(0.016)	(0.068)	(0.013)	(0.043)	(0.030)	(0.125)	(0.039)	(0.123)
P-Values of F-Tests		[0.000]		[0.000]		[0.000]		[0.022]
N	210,891	210,891	164,371	164,371	142,806	142,806	42,291	42,291
Level of Next-Course Passed	0.069***	0.006	0.086***	0.090*	0.101***	0.143	0.120***	0.204
	(0.017)	(0.076)	(0.014)	(0.051)	(0.029)	(0.126)	(0.040)	(0.140)
P-Values of F-Tests		[0.000]		[0.000]		[0.004]		[0.045]
N	185,246	185,246	143,918	143,918	125,679	125,679	39,609	39,609
Pass Next Course	0.076***	0.076	0.059***	0.043	0.07	-0.115	-0.117	-0.335*
	(0.017)	(0.052)	(0.018)	(0.039)	(0.043)	(0.065)	(0.094)	(0.198)
P-Values of F-Tests		[0.107]		[0.498]		[0.010]		[0.048]
N	96,260	96,260	67,989	67,989	54,698	54,698	11,928	11,928
Average Grade from Next Course	0.209***	0.08	0.206***	0.124	0.163	0.112	-0.358	-1.095*
	(0.048)	(0.139)	(0.049)	(0.106)	(0.114)	(0.431)	(0.277)	(0.576)
P-Values of F-Tests		[0.006]		[0.192]		[0.002]		[0.044]
N	96,260	96,260	67,989	67,989	54,698	54,698	11,928	11,928

Notes: The full sample and post-1998 results are based on the model specifications in columns 4 and 10 of Table 3, respectively. First-stage F statistics range from 37 to 2847. Standard errors (in parentheses) are clustered at course section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix

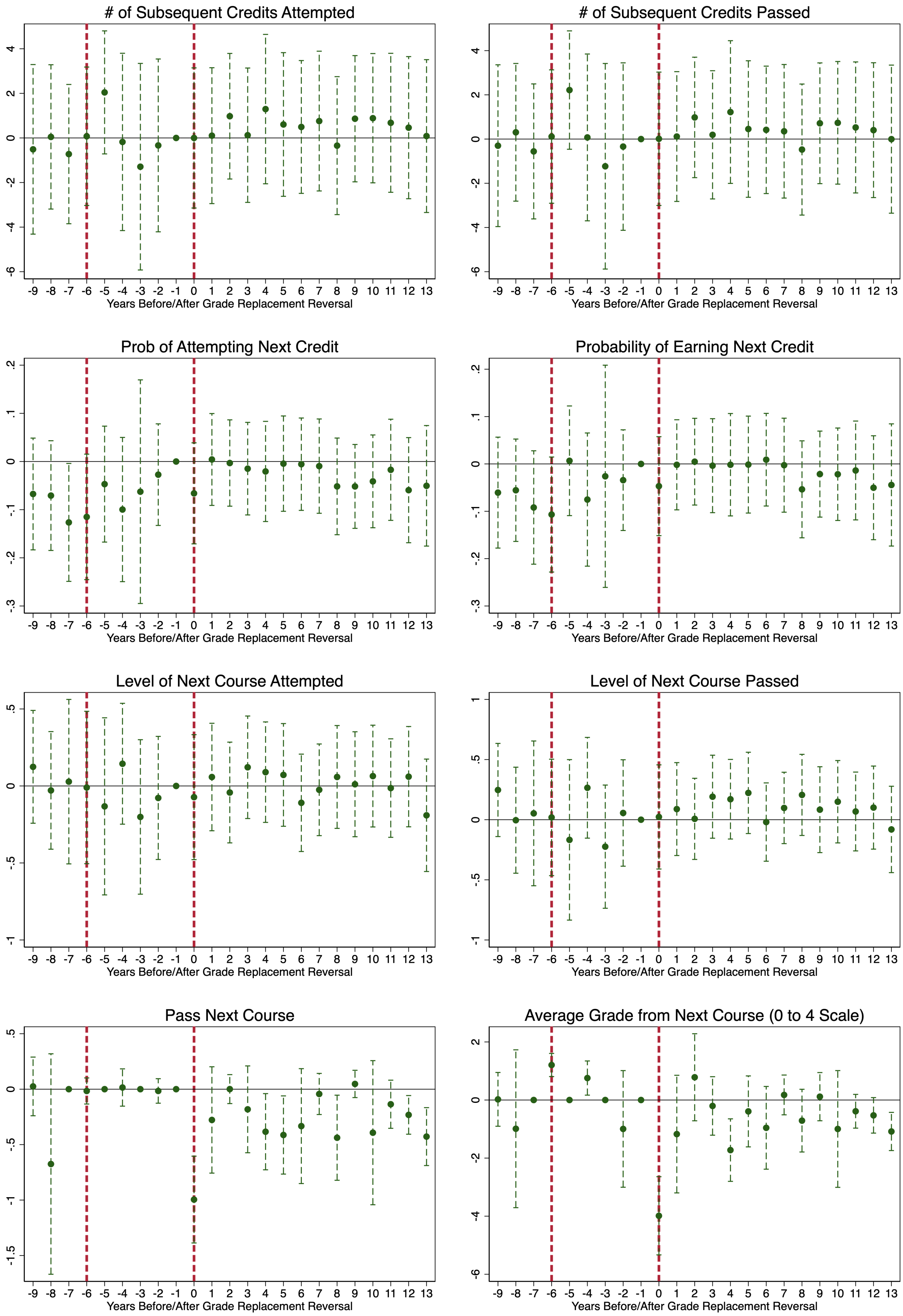


Figure A1: All Outcomes (P/F Courses)

Table A1: Event Study Analysis for Balanced Panels (1998-2011)

	First-Stage	Persistence				Course Choice		Subsequent Performance	
	Prob of Course Repetition	# of Credits Attempted	# of Credits Earned	Prob of Attempting Next Credit	Prob of Earning Next Credit	Level of Next-Course Attempted	Level of Next-Course Passed	Pass Next Course	Average Grade from Next Course
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean DV in 2001	0.133	4.131	3.233	0.411	0.347	2.01	2.082	0.706	1.716
(SD)	(0.340)	(8.719)	(7.670)	(0.492)	(0.476)	(1.051)	(1.070)	(0.456)	(1.354)
Years -4 to -3	-0.0226	-0.0101	0.00166	0.00000923	-0.000872	-0.0226	-0.0101	0.00166	0.00000923
	(0.014)	(0.086)	(0.078)	(0.007)	(0.006)	(0.014)	(0.086)	(0.078)	(0.007)
Year -2 (to -1)	-0.000104	-0.00345	-0.00811	-0.00295	-0.00312	-0.000104	-0.00345	-0.00811	-0.00295
	(0.015)	(0.074)	(0.066)	(0.006)	(0.006)	(0.015)	(0.074)	(0.066)	(0.006)
Years 0 to 1	0.0435***	0.0978*	0.0776	0.00573	0.00418	0.0435***	0.0978*	0.0776	0.00573
	(0.012)	(0.058)	(0.051)	(0.005)	(0.004)	(0.012)	(0.058)	(0.051)	(0.005)
Years 2 to 3	0.0745***	0.161***	0.150***	0.00549	0.00575	0.0745***	0.161***	0.150***	0.00549
	(0.012)	(0.060)	(0.054)	(0.005)	(0.004)	(0.012)	(0.060)	(0.054)	(0.005)
Years 4 to 5	0.104***	0.282***	0.274***	0.0123**	0.0138***	0.104***	0.282***	0.274***	0.0123**
	(0.012)	(0.061)	(0.055)	(0.005)	(0.005)	(0.012)	(0.061)	(0.055)	(0.005)
Years 6 to 7	0.124***	0.299***	0.291***	0.0141***	0.0155***	0.124***	0.299***	0.291***	0.0141***
	(0.012)	(0.063)	(0.057)	(0.005)	(0.004)	(0.012)	(0.063)	(0.057)	(0.005)
Years 8 to 9	0.139***	0.402***	0.388***	0.0160***	0.0174***	0.139***	0.402***	0.388***	0.0160***
	(0.012)	(0.063)	(0.057)	(0.005)	(0.005)	(0.012)	(0.063)	(0.057)	(0.005)
# Courses	2,461	2,461	2,461	2,461	2,461	2,031	1,982	1,360	1,360
# Course-Sections	27,219	27,219	27,219	27,219	27,219	19,346	18,410	8,540	8,540
N	594,819	594,819	594,819	594,819	594,819	207,392	182,045	94,486	94,486

Notes: This table reports the event study analysis results for courses offered every year from 1998-2011. All models control course-section fixed effects, academic progress fixed effects, and initial-attempt-performance fixed effects. Covariates include student gender, transfer status, declared major, the student's last-observed semester and cumulative GPAs prior to the time of observation, and their interactions with the student's initial-attempt grade. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Additional Robustness Tests

	Persistence				Course Choice		Performance	
	# of Credits Attempted	# of Credits Earned	Prob of Attempting Next Credit	Prob of Earning Next Credit	Level of Next-Course Attempted	Level of Next-Course Passed	Pass Next Course	Grade from Next Course
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Students								
GP x Repetition	0.805***	0.762***	0.040***	0.042***	0.090***	0.086***	0.059***	0.206***
	(0.094)	(0.089)	(0.005)	(0.005)	(0.013)	(0.014)	(0.018)	(0.049)
N	475,837	475,837	475,837	475,837	164,371	143,918	67,989	67,989
Panel B: Tracking Students for 10 Years								
GP x Repetition	0.739***	0.694***	0.042***	0.043***	0.094***	0.089***	0.062***	0.194***
	(0.090)	(0.084)	(0.005)	(0.004)	(0.013)	(0.014)	(0.019)	(0.052)
N	202,062	202,062	202,062	202,062	69,774	61,389	25,572	25,572
Panel C: Non-Transfer Students								
GP x Repetition	0.741***	0.705***	0.0418***	0.0438***	0.1000***	0.0961***	0.0733**	0.177**
	(0.119)	(0.112)	(0.006)	(0.006)	(0.018)	(0.021)	(0.032)	(0.082)
Mean of DV	4.106	3.234	0.412	0.357	1.887	1.939	0.691	1.6
(SD)	(8.484)	(7.350)	(0.493)	(0.480)	(1.038)	(1.050)	(0.466)	(1.241)
N	227,730	227,730	227,730	227,730	69,749	594,79	24,360	24,360

Notes: This table reports the results obtained using the model specification in column 6 of Table 3. Estimates in Panel A are identical to the estimates shown in Table 5. Results estimated for the full sample using the model specification in column 2 of Table 3 are qualitatively similar. To save space, they are not reported but are available upon request. First-stage F statistics range from 389 to 2444. Standard errors (in parentheses) are clustered at course-section level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$