HCEO WORKING PAPER SERIES

Lasting Effects of Retaking College Admission Exams

Veronica Frisancho
Sebastian Gallegos
Constanza González

Working Paper

2026-004

03-2026

HUMAN CAPITAL AND ECONOMIC OPPORTUNITY GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Lasting Effects of Retaking College Admission Exams*

Veronica Frisancho[a], Sebastian Gallegos[b] and Constanza González[c]

[a]CAF Development Bank of Latin America
[b]UAI Business School, IZA and HCEO-UChicago
[c]UAI Business School

March 10, 2026

Abstract

Do second chances at a high-stakes admission exam yield long-term gains? Leveraging fifteen years of Chilean administrative data and an RDD, we examine the causal effects of retaking on educational and labor market trajectories. Narrowly missing a preferred program cutoff triggers a 44% increase in retaking, leading to substantial score gains (0.27 SD) and improved placement and enrollment chances. However, these immediate gains do not persist. Retakers graduate at the same rate and from programs with similar earnings and employability profiles as their counterfactual peers. Our results suggest that retaking serves as a reshuffling mechanism yielding null net welfare gains.

Keywords: High-Stakes Exams, Developing Countries, Latin America, Educational Policy, College, Enrollment, Graduation, Choices.
JEL codes: J62; N36; I24; I25; I28.

*We thank seminar participants at the LACEA-LAMES Annual Meeting, the Global Labor Network (GLO) Seminar, SECHI Annual meeting, and academic seminars at CAF, UAI, U. de la Plata, UNAB, and PUC Chile for helpful comments. We also thank the Ministry of Education of the Government of Chile and the DEMRE for access to the administrative data.

(a) Frisancho: Chief Economist at CAF, Development Bank of Latin America and the Caribbean, Buenos Aires, Argentina.
(b) Gallegos (corresponding author): Associate Professor of Economics, Business School, Universidad Adolfo Ibañez, Chile, IZA@LISER Research Fellow and member of the Human Capital and Economic Opportunity Global Working Group (HCEO) at The University of Chicago. Email: sebastian.gallegos@uai.cl.
(c) Business School, Universidad Adolfo Ibañez, Chile.

1 Introduction

Most educational systems around the world grade and rank students according to their performance in one or more subjects. When these ranks are used to allocate scarce university seats, standardized exams become high-stakes pivots for a student's future. Consequently, scoring just below or above a given threshold can significantly alter educational trajectories (Kirkeboen et al., 2016; Machin et al., 2020a; Chetty et al., 2023). A relatively novel—and still thin—strand of literature examines how students respond to these high-stakes outcomes through retaking behavior. To date, this research has focused almost exclusively on immediate outcomes, consistently documenting gains in test scores and admission probabilities in the U.S., China, and Greece (Goodman et al., 2020; Kang et al., 2024; Bizopoulou et al., 2024). However, the literature remains largely silent on whether these initial gains translate into meaningful long-term success, leaving a gap in our understanding of how retaking ultimately shifts enrollment persistence, graduation rates, and eventual employment prospects.

This paper contributes to filling this gap by studying both the immediate and long-term causal effects of retaking on educational and labor market outcomes. We combine 15 years of administrative data from the centralized college admission system in Chile with a Regression Discontinuity Design (RDD). Because placement in this system is driven exclusively by exam performance and stated preferences for university-major combinations, we are able to isolate these two key channels and unravel the mechanisms behind our results. Beyond immediate placement, we track the impact of retaking on enrollment and graduation. Finally, by linking these records to program-level data on employability and wages, we provide some of the first evidence of the causal effects of retaking on long-term earnings and employment prospects.

Our analysis focuses on the 2010 cohort of first-time exam takers and applicants, for whom we reconstruct educational trajectories over a 15-year period. First, we use centralized admission records to observe retaking behavior, shifts in test performance, and evolving preferences, as well as final placement outcomes across successive attempts. Second, we link these files to longitudinal administrative records from the Ministry of Education (MINEDUC) for the period 2010–2025, allowing us to track enrollment and graduation across all public and private universities in Chile. Finally, we incorporate program-level labor market indicators derived from national tax records (Servicio de Impuestos Internos, SII). These data provide expected employment probabilities and wage distributions across tertiary programs, enabling us to estimate how retaking shapes long-term earning potential and labor market integration.

To address the non-random nature of the decision to retake, we exploit the exogenous variation provided by admission cutoffs. Following the empirical strategy used in recent studies of centralized systems in Greece, China, and the U.S. (Bizopoulou et al., 2024; Kang et al., 2024; Goodman et al., 2020), we compare students who narrowly missed the threshold for their most preferred program to those who barely cleared it. Our first-stage estimates show that "near-miss" applicants are 4.3 percentage points more likely to retake the exam—a 44% relative increase that aligns closely with findings in other international contexts. This retaking behavior leads to a significant performance gain of 0.27 standard deviations between the first and second attempts. Consequently, threshold-induced retakers see improvements in admission and enrollment rates. Crucially, we show that these gains are driven primarily by improved test scores rather than changes in application strategy or preferences.

Despite these substantial immediate gains, the advantages of retaking appear to dissipate over the long run. Following students for fifteen years after their first attempt, we find that the initial boost in placement does not translate into better educational or labor market outcomes; retakers graduate at the same rate as non-retakers and enter programs with nearly identical employability and expected earnings. These results suggest that while retaking helps marginally weaker students access their desired programs, it does not fundamentally alter their long-term economic trajectory. Notably, however, we find no evidence that these students "lag behind" once enrolled; the fact that graduation rates remain stable implies that retakers are as prepared for the rigors of university as their peers who were admitted on their first attempt. All in all, our findings provide a cautionary tale: second chances may improve admission and enrollment outcomes, but these seem to be decoupled from lasting gains in graduation rates and the labor market.

Students' responses to failure in the context of centralized admission systems remain an understudied topic. A nascent but growing literature explores the decision to retake exams and its subsequent impact on short-run outcomes. Early evidence from Turkey (Frisancho et al., 2016) and the United States (Goodman et al., 2020) suggests that the opportunity to retake leads to substantial score improvements, which consequently leads to better college admission prospects. The improvement in scores in Turkey and the US were particularly salient among less-advantaged students, thereby narrowing gaps in college access.

Building on these findings, a recent set of studies has leveraged the discrete nature of centralized admission systems to implement fuzzy Regression Discontinuity Designs (RDD). By exploiting exogenous variation in retaking probabilities around admission cutoffs, these papers isolate the causal gains of a second attempt. Bizopoulou et al. (2024) take advantage of a minimum qualifying score to get access to tertiary education in Greece and find that low-achieving students who retake national admission exams obtain better quality placements due to improved performance and more ambitious preferences the second time around. Similarly, in the competitive context of the French Grandes Écoles, Landaud and Maurin (2020) document that repeating a year allows marginal students to improve their performance by roughly 25 percentile ranks. Most recently, Kang et al. (2024) use data for China to exploit discontinuities in the retaking probability around the entry cutoff for selective universities (schools in the top two tiers, accessible to the top 25% of students). Again, retaking generates increases in exam performance (0.47 SD) and relative ranking (11 percentage points).

While the evidence consistently documents immediate gains in scores and placement, our paper extends this literature by tracking these effects into graduation and the labor market to determine if such "second chances" yield lasting economic returns.[1]

[1] Previous research highlights the complexities of these systems; for instance, Krishna et al. (2018) use a structural model of the Turkish exam system to show that *limiting* retakes can actually *improve* welfare for most agents. This is largely due to a general equilibrium effect, where retakers crowd the market and impose negative externalities by driving up admission cutoffs. Similarly, a recent study in Italy finds that restricting retakes at a major Italian university improved student efficiency, leading to lower dropout rates, higher exam pass rates, increased credit accumulation,

Our study builds on recent contributions and poses at least three key advantages. First, while the pioneering work of Landaud and Maurin (2020), Krishna et al. (2018), Kang et al. (2024), and Goodman et al. (2020) has established a robust understanding of the immediate effects of retaking on academic performance and initial placement, the longer-term trajectories regarding graduation rates and labor market outcomes remain an open empirical question. Our focus on these lasting outcomes allows for a more complete assessment of the welfare implications of retaking.

Second, our setting allows us to disentangle the underlying the drivers of change, i.e., performance and/or preference shifts. While the existing literature primarily focuses on test score gains, the Chilean centralized system provides a direct, unconditioned look at students' ranked preferences. By observing these lists across successive years, we can determine whether retakers secure better placements simply by scoring higher, or by strategically changing their university-major choices. This level of granularity is often unavailable in other contexts.

Third, other studies relying on RDDs focus on minimum entry cutoffs (Bizopoulou et al., 2024) or thresholds to ration access to elite universities (Kang et al., 2024), which produces results for students at the bottom or the top of the score distribution, respectively. By relying on cutoffs defined at the individual level, our analysis sample is likely more representative of a broader segment of the performance distribution, providing more generalizable insights.

The remainder of this article is organized into five sections. The following section presents the institutional background. Section 3 describes the data and defines the outcomes variables. Section 4 presents the estimation strategy while section 5 presents the results. Section 6 concludes by discussing the broader implications of our results.

2 College Admissions and Retaking in Chile

College Admissions. The Chilean college admission process is largely centralized, with coordinated application and selection procedures for the majority of higher education institutions across the country. This system was designed to ensure a standardized, transparent, and fair method of admitting students into various undergraduate programs. Central to this system is the Prueba de Selección Universitaria (PSU), similar to the SAT in the United States.[2]

These standardized tests are critical in determining student eligibility and placement in universities. Test takers complete exams in mathematics and language. Consistent with most OECD countries (Kirkeboen et al., 2016), the Chilean system is characterized by early specialization: students apply directly to a specific major upon entry. This stands in contrast to systems like that of the United States, where students generally enter a broad program and defer their major choice until later in their undergraduate studies. Prospective students apply to university-major combina-

and improved on-time graduation rates (Bratti et al., 2024).

[2]Chile possesses one of the world's longest-standing centralized admission systems, established in 1967 with the Academic Aptitude Test (PAA). Although the exam has undergone several structural and nomenclature revisions—evolving into the University Selection Test (PSU) in 2004, the Transition Test (PDT) in 2021, and the Higher Education Access Test (PAES) in 2023—the fundamental administrative cycle of examination, results disclosure, preference submission, and centralized enrollment has remained consistent for over five decades.

tions ranked based on personal preference, via a centralized process. The students' application score for each preference is a major-specific weighted average of the PSU test scores that are relevant to compete for a seat in each major (about 80%) and their high school GPA (about 20%). The allocation process ranks all the students applying to any given major using their respective major-specific scores.

Admission to Chilean tertiary education is governed by a centralized annual cycle. Each December, graduates take a suite of mandatory (Mathematics and Language) and elective (History or Science) exams. Crucially, the selection criteria—including the weights assigned to each exam and to high school records—are predetermined by the universities before the tests are administered. Once scores are disclosed in January, students use a centralized platform to apply to a ranked list of programs, with placement determined solely by their weighted scores and stated preferences.

The allocation process ranks all applicants to a given major by their weighted scores, following a student-proposing deferred acceptance algorithm (Abdulkadiroğlu et al., 2014; Rios et al., 2021). In the first round, students are tentatively assigned to their top-ranked preference. Each program then rejects students whose scores fall below the program's capacity limit. In each subsequent round r, students rejected in round $r - 1$ are considered for their next highest-ranked preference. Programs then re-evaluate all currently held and newly arriving applicants, again rejecting those in excess of capacity and updating their provisional admission lists. This iterative process concludes when all students are either matched to a program or have exhausted their preference list. This mechanism naturally defines a score cutoff for each program—the score of the last student admitted. Once finalized, students are matched to at most one program (their highest possible preference).At this stage, admitted students must choose whether to enroll, while those without an offer may seek placement in non-selective institutions outside the centralized system.

The centralized admission process is administered by the Department of Educational Evaluation, Measurement, and Registry (DEMRE). As of the 2025 cycle, 47 universities—representing approximately 81% of all universities in Chile—participate in the system, offering roughly 1,800 academic programs. Participation has expanded significantly over time; in 2010, the system was restricted to the 25 member institutions of the Council of Rectors of Chilean Universities (CRUCH). These long-standing institutions are widely regarded as highly selective and academically rigorous. In subsequent years, additional universities have joined the centralized system as they met the necessary participation requirements.

Unlike the flexible, multi-date schedule of the U.S. SAT, the Chilean system—similar to those in Greece and Turkey—relies on fixed, nationally synchronized testing windows that have historically occurred only once per year. Furthermore, during our study period, scores were non-transferable between admission cycles; applicants were required to retake the examination to participate in a subsequent year's process.

Retaking. Over the last ten years, the number of students taking the exam has fluctuated between 225,000 and 250,000. Retaking is prevalent in the Chilean context; each year, between 15% and 20% test-takers are making their second or subsequent attempt. The vast majority of these students sit for the exam in the year immediately following their first attempt. On average, three out of four retakers in any given round are individuals who first took the exam just one year prior. Over time, the retake rate in Chile has remained relatively stable, at levels comparable to other countries with single annual admission cycles, such as Greece (15%–20%) and China (13%–15%), yet significantly lower than in systems with multiple annual testing windows, such as the United States (54%).

The decision to retake the exam reflects a trade-off between expected benefits and costs and can be viewed as an investment in human capital. Students incur costs today in exchange for higher expected returns through improved college placement. In the Chilean context, the potential gain arises from obtaining a higher score in the next admission cycle, one year later, which may allow access to more selective programs with higher expected returns. There is no formal penalty for retaking. Universities do not distinguish between first-time and repeat applicants, and there are no limits on attempts or age restrictions.

Retaking, however, is costly. Students delay enrollment by one year, implying foregone time and potential earnings. They must also pay a registration fee of about USD 40–50; this fee is waived for eligible first-time applicants but not for repeat takers. In addition, many students purchase preparatory services, private tutoring, or paid online platforms.[3] Students will retake if the expected returns from improved placement exceed these time and monetary costs.

3 Data

3.1 Sources and Outcomes

Sources. Our primary data source consists of administrative records from the centralized admission system managed by DEMRE between 2010 and 2025, we supplement these data with further administrative records from the Ministry of Education (MINEDUC) and aggregated information from the Chilean Internal Revenue Service, as we describe below.

The DEMRE data cover the population of test-takers in Chile in each exam round and provide baseline demographic and academic information, including high school GPA, type of high school, gender, household income, and parental schooling. The records also include exam scores for each mandatory subject (language and mathematics), application scores, and the submitted rank-ordered list of preferred university–major combinations, as well as admission results.

We complement these records with administrative data from the MINEDUC, which compiles annual information from all higher education institutions in Chile. Using encrypted individual identifiers, we merge the DEMRE files with MINEDUC records that track students' enrollment and graduation histories. These records allow us to follow individuals over time within the higher education system.

Finally, we incorporate program-level labor market indicators from the Chilean Higher Education Information Service (SIES). The SIES aggregates individual tax records provided by the Chilean Internal Revenue Service (Servicio de Impuestos Internos, SII) to the university-program level. The

[3]These expenditures are private and are not recorded in administrative data, either in our dataset or in any centralized source. As a result, they cannot be directly measured.

data released for 2024 and 2025 report employability rates and average income levels for 1,090 distinct university programs. These indicators allow us to characterize the expected labor market prospects associated with different tertiary programs.

Outcomes. We categorize our outcomes into two primary groups: immediate academic results and long-term educational and labor market trajectories.

Immediate Outcomes. Immediate outcomes include admission exam scores, preferences, admission offers, and enrollment results. We use the scores standardized within the cohort of test takers. We use students' submitted preferences over university programs to measure application behavior proxied by the selectivity of each applicant's portfolio using the average cutoff score of the listed programs.[4] In additional analyses, we examine alternative metrics of application behavior, including the total number of programs listed, the selectivity of the top-ranked choice, the dispersion (interquartile range) of admission thresholds across the preference list, and the share of selective programs in the application portfolio.

The admission outcomes come directly from the centralized matching mechanism. Because admission offers are determined by exam performance and submitted preferences, these outcomes allow us to assess whether improvements in scores or changes in application behavior translate into improved placement.

Finally, we observe actual enrollment decisions using the MINEDUC administrative records. These data allow us to observe whether individuals enroll in tertiary education and the institution and program in which they enroll, providing a direct measure of realized educational choices rather than admission offers alone.

Long-Term Outcomes. Our long-term outcomes include graduation, time to graduation, and prospects of employability and earnings at the program level.

We construct graduation outcomes at the individual level using the MINEDUC administrative records for all available years. These data allow us to observe whether students graduate, the exact date of graduation, and the university-program of their degree. Because the records cover the full population of higher education institutions in Chile and are available annually from 2010 to 2025, we can follow individuals for up to fifteen years after their first attempt at the admission exam in 2010. This feature of the administrative data is particularly important for studying graduation outcomes. Graduation is inherently difficult to measure because students complete their programs at different times, and shorter observation windows may miss late graduates. By tracking the full population over a long horizon and observing the exact graduation date for each individual, our data allow us to construct precise measures of both degree completion and time to graduation.

We characterize students' labor market prospects using program-level indicators of employability and expected earnings from the SIES dataset. These measures include employment rates up to two

[4] To ensure our measure of selectivity is exogenous to the 2010 applicant pool, we rely on the 2009 cutoff scores as a fixed benchmark for program selectivity.

years after graduation and average earnings in the fourth year after graduation. Because these indicators are constructed using aggregated tax records from the Chilean Internal Revenue Service (SII), they provide a consistent measure of the expected labor market outcomes associated with each tertiary program. We use these outcomes allow us to assess whether the initial gains associated with retaking translate into meaningful differences in long-run educational attainment and expected labor market prospects.

3.2 Sampling

Our main analysis sample comprises applicants who took the university entrance exam in 2010 for the first time. We exclude test-takers who participated in a previous admission process. These criteria define a working sample of about 46 thousand students. We follow these students over time and classify their retaking status based on whether they retake the admission exam in 2011, the year immediately following their first attempt.[5]

We define our running variable as the margin between a student's application score for their top-ranked program and the score of the last student admitted to that same program. Our RDD analysis uses a sample of 19,387 observations falling within a ±27.5-point bandwidth of the admission threshold. This bandwidth is selected using the MSE-optimal selector proposed by Calonico et al. (2019) for the first-stage specification. We note that, in principle, the optimal bandwidth in a regression discontinuity design is outcome-specific, as it depends on the variance of the corresponding dependent variable. To facilitate comparability across specifications and maintain a common analytical bandwidth sample, we adopt the first-stage optimal bandwidth for all outcome estimations. Importantly, our results are robust to the use of outcome-specific optimal bandwidths and to alternative bandwidth choices (see Table B.3 in Appendix B).

3.3 Descriptive Statistics

Table 1 presents the mean sociodemographic characteristics for students in both the working and the bandwidth samples, further disaggregated by retaker status. Overall, both samples exhibit similar demographic profiles and high school backgrounds. However, students in the bandwidth sample appear slightly more advantaged in terms of household income and parental education. In both the working and bandwidth samples, retakers are more likely to have lower high school GPAs and come from more advantaged socioeconomic backgrounds compared to non-retakers.

The last column of Table 1 presents the mean characteristics of compliers, computed following Goodman et al. (2020) and the methods developed by Abadie et al. (2002) and Abadie (2003). In our setting, compliers are the students who only retake because they narrowly missed the cutoff of their top preference. If they had barely passed, they would have enrolled instead. Overall, we find that compliers are relatively similar to all retakers along observable dimensions. There are modest

[5] We do so because most retaking behavior occurs immediately after the first attempt, while delayed retakes represent a small and quickly diminishing share of the cohort.

differences in a few covariates: compliers are more likely to be female, live in the capital city, and have higher household income.

Table 2 presents descriptive statistics on admission exam performance, placement and enrollment, comparing the full sample (Columns 1–3) with the bandwidth sample (Columns 4–6). As expected, first-attempt scores are higher on average within the bandwidth sample, reflecting the fact that these students scored closer to the admission cutoffs of their top-ranked programs.

Table 1: Baseline Characteristics (Working and Bandwidth Samples)

	Sample			Bandwidth Sample			
	(1) All	(2) Non-retaker	(3) Retaker	(4) All	(5) Non-retaker	(6) Retaker	(7) Complier
Retakers	0.18	-	1.00	0.15	-	1.00	-
Female	0.51	0.50	0.52	0.49	0.49	0.48	0.54
GPA Score	610	612	603	620	621	614	618
Type of School							
Public School	0.30	0.30	0.29	0.29	0.29	0.29	0.27
Voucher School	0.49	0.48	0.53	0.48	0.47	0.50	0.50
Private School	0.21	0.21	0.18	0.23	0.23	0.20	0.23
No information	0.01	0.01	0.00	0.01	0.00	0.01	0.00
Capital Region	0.31	0.30	0.33	0.34	0.34	0.35	0.41
Parents' Schooling							
Mother's years of education	12.56	12.51	12.76	12.65	12.60	12.92	12.94
Father's years of education	12.75	12.69	12.99	12.85	12.80	13.10	13.55
Household Income	566.05	565.31	569.32	589.19	586.28	605.20	679.29
Observations	46,367	37,810	8,557	19,387	16,400	2,987	

Notes: Table 1 reports mean characteristics for our working sample (columns 1 to 3), bandwidth sample (columns 4 to 6), and for the compliers (cutoff-induced retakers) in column 7, computed following Abadie (2003) and Abadie et al. (2002).The GPA Score represents the high school grade point average, and is reported as a standardized score ranging from roughly 150 to 850 points. Household income is reported in monthly thousands of Chilean pesos (at the time, roughly 500 CLP = 1 USD).

Looking at the first attempt in 2010, non-retakers perform 0.08 standard deviations (SD) above prospective retakers in the working sample.[6] In their second attempt, retakers exhibit a sizable (descriptive) increase of about 0.21 SD on standardized average scores. With this improvement, the average retaker in their second attempt is able to score above the average non-retaker. In the bandwidth sample, the initial gap between retakers and non-retakers is smaller, but the former also experience substantial increases in their scores as they give the exam another try. In this sub-sample, retakers also surpass the performance of non-retakers their second time around.[7]

In terms of admission outcomes, retakers are at a clear disadvantage following their first attempt compared to those who do not retake. In their initial application cycle, only 17% of future retakers are placed in their first-preference program, whereas 43% of non-retakers achieve this placement. This low admission rate to top-choice programs likely motivates the decision to retake the exam the following year. Notably, during their subsequent attempt in 2011, the placement rate into

[6] Test scores are standardized within each cohort of test-takers.

[7] The results are qualitatively the same when examining raw scores. See Table A.1 in Appendix A.

most-preferred programs for retakers rises to 39%. A similar pattern emerges within the bandwidth subsample: only 28% of future retakers are placed in their top-choice program in 2010, compared to a 49% placement rate among non-retakers. However, by their second attempt, retakers effectively close this gap, achieving a 50% placement rate in their first-preference programs.

Table 2: Scores, Placement, and Enrollment (Working and Bandwidth Samples)

	Sample			Bandwidth Sample		
	(1) All	(2) Non-retaker	(3) Retaker	(4) All	(5) Non-retaker	(6) Retaker
Average Score						
Standardized Score in 2010	0.90	0.92	0.84	1.00	1.00	0.98
	(0.71)	(0.73)	(0.61)	(0.71)	(0.73)	(0.61)
	[0.00]	[0.00]	[0.01]	[0.01]	[0.01]	[0.01]
Standardized Score in 2011			1.05			1.15
			(0.66)			(0.66)
			[0.01]			[0.01]
Admission						
Admitted in 2010	0.77	0.79	0.64	0.88	0.89	0.81
Admitted Top-Ranked Preference in 2010	0.38	0.43	0.17	0.45	0.49	0.28
Admitted in 2011			0.85			0.90
Admitted Top-Ranked Preference in 2011			0.39			0.50
Enrollment						
Enrolled in 2010	0.58	0.63	0.33	0.69	0.73	0.46
Enrolled Top-Ranked Preference in 2010	0.36	0.41	0.13	0.43	0.47	0.21
Enrolled in 2011			0.65			0.72
Enrolled Top-Ranked Preference in 2011			0.38			0.48
Total observations	46,367	37,810	8,557	19,387	16,400	2,987

Notes: Table 2 reports descriptive statistics of the scores, admission and enrollment outcomes for our working and bandwidth sample. Standard deviations are reported in parentheses and the standard error of the mean in square brackets.

The significant improvement in placement across attempts, in turn, drives better enrollment outcomes. In the working sample, retakers increase their overall enrollment rate by 32 percentage points (pp) between their first and second attempt, rising from 33% to 65%. A similar trend exists for the bandwidth sample, where enrollment increases by 26 pp (from 46% to 72%). Across both samples, the enrollment gap between non-retakers and retakers effectively closes over time; by 2011, both groups exhibit nearly identical overall enrollment rates.

4 Empirical Strategy

Our empirical strategy builds on recent work that uses RDDs around admission cutoffs to study exam retaking behavior and its consequences (Goodman et al., 2020; Bizopoulou et al., 2024; Kang et al., 2024). We essentially implement the same identification strategy relying on top-preference cutoffs and extend the analysis beyond immediate outcomes (i.e., exam performance, admissions, and enrollment) to longer-run results like graduation and labor market indicators. Focus beyond admission and enrollment is key to fully assess the sustained net effects of retaking admission exams on individual outcomes.

Setup. Our research design exploits a setting that resembles a local experiment where test-takers with nearly identical scores have different probabilities of retaking the exam due to chance. We use a RDD around the cutoff of the program ranked at the top of each student preference list. The logic is that test-takers who barely miss their top preference would be – by chance– more likely to retake the exam than other, otherwise similar, test-takers. Let s_i denote the application score and c the cutoff of the top program in the ranked ordered list. Define $Z_i = 1(s_i < c_i)$ as an indicator variable for narrowly missing admission to the top program, and let R_i indicate whether the student retakes the exam the following year. Our first-stage equation can be formalized as follows:

$$R_i = \gamma_0 + \gamma_1 Z_i + \gamma_2 f(s_i - c_i) + \gamma_3 Z_i \times f(s_i - c_i) + X_i' \gamma_4 + \eta_i \quad (1)$$

where $f(\cdot)$ is a flexible function of the (centered) score, which we interact with Z_i to allow for different slopes on each side of the cutoff. We also include a set of predetermined variables as controls in X_i such as gender, measures of household socioeconomic status, and high school characteristics. All these control variables behave smoothly near the cutoff and serve mainly to improve precision of our RD estimates.

In the case of immediate outcomes (i.e., scores, preferences, admission, and enrollment), we focus on identifying the causal effects of retaking on outcome improvements. We borrow from Kang et al. (2024) and measure the improvements using differences between final and initial realizations of the outcomes, $\Delta Y_i = Y_{i,F} - Y_{i,I}$, where $Y_{i,I}$ is the outcome observed at the time of the first attempt and $Y_{i,F}$ is the outcome observed after the subsequent attempt. For individuals who do not retake there is no new realization of the outcome in the subsequent year. Therefore, this specification of the outcome measures how individuals induced to retake gain relative to their own baseline. This distinction matters because applicants who narrowly miss the cutoff for their top-preference program perform worse on a range of initial outcomes by construction. Estimating effects on final outcomes alone would mix this baseline disadvantage with any gains generated by retaking. Using outcome differences isolates improvement.

When evaluating long-term trajectories—specifically degree completion and labor market performance—we focus on absolute outcomes. From an individual welfare perspective, the returns to retaking should be measured in absolute terms rather than relative to own baseline. This approach ensures that we capture the total human capital gain and the final market value of the degree obtained, rather than just the marginal shift for retakers.

The reduced-form equation estimating effects of narrowly missing the top-preference cutoff on outcome improvements ΔY_i is given by

$$\Delta Y_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 f(s_i - c_i) + \alpha_3 Z_i \times f(s_i - c_i) + X_i' \alpha_4 + \varepsilon_i, \quad (2)$$

The coefficient α_1 captures the intention-to-treat effect on each outcome improvement. We combine these equations in a fuzzy RD two-stage least squares framework:

$$\Delta Y_i = \kappa_0 + \kappa_1 \widehat{R}_i + \kappa_2 f(s_i - c_i) + \kappa_3 \widehat{R}_i \times f(s_i - c_i) + X_i' \kappa_4 + u_i. \tag{3}$$

Identification. Our identifying assumptions are standard in RD designs (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). Test-takers should not be able to manipulate their application score around the cutoff, and potential outcomes and covariates should vary smoothly at the threshold. Also, our instrument should affect outcomes only through retaking behavior. We follow (Goodman et al., 2020; Bizopoulou et al., 2024) to argue that, for intermediate outcomes such as exam scores and admissions, the exclusion restriction is satisfied because retaking is the only channel through which such outcomes can change. Following Bizopoulou et al. (2024), we further provide empirical support for the exclusion restriction by examining students near the cutoff (see Sub-section 5.1).

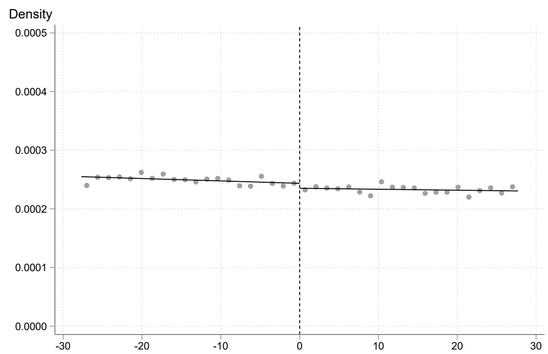
Estimation. All regressions are estimated using observations within the optimal bandwidth for the first stage, selected following Calonico et al. (2019). This optimal window is 27.5 points around the cutoff. Results are robust to alternative bandwidth choices, the inclusion of covariates, and different functional-form specifications (see Table B.3 in Appendix B). Estimation uses local linear point estimators with robust bias-corrected confidence intervals following Calonico et al. (2019, 2020).

5 Results

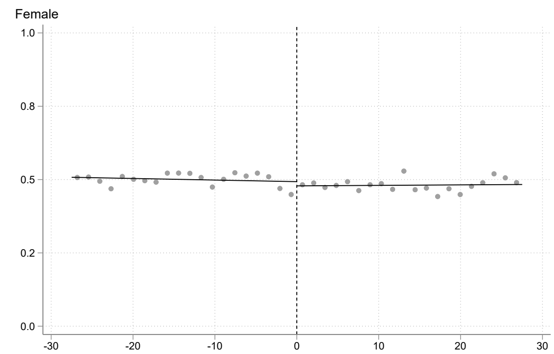
5.1 First stage

No Manipulation and Covariates Smoothness. We begin this section presenting two pieces of evidence supporting the validity of our RDD. First, we find no evidence of manipulation of our running variable. Column (1) in Table 3 and its graphical analog in Figure 1a show no evidence of a discontinuity in the density of the top preference application score at the threshold. Second, we find no differences in a host of several covariates at the cutoff. Figures 1b, 1c and 1d display a smooth behavior of gender, graduation from a public high school and mother's schooling as examples. Table 3 reports the estimates from equation (2) for these covariates, along with additional baseline characteristics such as high school location (capital region and rural status), paternal education, and household income. The corresponding graphical evidence for the remaining covariates is presented in Figure B.1. Across all specifications, the coefficients are precisely estimated zeros of small magnitude and are statistically insignificant at conventional levels.

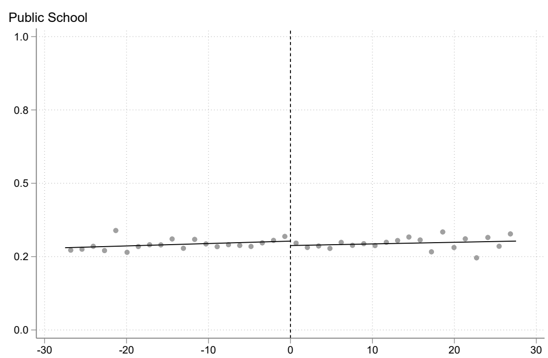
Figure 1: No Manipulation and Covariates Smoothness



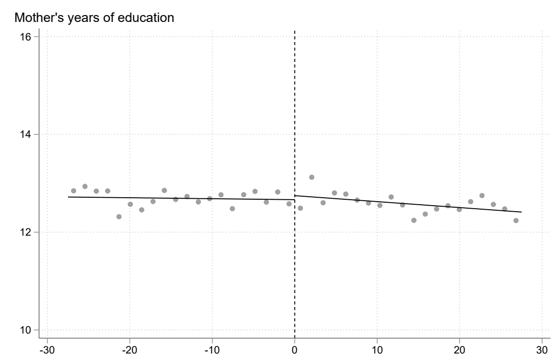
(a) Density of the Running Variable



(b) Female



(c) Public High School



(d) Mothers' Years of Education

Notes: Figures 1a, 1b, 1c and 1d are graphical analog to the corresponding estimates in Table 3 (columns 1, 2, 3 and 6, respectively). We provide the remaining graphs in Figure B.1. The sample consists of observations within an optimal bandwidth of 27.5 points around the cutoff. Each graph plots the mean of the y-axis variable within bins of the top preference application score and fits estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff.

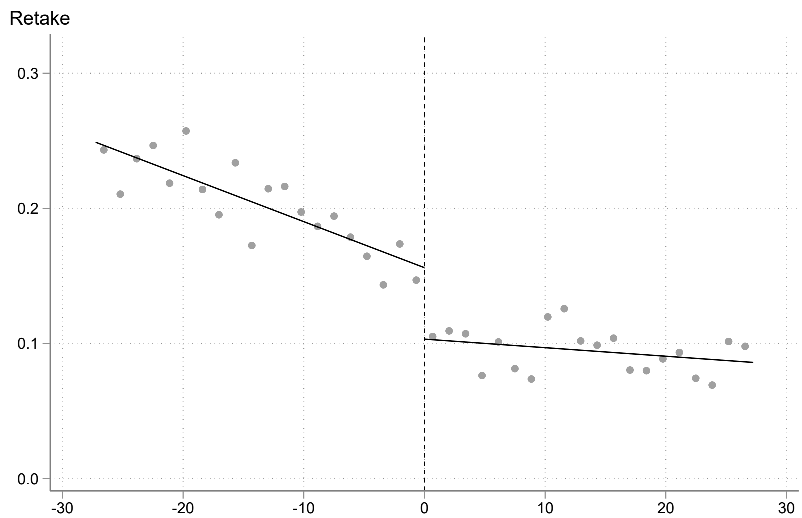
Table 3: No Manipulation and Covariates Smoothness

	(1) Density	(2) Female	(3) High School GPA	(4) Public High School	(5) Capital Region High School	(6) Rural High School	(7) Mothers' Education	(8) Fathers' Education	(9) Household Income
$\widehat{\alpha_1}$	0.000	-0.001	-1.318	0.018	-0.024	-0.008	-0.102	-0.052	-23.808
	(0.000)	(0.017)	(3.017)	(0.016)	(0.016)	(0.005)	(0.108)	(0.112)	(17.457)
Mean	0.000	0.491	619.741	0.293	0.341	0.021	12.648	12.847	589.191
Mean Above Cutoff	0.000	0.481	627.840	0.295	0.323	0.021	12.595	12.817	601.800
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548	27.548	27.548	27.548	27.548	27.548
Observations	19,387	19,387	19,387	19,387	19,387	19,387	19,387	19,387	19,387

Notes: Table 3 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 for the following dependent variables: the density of the top preference application score (column 1); indicators for female (column 2), graduating from a public high school (column 3), graduating from a high school in the capital region (column 4) and graduating from a rural high school (column 5); maternal and paternal years of schooling (columns 6 and 7), and household income, measured in thousands of Chilean pesos (column 8). The running variable is the (centered) application score to the top preference. Robust standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Retaking the admission exam near the threshold. Figure 2 illustrates the jump in the probability of retaking the exam at the admission threshold. Estimating equation (1), we find a 4.3 percentage point (pp) increase at the cutoff, representing a 44% surge over a baseline retake rate of 9.9%. This discontinuity shows that test-takers who fall just below the threshold are significantly more likely to retake the exam the following year. Our estimates are remarkably consistent with findings from analogous empirical strategies in China (a 5.1 pp increase over a 10% baseline; Kang et al., 2024) and Greece (a 7.9 pp increase over a 28% baseline; Bizopoulou et al., 2024).[8]

Figure 2: Change in the Probability of Retaking at the Threshold



Notes: Figure 2 plots the mean of retaking the exam the next year within bins of the top preference application score and fits estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff. The sample consists of observations within an optimal bandwidth of 27.5 points around the cutoff. The difference at the cutoff, estimated from the equation (1), is $\widehat{\alpha} = 0.043$ ($SE = 0.011$).

Following the logic of a fuzzy RDD, we exploit the quasi-random assignment of students to either side of the admission threshold as an instrument for retaking the exam. Because falling narrowly below the cutoff provides a sharp, exogenous shock to the incentive to retake—unrelated to a student's underlying ability or demographics—this variation allows us to identify the causal effect of retaking on subsequent outcomes for the "marginal" student. We begin by documenting the impact on immediate academic results, but our analysis goes beyond the existing literature by tracking these individuals over a 15-year horizon to identify the long-term consequences for university graduation and labor market success.

5.2 Immediate Effects

We begin by documenting the immediate impact of retaking across the primary stages of the university transition process. Table 4 reports the estimated effects on test performance, application be-

[8]In Figure C.2 we provide a visual comparison plotting the first-stage graph for the three countries side by side.

havior, admission outcomes, and subsequent enrollment, while Figure 3 visualizes the corresponding reduced-form discontinuities at the threshold. Taken together, these results provide a comprehensive view of how a second attempt reshapes a student's academic trajectory—starting with the accumulation of test-specific human capital, potentially affecting application choices, and culminating in the realization of high-stakes educational choices.

Effects on Scores. Our RD estimates in Table 4 indicate that retaking leads to a 0.27 SD increase on test scores for cutoff-induced retakers (see column 1). These score gains are well within the range of estimates reported in related studies using similar empirical designs. In the United States, Goodman et al. (2020) find that threshold-induced retaking increases SAT scores by 0.15–0.30 SD. Similarly, Kang et al. (2024) document that retaking in China increases standardized exam scores by 0.47 SD, while Bizopoulou et al. (2024) report even larger gains of approximately 0.60 SD in the Greek context. Our findings thus align with the international evidence suggesting that a second attempt yields substantial improvements in exam performance.

Effects on Admissions and Enrollment. We next examine how the significant score gains documented in the previous section translate into changes in admission and enrollment. This transition is critical: the improved exam performance effectively unlocks access to more selective tiers of higher education that were previously out of reach. By analyzing both outcomes, we can distinguish between the expansion of a student's opportunity set—as reflected in admission offers—and their realized choice, as captured by actual enrollment. This allows us to observe not only the enhanced academic supply available to retakers, but also their demand for potentially higher-quality or better-matched programs.

The reduced-form estimates in the first row in Table 4 indicate that eligibility for retaking improves admission and enrollment probabilities by 3.2 and 3.0 percentage points, respectively. Given a 4.3 percentage point discontinuity in the probability of retaking, the implied fuzzy RD estimates indicate that retaking improves placement and enrollment chances by approximately 0.735 and 0.687, respectively among cutoff-induced retakers (see columns 3 and 4). These sizable improvements indicate that retaking affects not only formal eligibility or offers, but also downstream decisions that directly shape students' realized educational trajectories.

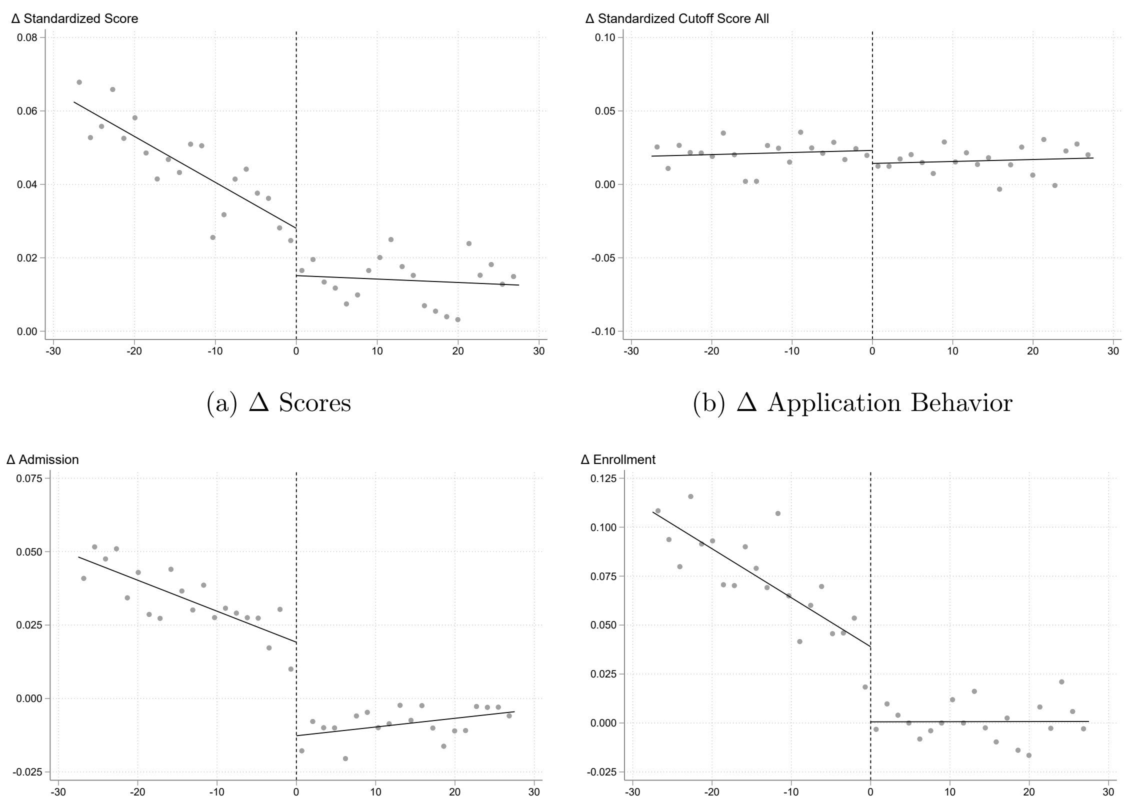
The magnitude of these effects are again consistent with recent prior evidence. Kang et al. (2024) find that retaking China's National College Entrance Examination increases eligibility to apply to universities by up to 62 percentage points for cutoff-induced retakers. Bizopoulou et al. (2024) report an increase of up to 40 percentage points in the likelihood of receiving a university offer in Greece; this surge raises the probability of admission for retakers to nearly 50–60%, bringing them close to the applicant population mean. The evidence from the United States points in the same direction: Goodman et al. (2020) report that retaking the SAT increases four-year college enrollment by 13 percentage points—nearly a 20 percent increase relative to control compliers.

Table 4: Estimated Effects on Scores, Preferences, Admission, and Enrollment Improvements

	(1) Δ Scores	(2) Δ Application Behavior	(3) Δ Admission	(4) Δ Enrollment
$\widehat{\alpha_1}$	0.012**	0.008	0.032***	0.030***
	(0.004)	(0.007)	(0.006)	(0.008)
Retaking	0.271***	0.189	0.735***	0.687***
	(0.081)	(0.142)	(0.195)	(0.221)
Mean	0.031	0.019	0.014	0.040
Mean Above Cutoff	0.014	0.016	-0.009	0.001
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548
Observations	19,387	19,387	19,387	19,387

Notes: Table 4 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. Application behavior is measured using the average cutoff score of the preferences listed in the applicants' ranked-ordered list. All estimations include the controls described in Table 3.

Figure 3: Reduced-Form Effects on Scores, Preferences, Admission, and Enrollment Improvements



(a) Δ Scores (b) Δ Application Behavior

(c) Δ Admissions (d) Δ Enrollment

Notes: All graphs in Figure 3 are graphical analog to the corresponding estimates in Table 4. The sample consists of observations within an optimal bandwidth of 27.5 points around the cutoff. Each graph plots the mean of the y-axis variable within bins of the top preference application score and fits estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff.

Effects on Application Behavior. Besides increased scores, the improvements on admissions and enrollment could also be driven by changes in application behavior. We do not find evidence that retaking affects students' application behavior. Results presented in Column 2 in Table 4 focus on the average cutoff score of the preferences listed in the ranked-ordered list as an indicator of the selectivity of the students' portfolios. We do not find significant effects on application choices. Indeed, the reduced-form estimates reveal a point estimate very close to zero. We conduct robustness checks on several other measures of selectivity as well as the size of the submitted ranked-ordered list and confirm that portfolios are not significantly changing between attempts (see Table A.2 in Appendix A). This result indicates that improved immediate outcomes following a retake are not driven by changes in students' application strategies, but rather by higher exam performance.

This finding is consistent with results from related studies. Bizopoulou et al. (2024) examine several dimensions of application behavior in Greece, including the number and quality of college–course choices, and report limited evidence of systematic changes following retaking. Similarly, Goodman et al. (2020) relies on the number of colleges to which students send their scores as a proxy for application behavior. In general, their estimates for the effect of retaking on these choices are small, negative, and statistically insignificant, suggesting that retaking does not meaningfully alter application patterns. Taken together, the evidence across settings indicates that retaking affects educational outcomes primarily by increasing exam scores and admission probabilities, rather than by inducing students to apply to different or more selective programs. This may suggest persistence towards a well-defined goal rather than a moving target over attempts.

5.3 Long-Term Effects

While the immediate gains in test scores, admissions, and enrollment are substantial, the ultimate value of a second attempt depends on whether these improvements translate into long-term human capital accumulation and labor market success. We extend our analysis beyond the typical short-run focus of the literature by tracking students over a 15-year horizon. Table 5 presents the estimated effects of retaking on the probability of degree completion and time-to-graduation, as well as proxies for labor market performance, including formal employment rates and expected earnings. Figure 4 plots the corresponding reduced-form discontinuities at the threshold.

From an individual perspective, these results are essential for evaluating the net welfare implications of retaking. While the immediate opportunity cost of a delayed career start may be offset by improved educational trajectories, the long-term consequences remain an empirical question. Specifically, while retaking unlocks more selective academic paths, it remains unclear whether these placements lead to sustainable human capital gains or, conversely, to mismatch effects if students are pushed into overly demanding programs. Measuring the impact on degree completion and labor market outcomes thus provides a vital proxy for the total private returns of the retake investment.

Effects on Graduation and Time to Graduate. Column 1 in Table 5 shows that retaking does not impact overall graduation rates. This result persists when measuring both graduation from the individual's enrollment program (panel A) as well as overall graduation from any program (panel B). Both the reduced-form and the treatment effect estimates are precisely estimated zeros; cutoff-induced retakers graduate from their enrollment program at a rate of 57.5%, which is very close to the counterfactual mean of 58.7% (see panel A). Similarly, panel B shows that overall graduation rates are 83.4%, a figure statistically indistinguishable from the counterfactual mean of 82.8%. Column 2 further shows that we find no significant impact on the time required to complete a degree, irrespective of the graduation program considered (i.e., same as enrollment or any program). Retakers take approximately 6.5 (or 6.86) years to graduate, effectively mirroring the trajectory of their peers who did not retake.

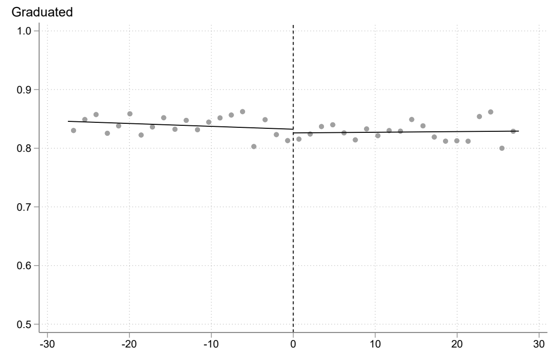
These results contribute new evidence to a literature in which long-term graduation data are often limited or inconclusive. For example, while Bizopoulou et al. (2024) study similar dynamics in Greece, their data do not allow for an analysis of degree completion. Goodman et al. (2020) examine potential effects on graduation, but they acknowledge that their follow-up period is short and their estimates are too noisy to draw definitive conclusions. Using fifteen years of administrative follow-up, we show that although retaking moves students into more selective programs, it does not increase dropout risk or generate evidence of mismatch. Retakers appear able to meet the academic standards of the programs they access due to the additional attempt they undertake.

Table 5: Estimated Effects on Graduation and Labor Market Outcomes

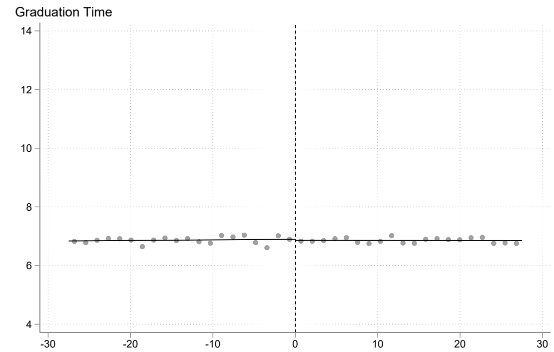
Panel A. Long-term Outcomes: Same Enrollment Program				
	(1) Graduated	(2) Graduation Time	(3) Employment Rate	(4) Expected Earnings
$\widehat{\alpha_1}$	0.001	0.002	-0.008	-0.010
	(0.016)	(0.075)	(0.005)	(0.013)
Retaking	0.014	0.031	-0.223	-0.291
	(0.337)	(1.585)	(0.161)	(0.341)
Mean	0.575	6.513	0.830	7.284
Mean Above Cutoff	0.587	6.545	0.836	7.293
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548
Observations	19,387	11,154	8,830	8,830
Panel B. Long-term Outcomes: Any Program				
	(1) Graduated	(2) Graduation Time	(3) Employment Rate	(4) Expected Earnings
$\widehat{\alpha_1}$	0.003	0.047	-0.008	-0.013
	(0.012)	(0.073)	(0.005)	(0.011)
Retaking	0.062	0.959	-0.169	-0.289
	(0.263)	(1.408)	(0.106)	(0.241)
Mean	0.834	6.862	0.824	7.287
Mean Above Cutoff	0.828	6.855	0.831	7.306
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548
Observations	19,387	16,166	12,546	12,546

Notes: Table 5 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. All estimations include the controls described in Table 3.

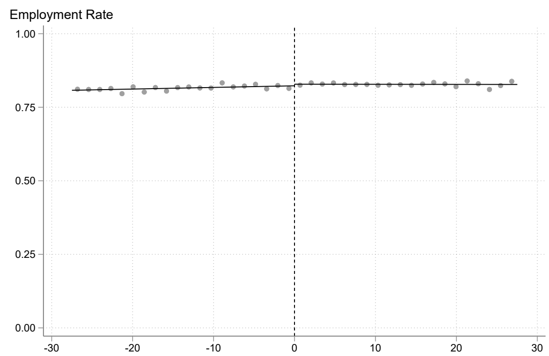
Figure 4: Reduced-Form Effects on Graduation and Labor Market Outcomes



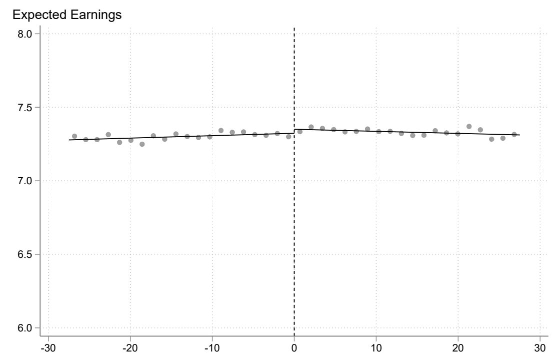
(a) Graduation from All Universities



(b) Graduation Time from All Universities



(c) Expected Employment Rate from All Universities



(d) Expected Earnings from All Universities

Notes: All graphs in Figure 4 are graphical analog to the corresponding estimates in Table 5. The sample consists of observations within an optimal bandwidth of 27.5 points around the cutoff. Each graph plots the mean of the y-axis variable within bins of the top preference application score and fits estimated lines using all the underlying data, allowing for different slopes on each side of the cutoff.

Effects on Employability and Earnings. Beyond academic persistence, we evaluate whether the improved admission and enrollment outcomes translate into better labor market prospects. Table 5 (Columns 3 and 4) reports the estimated effects on the employment rates and average earnings associated with the student's graduation program, utilizing data from the Chilean Internal Revenue Service aggregated at the university-field level. Specifically, we examine the expected employability two years post-graduation and average earnings in the fourth year following degree completion.

Despite the substantial upgrade in initial placement documented earlier, we find no statistically significant impact on these program-level proxies. The coefficients are precisely estimated zeros, suggesting that retaking does not systematically move students into programs with higher average market returns. While retakers access more selective programs, those programs do not appear to offer significantly different employment or earnings profiles than the ones they would have otherwise attended.

These findings provide a rare glimpse into the long-term economic returns of retaking—a dimension largely absent from the existing literature due to data constraints. While the aggregated measures miss within-program individual differences, our results suggest that retaking unlocks previously inaccessible academic paths but these gains do not translate into a detectable shift in expected market value based on program-level benchmarks.

Heterogeneity. We examine heterogeneous effects across several observable characteristics, including gender, parental education, household income, socioeconomic status, prior test scores, high school GPA, and geographic location. The results are reported in Table B.4 and Table B.5. In most cases, we lack sufficient statistical power to detect meaningful differences in retaking effects across subgroups.[9] This is a common challenge in regression discontinuity designs, where identification relies on observations near the administrative cutoff and subgroup partitioning further reduces the effective sample size (Calonico et al., 2020).

That said, a few patterns emerge that are worth mentioning with caution. First, students from more advantaged socioeconomic backgrounds—such as those with higher parental education, higher household income, or higher SES indices—tend to exhibit relatively larger score gains over attempts. However, these academic gains do not translate into disproportionately larger admission or enrollment gaps. While the intensive margin (score improvement) is stronger for high-SES students, the extensive margin (actual university entry) appears more uniform across socioeconomic strata. This result suggests that the admission thresholds may act as a leveling mechanism, where even the smaller score gains achieved by lower-SES students are sufficient to cross critical participation cutoffs.

Second, we find that students across the high school GPA distribution experience similar score gains, but those with lower GPAs realize relatively larger improvements in admission outcomes. For these students, the retake appears to serve as a compensating mechanism, allowing a higher test score to offset a weaker high school record within the centralized ranking system.

Third, the first-stage estimates reveal that the discontinuity in the probability of retaking is significantly more pronounced among students from more advantaged socioeconomic backgrounds. This suggests that students from advantaged backgrounds are more responsive to just missing the cutoff of their preferred program. For this group, the administrative threshold acts as a powerful trigger for a second attempt, whereas lower-SES students are less likely to transition into a retake regardless of their proximity to the cutoff. This differential responsiveness suggests that the ability to utilize a second chance is socially stratified. This retaking gap may explain why we do not observe the same pro-poor benefits found in Goodman et al. (2020) and Frisancho et al. (2016): if the most vulnerable students do not comply with the encouragement to retake, they cannot reap the associated long-term rewards.

[9] In fact, note that the RD estimates for male, low mother's education, low household income, middle SES, and low initial PSU score should be interpreted with caution, as the first-stage coefficient fails to reach statistical significance, suggesting a lack of identifying variation for this partition of the sample.

6 Discussion

This study provides some of the first empirical evidence on the long-term causal effects of exam retaking, leveraging administrative records to track students for fifteen years following their initial attempt. Our analysis reveals a striking decoupling between immediate performance gains and later-life economic outcomes. While retaking yields significant short-run benefits—specifically, cutoff-induced retakers achieve higher scores and see a marked increase in the probability of university admission and enrollment—these initial successes do not translate into improved long-term economic returns.

Our results show that retaking does not impact graduation rates, time to graduate, employability, or expected earnings. These null findings provide evidence against a mismatch interpretation in this context; despite their weaker initial performance, retakers do not face an elevated risk of dropout or prolonged completion times once admitted. However, the catch up they are able to transit leveraging on retaking –as well as an additional year of preparation– does not lead to a corresponding labor market premium. Ultimately, the degrees obtained by retakers yield a similar market value to those earned by comparable non-retakers that did not delayed their enrollment.

A back-of-the-envelope comparison highlights the economic magnitude of this null result. Our fuzzy RD estimates imply that retaking increases tertiary enrollment by 68 percentage points among cutoff-induced retakers. Using Chile's centralized admissions as a benchmark, where crossing an admission cutoff increases earnings by 4.5% to 9.1% (Hastings et al., 2013), a 0.68 increase in enrollment would imply expected earnings gains of approximately 3–6%. Instead, our estimated earnings effects are near zero. While these calculations are illustrative rather than a claim of external validity, they demonstrate that the absence of long-run returns in our setting is economically meaningful.

The observed score gain of 0.27 standard deviations is significant in the short run and comparable to gains documented in other high-stakes testing contexts. Yet, the lack of long-term returns suggests that these improvements likely reflect enhanced test-taking proficiency rather than a substantial shift in underlying skills. Even when large in standardized units, score gains alone may not translate into sustained human capital accumulation. Consequently, retaking appears to primarily reshuffle access within the system—expanding entry into more selective programs—without increasing degree completion or the market value of the credentials obtained.

Furthermore, our heterogeneity analysis suggests that the reshuffling that is induced by retaking is unlikely to erode existing socioeconomic gaps in university access. While higher-SES students achieve larger raw score gains, these do not translate into disproportionately higher admission rates compared to their lower-SES counterparts. However, the primary driver of inequality in this setting appears to be the socially stratified second-chance mechanism: advantaged students are significantly more likely to utilize the retake option following a marginal failure. Thus, even if the returns to a second attempt are uniform across groups, the disproportionate take-up among the more affluent students would generate systemic benefits of retaking remain concentrated at the top of the distribution.

Our findings also speak to broader equilibrium considerations within a centralized admission system with fixed program capacities. In such an environment, improved access for retakers may reflect a reallocation of seats across applicants rather than an expansion of total capacity. This aligns with structural evidence from other centralized contexts suggesting that retaking can impose negative externalities—crowding the market and artificially inflating admission cutoffs without improving aggregate welfare (Krishna et al., 2018). Within this framework, retaking may effectively reshuffle the queue for higher education without generating additional aggregate human capital. While our estimates apply specifically to a local margin—students whose retaking decision is sensitive to narrowly missing a cutoff—this margin remains highly policy-relevant, as it captures the population most affected by shifts in admission thresholds and institutional retaking rules.

From a policy perspective, these findings offer a nuanced message. Retaking may enhance perceived fairness and expand opportunities for students who narrowly miss admission thresholds. Even in the absence of measurable labor market gains, retaking may increase private utility through access to preferred fields of study or institutions—dimensions not captured by earnings proxies. However, the fact that higher-SES students are significantly more likely to utilize this second chance suggests that retaking may inadvertently reinforce socioeconomic disparities in university access. Furthermore, its contribution to aggregate human capital formation appears limited. Policies that focus exclusively on retaking opportunities may therefore yield weaker long-run effects than those targeting earlier skill formation or post-enrollment academic support.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263.

Abadie, A., Angrist, J., and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica*, 70(1):91–117.

Abdulkadiroğlu, A., Angrist, J., and Pathak, P. (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, 82(1):137–196.

Arcidiacono, P., Aucejo, E. M., and Spenner, K. (2012). What happens after enrollment? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics*, 1:5.

Arteaga, F., Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2022). Smart matching platforms and heterogeneous beliefs in centralized school choice*. *The Quarterly Journal of Economics*, 137:1791–1848.

Astorne-Figari, C. and Speer, J. D. (2019). Are changes of major major changes? the roles of grades, gender, and preferences in college major switching. *Economics of Education Review*, 70:75–93.

Barrios Fernandez, A., Nielson, C., and Zimmerman, S. (2024). Elite universities and the intergenerational transmission of human and social capital. Technical report.

Bizopoulou, A., Megalokonomou, R., and Simion, S. (2024). Do second chances pay off? evidence from a natural experiment with low-achieving students. *Journal of Public Economics*.

Bratti, M., Granato, S., and Havari, E. (2024). Another chance: Number of exam retakes and university students' outcomes. Technical report, IZA Discussion Paper No. 17400.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). Rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression discontinuity designs using covariates. *The Review of Economics and Statistics*, 101(3):442–451.

Caner, A. and Okten, C. (2010). Risk and career choice: Evidence from turkey. *Economics of Education Review*, 29:1060–1075.

Caner, A. and Okten, C. (2013). Higher education in turkey: Subsidizing the rich or the poor? *Economics of Education Review*, 35:75–92.

Cattaneo, M., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, 18:234–261.

Celhay, P. and Gallegos, S. (2022). Early skill effects on parental beliefs, investments and children long-run outcomes1. *Journal of Human Resources*.

Chetty, R., Deming, D. J., and Friedman, J. N. (2023). Diversifying society's leaders? the determinants and causal effects of admission to highly selective private colleges. Working Paper 31492, National Bureau of Economic Research.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8:36–65.

Frisancho, V., Krishna, K., Lychagin, S., and Yavas, C. (2016). Better luck next time: Learning through retaking. *Journal of Economic Behavior & Organization*, 125:120–135.

Fu, Q. (2006). A theory of affirmative action in college admissions. *Economic Inquiry*, 44:420–428.

Goodman, J., Gurantz, O., and Smith, J. (2020). Take two! sat retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, 12:115–158.

Hastings, J. S., Neilson, C. A., Ramirez, A., and Zimmerman, S. D. (2016). (un)informed college and major choice: Evidence from linked survey and administrative data. *Economics of Education Review*, 51:136–151. Access to Higher Education.

Hastings, J. S., Neilson, C. A., and Zimmerman, S. D. (2013). Are some degrees worth more than others? evidence from college admission cutoffs in chile. Working Paper 19241, National Bureau of Economic Research.

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

Kang, L., Lei, Z., Song, Y., and Zhang, P. (2024). Gender differences in reactions to failure in high-stakes competition: Evidence from the national college entrance exam retakes. *Journal of Political Economy Microeconomics*, 2:355–397.

Kapor, A., Karnani, M., and Neilson, C. (2024). Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms. *Journal of Political Economy*, 132(7):2346–2395.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection*. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Krishna, K., Lychagin, S., and Frisancho, V. (2018). Retaking in high stakes exams: Is less more? *International Economic Review*, 59:449–477.

Landaud, F. and Maurin, E. (2020). Aim High and Persevere! Competitive Pressure and Access Gaps in Top Science Graduate Programs. working paper or preprint.

Larrocau, T. and Rios, I. (2022). Dynamic college admissions and the determinants of students' college retention. Technical report, Cowles Foundation for Research in Economics, Yale University.

Larroucau, T., Rios, I., Fabre, A., and Neilson, C. (2024). College application mistakes and the design of information policies at scale.

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

Machin, S., McNally, S., and Ruiz-Valenzuela, J. (2020a). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics*, 190:104224.

Machin, S., McNally, S., and Ruiz-Valenzuela, J. (2020b). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics*, 190:104224.

Papay, J. P., Murnane, R. J., and Willett, J. B. (2016). The impact of test score labels on human-capital investment decisions. *Journal of Human Resources*, 51(2):357–388.

Rios, I., Larroucau, T., Parra, G., and Cominetti, R. (2021). Improving the chilean college admissions system. *Operations Research*, 69(4):1186–1205.

Saygin, P. O. (2016). Gender differences in preferences for taking risk in college applications. *Economics of Education Review*, 52:120–133.

Simion, S., Bizopoulou, A., and Megalokonomou, R. (2023). Do second chances pay off? evidence from a natural experiment with low-achieving students. *Evidence from a Natural Experiment with Low-Achieving Students*.

Smith, J., Hurwitz, M., and Avery, C. (2017). Giving college credit where it is due: Advanced placement exam scores and college outcomes. *Journal of Labor Economics*, 35(1):67–147.

Tansel, A. and Bircan, F. (2006). Demand for education in turkey: A tobit analysis of private tutoring expenditures. *Economics of Education Review*, 25(3):303–313. Special Issue: In Honor of W. Pierce Liles.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754.

Lasting Effects of Retaking College Admission Exams

Online Appendix

Veronica Frisancho Sebastian Gallegos Constanza González

March 10, 2026

Contents

Appendix A: Additional Results

Table A.1: Descriptive Statistics of Test Scores

	Sample			Bandwidth Sample		
	(1) All	(2) Non-retaker	(3) Retaker	(4) All	(5) Non-retaker	(6) Retaker
Average Score						
2010	594	596	588	604	604	602
	(73.78)	(75.88)	(63.28)	(73.37)	(75.03)	(63.47)
	[0.34]	[0.39]	[0.68]	[0.53]	[0.59]	[1.16]
2011			609			619
			(67.09)			(67.71)
			[0.73]			[1.24]
Language Score						
2010	591	592	586	600	601	600
	(79.11)	(80.76)	(71.17)	(78.10)	(79.17)	(72.00)
	[0.37]	[0.42]	[0.77]	[0.56]	[0.62]	[1.32]
2011			606			614
			(75.66)			(76.93)
			[0.82]			[1.41]
Math Score						
2010	597	599	589	607	608	604
	(84.46)	(87.08)	(71.24)	(85.75)	(87.86)	(72.96)
	[0.39]	[0.45]	[0.77]	[0.62]	[0.69]	[1.33]
2011			612			624
			(77.93)			(79.76)
			[0.84]			[1.46]
Elective Score						
2010	593	594	588	605	605	603
	(85.75)	(87.59)	(76.82)	(85.31)	(86.57)	(78.01)
	[0.40]	[0.45]	[0.83]	[0.61]	[0.68]	[1.43]
2011			617			627
			(80.77)			(82.88)
			[0.87]			[1.52]
GPA Score	610	612	603	620	621	614
	(94.90)	(96.07)	(89.20)	(90.35)	(91.00)	(86.50)
	[0.44]	[0.49]	[0.96]	[0.65]	[0.71]	[1.58]
Observations	46,367	37,810	8,557	19,387	16,400	2,987

Notes: Table A.1 reports descriptive statistics for the language and mathematics tests, their average, and the elective test (history or science), as well as High School GPA. Standard deviations are reported in parentheses and standard errors of the mean are reported in brackets.

Table A.2: Estimated Effects on Preferences

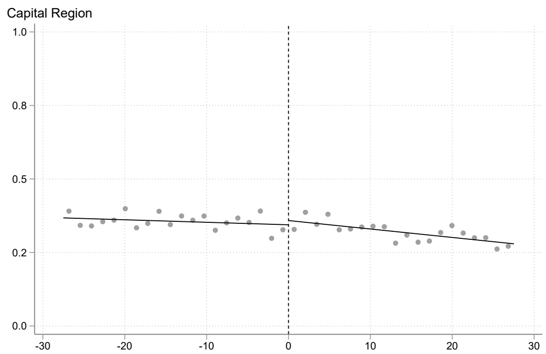
	(1) Δ Application Behavior (Top-Ranked)	(2) Δ Application Behavior (Top 3)	(3) Δ IQR	(4) Δ Selective Portafolio	(5) Δ Portfolio Size
$\widehat{\alpha_1}$	0.014	0.009	0.002	0.003	-0.009
	(0.009)	(0.007)	(0.007)	(0.004)	(0.031)
Retaking	0.322*	0.201	0.047	0.075	-0.198
	(0.174)	(0.152)	(0.146)	(0.078)	(0.649)
Mean	0.029	0.021	0.000	0.006	-0.033
Mean Above Cutoff	0.032	0.020	0.000	0.006	-0.036
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548	27.548
Observations	19,387	19,387	19,387	19,387	19,387

Notes: Table A.2 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. All estimations include the controls described in Table 3.

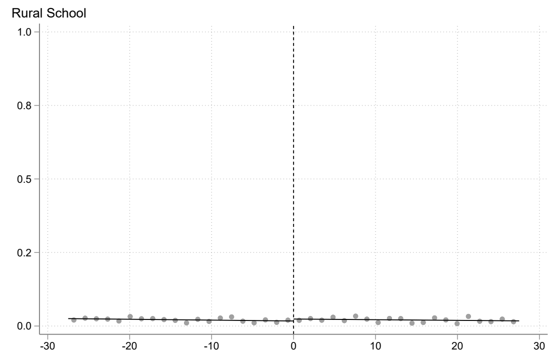
Appendix B: Robustness Checks

B.1 Covariates Smoothness

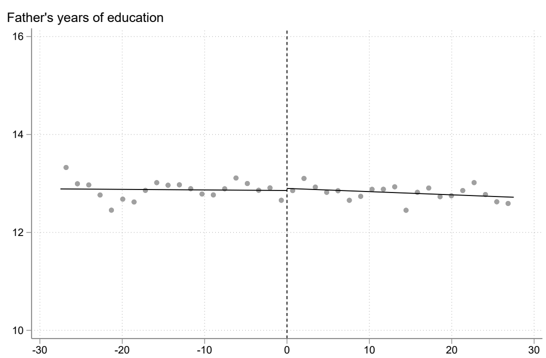
Figure B.1: No Manipulation and Covariates Smoothness



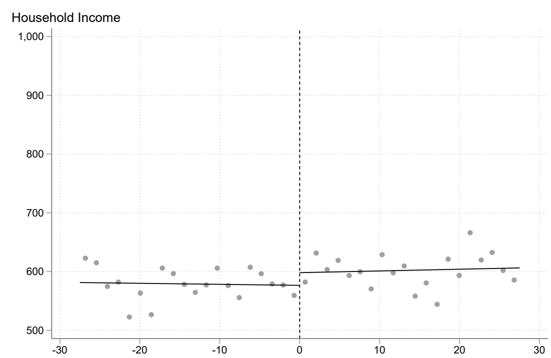
(a) Capital Region High School



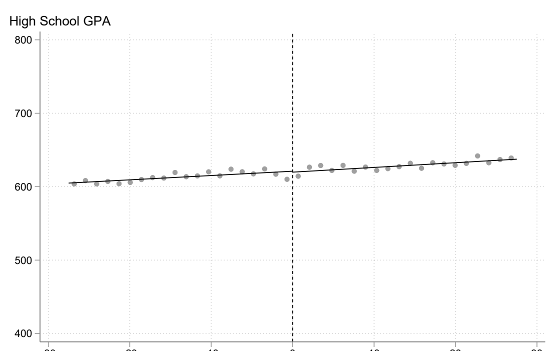
(b) Rural High School



(c) Fathers' Education



(d) Household Income



(e) High School GPA

Notes: The figures in Figure B.1 are graphical analog to estimates in Table 3. Each graph plots the probability of the y-axis variable within the difference between the application score for their first preference and the minimum to be accepted into the program in 2010. The dependent variables are described in Table 3.

B.2 Alternative specifications

Table B.3: Estimated Effects on Scores, Preferences, Admission, and Enrollment Improvements

	(1) Δ Scores	(2) Δ Application Behavior	(3) Δ Admission	(4) Δ Enrollment
Panel (a): No Covariates				
$\widehat{\alpha_1}$	0.012**	0.008	0.032***	0.030***
	(0.004)	(0.007)	(0.006)	(0.008)
Retaking	0.264***	0.182	0.734***	0.683***
	(0.081)	(0.141)	(0.193)	(0.220)
Mean	0.031	0.019	0.014	0.040
Mean Above Cutoff	0.014	0.016	-0.009	0.001
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548
Observations	19,387	19,387	19,387	19,387
Panel (b): Optimal Bandwidth				
$\widehat{\alpha_1}$	0.012***	0.009*	0.032***	0.028***
	(0.004)	(0.006)	(0.005)	(0.009)
Retaking	0.263***	0.197	0.739***	0.657***
	(0.079)	(0.138)	(0.213)	(0.205)
Mean	0.031	0.019	0.014	0.040
Mean Above Cutoff	0.014	0.016	-0.009	0.001
BW Loc. Poly. (h)	30.983	44.931	37.248	23.388
Observations	21,396	28,308	24,636	16,814
Panel (c): Local Quadratic				
$\widehat{\alpha_1}$	0.011*	0.009	0.031***	0.023**
	(0.006)	(0.009)	(0.007)	(0.011)
Retaking	0.268**	0.223	0.743**	0.555*
	(0.122)	(0.221)	(0.323)	(0.314)
Mean	0.031	0.019	0.014	0.040
Mean Above Cutoff	0.014	0.016	-0.009	0.001
BW Loc. Poly. (h)	27.548	27.548	27.548	27.548
Observations	19,387	19,387	19,387	19,387

Notes: Table B.3 replicates the results in Table 4 under three alternative specifications: without covariates (Panel a), using the optimal bandwidth of each outcome (Panel b), and using a different functional-form specification (Panel c). Table B.3 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. All estimations include the controls described in Table 3.

B.3 Heterogeneous Effects

Table B.4: Estimated Effects of Top-Preference Cutoff by Groups

	(1) Retaker	(2) Δ Scores	(3) Δ Application Behavior	(4) Δ Admission	(5) Δ Enrollment
Panel A. Gender					
Female	0.063***	0.014**	0.012	0.023***	0.023*
	(0.016)	(0.006)	(0.009)	(0.007)	(0.012)
Observations	9,527	9,527	9,527	9,527	9,527
Male	0.026	0.010	0.004	0.041***	0.037***
	(0.016)	(0.006)	(0.010)	(0.008)	(0.012)
Observations	9,860	9,860	9,860	9,860	9,860
H_0: $\alpha_{female} = \alpha_{male}$	0.026	0.349	0.599	0.142	0.820
Panel B. Father's Education					
Low	0.029**	0.005	0.001	0.034***	0.019*
	(0.014)	(0.005)	(0.009)	(0.007)	(0.011)
Observations	11,835	11,835	11,835	11,835	11,835
High	0.067***	0.023***	0.020*	0.030***	0.047***
	(0.019)	(0.007)	(0.011)	(0.009)	(0.014)
Observations	7,552	7,552	7,552	7,552	7,552
H_0: $\alpha_{low} = \alpha_{high}$	0.144	0.003	0.082	0.221	0.081
Panel C. Mother's Education					
Low	0.024	0.010	0.010	0.038***	0.024**
	(0.016)	(0.006)	(0.010)	(0.008)	(0.012)
Observations	10,342	10,342	10,342	10,342	10,342
High	0.064***	0.014**	0.007	0.026***	0.037***
	(0.016)	(0.006)	(0.009)	(0.007)	(0.012)
Observations	9,045	9,045	9,045	9,045	9,045
H_0: $\alpha_{low} = \alpha_{high}$	0.139	0.265	0.621	0.568	0.255
Panel D. Household Income					
Low	0.022	0.011*	0.006	0.032***	0.024**
	(0.015)	(0.005)	(0.009)	(0.007)	(0.011)
Observations	10,738	10,738	10,738	10,738	10,738
High	0.070***	0.014**	0.012	0.033***	0.039***
	(0.017)	(0.007)	(0.010)	(0.009)	(0.013)
Observations	8,649	8,649	8,649	8,649	8,649
H_0: $\alpha_{low} = \alpha_{high}$	0.006	0.153	0.434	0.630	0.100

Notes: Table B.4 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. Panels B, C, and D split the sample according to whether the relevant characteristic lies below or above the sample median, defining the Low and High groups, respectively. The last row of each panel displays the corresponding p-value testing differences across subgroups. All estimations include the controls described in Table 3.

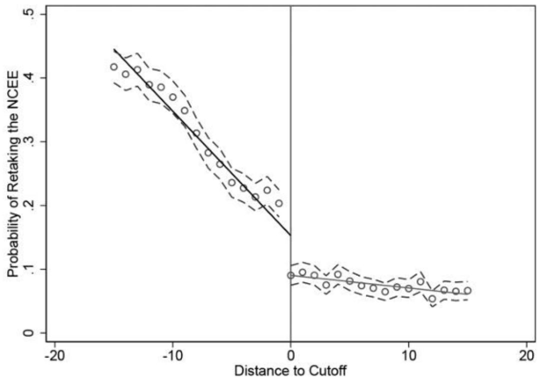
Table B.5: Estimated Effects of Top-Preference Cutoff by Groups (Continued)

	(1) Retaker	(2) Δ Scores	(3) Δ Application Behavior	(4) Δ Admission	(5) Δ Enrollment
Panel E.: SES					
Low	0.035*	0.012	0.013	0.031***	0.023
	(0.020)	(0.007)	(0.012)	(0.010)	(0.014)
Observations	6,225	6,225	6,225	6,225	6,225
Middle	0.019	0.003	0.002	0.032***	0.023
	(0.021)	(0.008)	(0.013)	(0.010)	(0.016)
Observations	6,323	6,323	6,323	6,323	6,323
High	0.073***	0.021***	0.010	0.034***	0.043***
	(0.019)	(0.007)	(0.011)	(0.009)	(0.014)
Observations	6,839	6,839	6,839	6,839	6,839
H_0: $\alpha_{middle} = \alpha_{low}$	0.864	0.780	0.209	0.234	0.635
H_0: $\alpha_{high} = \alpha_{low}$	0.196	0.161	0.794	0.521	0.146
H_0: $\alpha_{high} = \alpha_{middle}$	0.156	0.113	0.108	0.508	0.372
Panel F. Initial PSU Score					
Low	0.026	0.005	-0.002	0.044***	0.021*
	(0.016)	(0.006)	(0.011)	(0.009)	(0.012)
Observations	9,696	9,696	9,696	9,696	9,696
High	0.059***	0.019***	0.018**	0.022***	0.039***
	(0.016)	(0.006)	(0.009)	(0.007)	(0.012)
Observations	9,691	9,691	9,691	9,691	9,691
H_0: $\alpha_{low} = \alpha_{high}$	0.183	0.038	0.167	0.194	0.036
Panel G. GPA Score					
Low	0.035**	0.008	-0.001	0.060***	0.040***
	(0.016)	(0.006)	(0.011)	(0.009)	(0.012)
Observations	9,813	9,813	9,813	9,813	9,813
High	0.052***	0.016**	0.017*	0.004	0.019
	(0.016)	(0.006)	(0.009)	(0.007)	(0.012)
Observations	9,574	9,574	9,574	9,574	9,574
H_0: $\alpha_{low} = \alpha_{high}$	0.459	0.255	0.292	0.000	0.676
Panel H. Capital Region					
No	0.047***	0.013**	0.014	0.032***	0.030***
	(0.014)	(0.005)	(0.009)	(0.007)	(0.010)
Observations	12,778	12,778	12,778	12,778	12,778
Yes	0.037*	0.009	-0.002	0.032***	0.029*
	(0.019)	(0.008)	(0.011)	(0.010)	(0.014)
Observations	6,609	6,609	6,609	6,609	6,609
H_0: $\alpha_{no} = \alpha_{yes}$	0.884	0.463	0.347	0.861	0.878

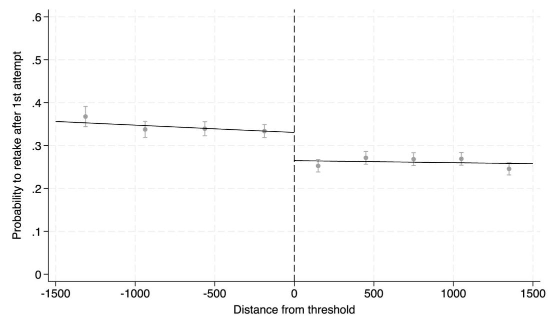
Notes: Table B.5 reports the results from estimating equations 2 and 3 within an optimal bandwidth of 27.5 points around the cutoff for each outcome in the columns. Panel E splits the sample into three roughly equal-sized groups based on the family socioeconomic status index: Low, Middle, and High. The index is constructed using parents' education, household income, and type of high school attended (public, private, or voucher). Panels F and G split the sample according to whether the relevant characteristic lies below or above the sample median, defining the Low and High groups, respectively. Panel H splits the sample by whether students graduated from a high school located in the Capital Region (Yes) or outside (No). The last row of each panel displays the corresponding p-value testing differences across subgroups. All estimations include the controls described in Table 3.

Appendix C: Comparative First-Stage Evidence from the Literature

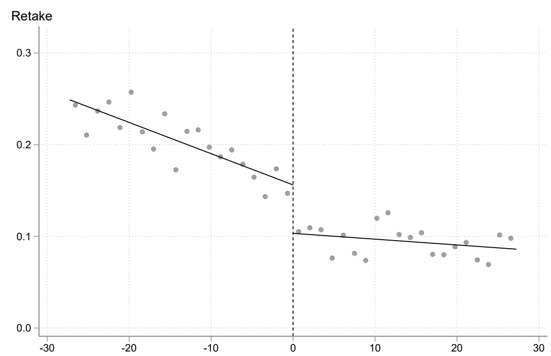
Figure C.2: First-stage evidence across related studies



(a) China (Kang et al., 2024)



(b) Greece (Bizopoulou et al., 2024)



(c) Chile (This paper)

Notes: Figure C.2 plots the probability of retaking the exam against distance to the cutoff in the running variable. Panel (a) reproduces results from Kang et al. (2024) for China; panel (b) from Bizopoulou et al. (2024) for Greece; panel (c) reports the corresponding first stage for Chile (this paper). Lines show local polynomial fits estimated separately on each side of the cutoff. The vertical dashed line indicates the cutoff.