Targeting, Selection, and the Impact of Grant Aid on Student Outcomes

Michael Galperin*
University of Chicago

October 30, 2023

Please click here for the latest draft.

Abstract

For some students, grant aid awards increase college attendance. However, existing methods are unable to separately estimate the impact of aid on these marginal students from its impacts on inframarginal students who would attend college even without additional aid. In this paper, I develop an econometric framework for separately identifying the effect of grant aid on these two groups. I use administrative data on the universe of students in Texas to apply the framework to two grant programs: the TEXAS Grant (the state of Texas's flagship need-based financial aid policy), and a cutoff in federal grant aid eligibility rules that automatically lowers students' Expected Family Contribution to zero. For middle-income students at the margin of receiving a TEXAS Grant, the grant is effectively a transfer, with no detectable impact on enrollment or postsecondary outcomes but a significant negative effect on loans and earnings during college. In contrast, for lower-income students at the margin of reciving an Auto-Zero EFC, I replicate previous findings that show significant increases in degree attainment, future earnings, and community college enrollment. Targeting of grant aid at the margin shapes these results: the TEXAS Grant has null effects on student outcomes in part because schools give grants to students with high ex-ante graduation probabilities. Finally, I show that students brought into community college attendance by receiving an Automatic Zero EFC are negatively selected. I use this fact to tighten nonparametric bounds on the grant's treatment effects, overcoming sample selection bias due to data being selected on enrollment. This method reveals positive treatment effects where standard no-assumptions bounds are too wide to be informative.

*Job Market Paper. I thank my advisors Michael Greenstone, Magne Mogstad, and Jack Mountjoy for their advice and encouragement. Thank you to Claire Bergey, Marianne Bertrand, Manasi Deshpande, Michael Dinerstein, Juanna Joensen, Thibaut Lamadon, Ruchi Mahadeshwar, Anya Marchenko, Lucy Msall, Suresh Naidu, Derek Neal, Vishan Nigam, Johanna Rayl, Evan Rose, Jordan Rosenthal-Kay, Francesco Ruggieri, Lillian Rusk, Karthik Srinivasan, and Alex Torgovitsky for helpful comments. I thank Celeste Alexander, Wintana Hansen, and Andres Rodriguez of the UT-Austin Education Research Center for their vital help and support in understanding and accessing the Texas ERC Data. I gratefully acknowledge financial support from the Becker Friedman Institute for Research in Economics and the Energy Policy Institute at the University of Chicago. The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas.

1 Introduction

Grant-based financial aid is the dominant form of financial assistance to college students in the United States. Annual spending on grants by federal and state governments and college institutions exceeds $240 billion (Ma and Pender 2022). Yet since the seminal investigation of Fuller et al. (1982), economists have questioned whether grant aid causes students to enroll in college or primarily acts as a transfer to inframarginal students who would attend college even without additional financial assistance.

Because students are heterogeneous, the answer to this question depends on how grants are targeted. For some students, grant aid *causes* enrollment in college, and therefore has potentially large benefits relative to these students' no-college counterfactual (Autor 2014). Yet even for inframarginal students who would attend college without additional aid, increased financial support may improve student outcomes by loosening budget constraints during college (Denning et al. 2019; Scott-Clayton and Zafar 2019). This dynamic has important parallels to many public subsidies: when take-up responses in the population are heterogeneous, targeting is centrally important to shaping welfare impacts (Finkelstein and Notowidigdo 2019; Ida et al. 2023; Ito et al. 2021). For which students are the benefits of grant aid the largest, and how does the targeting of grant programs affect their overall impact on student outcomes?

This paper develops a new econometric framework and uses detailed administrative data on the universe of college students in Texas to investigate how the targeting of marginal grant aid determines its effects on student outcomes. I leverage exogenous variation in aid stemming from several discontinuities in grant award rules to study how the impacts of additional grant funding depend on students' family income and pre-college preparedness. To estimate these heterogeneous effects, I develop a bounding approach to overcome sample selection bias stemming from the grant's impact on enrollments.[1] I find that the impacts of grant aid on student outcomes are higher for students with lower incomes and lower levels of pre-college preparedness, as measured by tenth grade test scores. However, an analysis of a grant program that gives schools discretion in awarding grants reveals that schools prefer to give grants to *ex-ante* higher-achieving students, dampening their average impact.

I apply the framework to two grant programs which differ significantly in their income thresholds and generosity, allowing me to study how the targeting of aid influences its effects on student outcomes. The first is the TEXAS Grant, the state of Texas's flagship need-based grant aid program. The program covers the entirety of grant recipients' tuition and fees, making it unusually generous

[1] This sample selection bias stems from a missing data problem: I observe all enrolled students, not all applicants, meaning that comparisons of marginal grant recipients and nonrecipients may conflate the causal effects of grants with compositional changes caused by their enrollment impacts.

compared to other state and federal sources of need-based aid.[2] I benchmark the TEXAS Grant's impacts against a second source of variation stemming from the federal government's formula for determining a student's Expected Family Contribution (EFC), initially studied by Denning et al. (2019). Students at the margin of receiving an "Automatic Zero EFC" have substantially lower incomes than students near the margin of receiving a TEXAS Grant, allowing me to compare how the impacts of grant aid on student outcomes depend on the income of aid recipients.[3]

Comparing these grant programs yields three main conclusions. First, for higher-income students at the margin of receiving a TEXAS Grant, I find that additional grant aid is essentially a transfer. TEXAS Grants significantly impact student finances during college, with large increases in grant aid partially offset by reductions in loans and labor market earnings among grant recipients.[4] Despite these large increases, I find economically small and statistically insignificant impacts of the program on recipients' graduation probability, persistence in college, future earnings, or course completion. The results are robust to a number of specification checks and hold at a variety of time horizons for measuring impacts on student achievement, and indicate that for middle-income students at the margin of receiving a TEXAS Grant, additional aid does not improve academic or post-college labor market outcomes.

Second, I show that incentives created by the TEXAS Grant's design are partially responsible for these null effects. Schools have discretion in deciding who among financially eligible students receives TEXAS Grants. Because schools must supplement governmental funds with institutional aid to fully cover TEXAS Grant recipients' tuition and fees, they face an incentive to award TEXAS Grants to students with high baseline grant aid awards. Moreover, because state regulators track raw graduation rates among grant recipients rather than estimates of the grant's causal impacts,[5] schools may face incentives to target the grant to *ex-ante* high-achieving students. Analyzing complier characteristics at the grant threshold (Abadie 2002), I find evidence of both phenomena. Schools award grants at the cutoff to students with substantially higher preexisting (i.e., non-TEXAS Grant) aid awards, and these grant compliers have graduation rates 14 percentage points higher than "never-taker" students who do not receive grants despite being financially eligible.

[2] An exception is the Cal Grant (Bettinger et al. 2019), which provides similar levels of renewable support on the basis of need to California public college students. Large aid increases have also been studied in experimental evaluations of privately-funded aid programs (Anderson et al. 2020; Angrist et al. 2022) and state merit aid programs (Cohodes and Goodman 2014).

[3] Importantly, the running variables for both of these programs are difficult to manipulate. The EFC, the running variable for the TEXAS Grant, is calculated by the federal government from information reported by students on the Free Application for Federal Student Aid (FAFSA). The mapping from FAFSA information to aid awards is complex, and much of the information reported on the form is easily verifiable from federal tax records, creating a deterrent for misreporting.

[4] Similar trade-offs between grants and loans have been documented in other settings. Marx and Turner (2018) find that among City University of New York students, an additional dollar of Pell Grant aid causes a $1.80 reduction in student loans. In contrast, I find that loans fall by about half of the increase in grants.

[5] see https://reportcenter.highered.texas.gov/meeting/board-supporting-documents/vii-d-texas-grant-program-report/, Appendix I.

Because treatment effects for compliers are statistically indistinguishable from zero, this difference operates entirely through selection.[6]

Third, I replicate previous findings (Denning et al. 2019) to show that for students at the margin of receiving an Automatic Zero EFC, nominally smaller increases in grant aid have significant positive impacts on persistence, degree attainment, and later-life earnings for students enrolled in four-year colleges. These impacts emerge despite the fact that the increase grant aid at the Auto-Zero threshold is less than one-fifth of the increase in aid for TEXAS Grant compliers. Gains for community college students are positive but more muted, suggesting that aid for lower-income students at the margin of receiving an Automatic Zero EFC is better targeted at students entering four-year institutions.

In estimating these results, I confront an important estimation challenge caused by sample selection, described in Section 3. While I observe outcomes for all Texas high school graduates, I only observe the running variable for my regression discontinuity (RD) analyses in the select sample of students who enroll in college. Because some students enroll in college as a result of receiving aid, standard IV methods applied across the policy threshold in enrollment-selected data will be biased whenever grant aid has enrollment effects. Empirically, I find that this concern is significant, showing that one of my sources of variation in grant aid — the Auto-Zero EFC threshold in the sample of community college students — has significant enrollment impacts, invalidating standard IV approaches to estimating the impacts of grant aid for this population.

To solve this problem, I develop a nonparametric approach that separately bounds treatment effects for two groups of students: "extensive-margin" compliers induced by grant aid to enroll in college, and "intensive-margin" compliers who would attend college regardless of the additional grant aid at the threshold. My approach extends sample-trimming methods which use identified complier shares to place best-case and worst-case bounds on treatment effects for inframarginal students (Gerard et al. 2020; Lee 2009; Manski 1989). I depart from these studies by showing that the pre-college characteristics of extensive- and intensive-margin compliers are separately point identified, even with selected data conditioned on college enrollment. Using this result, I demonstrate that extensive-margin students induced to attend college by aid awards are heavily negatively selected, with average tenth grade test scores far below those of intensive-margin compliers. This finding motivates a novel mean-dominance assumption which substantially tightens nonparametric bounds on grant treatment effects, yielding useful bounds when no-assumptions bounds (Gerard et al. 2020; Lee 2009) are too wide to be informative.

This approach overcomes a common challenge in studies of higher education programs, which frequently lack access to the full set of data on all college applicants and instead use data on the selected sample of college enrollees (Rothstein 2004). In the context of grant aid, the vast

[6] I find that this difference is essentially unchanged by the introduction of merit standards to the TEXAS Grant in 2014, suggesting that schools independently value ex-ante high achieving students when awarding grants.

majority of regression-discontinuity studies with enrollment-selected data restrict their attention to programs with zero extensive-margin enrollment effects.[7] This restriction allows for valid RD estimation but necessarily limits attention to the effects of grant aid on inframarginal students who are not induced by aid awards to enroll in college. Even when researchers have access to data on all applicants (Angrist et al. 2022; Bettinger et al. 2019; Cohodes and Goodman 2014), the difficulty of identifying precisely who among the compliers is an extensive-margin complier leads them to analyze effects on enrollment and outcomes separately.[8] I obtain separate bounds on treatment effects for students at both complier margins — those induced by aid to attend college, and those who would attend college anyway — allowing comparisons of the impact of grant aid on these two groups. Furthermore, I show how my framework recovers estimates of grant aid's total effect, pooled across these two groups, providing a policy-relevant parameter for evaluating expansions of grant aid even when data is selected and grants affect enrollment.

Section 8 applies the bounding method to separately estimate the impacts of receiving a Zero EFC on the B.A. completion rates and future earnings of intensive-margin and extensive-margin community college students. I begin by showing that standard no-assumptions bounds are too wide to be informative, failing to reject either positive or negative effects of grant aid on students' B.A. completion at any time horizon. However, applying the mean-dominance assumption developed in Section 7 reveals small positive effects of zero-EFC eligibility on intensive-margin community college students' probability of eventual B.A. completion and early-career earnings. The results indicate that failing to correct for sample selection bias masks positive grant impacts on inframarginal students with negative compositional bias caused by the entry of extensive-margin students.

Taken together, the TEXAS Grant and Auto-Zero results suggest large impacts of relatively small increases in grant aid on the outcomes of lower-income students, but null effects of full scholarships for moderately higher-income students. Section 9 unifies these results with a cost-benefit analysis that compares the impact of additional grant aid on student outcomes across the three treatment margins. I derive the costs and benefits of a marginal expansions of grant aid eligibility, and show how these parameters can be expressed in terms of complier shares and margin-specific LATEs. Applying the empirical results to my framework, I find that while the TEXAS Grant is essentially a transfer to middle-income students, the benefits of providing grant aid to lower-income students enrolled in four-year colleges exceed their costs. An implication of these results is that the current policy environment provides aid inefficiently, suggesting potentially

[7] A notable exception is Park and Scott-Clayton (2018), who analyze the impact of the Pell Grant on the outcomes of community college students in an enrollment-selected sample. They consider two approaches: restricting attention to the subsample of community colleges with no apparent enrollment effect, and using Lee (2009)-style bounds to estimate the Pell Grant's impact on intensive-margin students (Gerard et al. 2020). The latter strategy is the starting point for the extended bounding analysis considered in my paper.

[8] Denning (2018) discusses this concern, noting that many published estimates of the effect of grant aid on college graduation may function primarily through their impact on initial college enrollments.

large gains from budget-neutral reallocations of grant aid towards lower-income students enrolled in four-year colleges.

Related Literature The empirical results contribute to a large literature studying the impact of grant aid on student outcomes (see Deming and Dynarski (2010), Dynarski and Scott-Clayton (2013), and Dynarski et al. (2022) for reviews). The literature has consistently found that aid can have both modest impacts on students' college enrollment (Angrist et al. 2022; Dynarski 2003) and improve academic outcomes conditional on entry (Denning 2018; Scott-Clayton 2011), although findings vary considerably by the program studied. I contribute to this literature in several ways. First, I present a novel method for separately estimating the impact on grant aid on the outcomes of inframarginal students and students induced by aid to enter college. Previous studies using enrollment-selected data focus near-exclusively on grant programs with zero enrollment effects due to concerns over selection bias stemming from the entry of students into the sample (Denning et al. 2019; Scott-Clayton and Zafar 2019). Even when researchers have access to data on the full set of a particular program's recipients (Angrist et al. 2022; Bettinger et al. 2019; Castleman and Long 2016; Cohodes and Goodman 2014), they typically estimate effects on enrollment separately from effects on academic outcomes. As a result, it is unclear whether impacts on graduation are larger for inframarginal students or for students induced by aid to enroll.[9]

In addition, both grant aid programs I study have been the subject of prior work. Two prior studies examine the TEXAS Grant, finding mixed results; while Villareal (2018) finds large increases in graduation and future earnings among TEXAS Grant recipients, Montenegro (2020) finds evidence of decreases in loans and GPA among grant recipients.[10] Rather than focusing on outcomes alone, this paper analyzes the targeting mechanisms behind the TEXAS Grant's null effects, finding that the grant's unique allocation system, which allows schools to decide grant aid recipients, disproportionately targets aid at the margin to students with high preexisting graduation probabilities. Relative to previous studies of the Automatic Zero EFC threshold (Denning et al. 2019; Eng and Matsudaira 2021), I show how even when grants have substantial enrollment impacts, information on the characteristics of students who enroll due to grant aid can be used to tighten nonparametric bounds on the impact of aid on student outcomes.

Methodologically, this paper contributes to the literature on estimating nonparametric bounds on treatment effects (Manski 1989, 1990, 1997; Manski and Pepper 2000). The bounds I develop take Lee (2009) bounds as a foundation, and I use methodological results from Gerard et al. (2020) to

[9] A limited number of studies have examined the relationship between enrollment impacts and treatment effects on outcomes by estimating separate models by subgroups and examining the correlation between the two effects. These studies generally find that graduation impacts are larger among groups with larger enrollment impacts (Angrist et al. 2022; Cohodes and Goodman 2014).

[10] I am unable to replicate the results of Villareal (2018), who uses a similar research design to find large positive impacts of the TEXAS Grant on student outcomes.

implement them in the regression-discontinuity context. I show how the "no-assumptions" bounds considered in these studies can be substantially tightened by imposing assumptions on the relative potential outcomes of students at different complier margins, and show how these assumptions can be motivated by separately identifying the pretreatment characteristics of different complier groups. While similar mean dominance assumptions have been studied in empirical settings where data on untreated individuals is available (Chen et al. 2018), I show how the complier characteristics required to motivate them are identified even with data selected on enrollment. Other work uses parametric assumptions (Frangakis and Rubin 2002; Kline and Walters 2016) or multiple instruments (Kirkeboen et al. 2016; Mountjoy 2022) to disentangle causal effects across different complier margins, but these require both access to data on untreated units and several sources of variation, which are frequently unavailable in settings where data availability depends on take-up.

The paper proceeds as follows. After describing the setting and data in Section 2, I turn in Section 3 to describing the empirical challenge posed by enrollment effects when estimating grant treatment impacts. I evaluate these enrollment effects in Section 4, finding that they are empirically large in one of my settings. Section 5 estimates the causal effects of the TEXAS Grant and compares them to the impact of grant aid for lower-income students, and Section 6 uses observed patterns of grant allocation to discern mechanisms behind the TEXAS Grant's null effects. Section 7 develops the bounding argument, and Section 8 presents the main results from implementing it. Section 9 unifies the estimates with a cost-benefit analysis. Section 10 concludes.

2 Setting and Data

The setting for my analysis is the public higher education system in Texas. Several features of this environment make it particularly well-suited to studying the effects of grant-based financial aid on students' enrollment choices and labor market outcomes. First, the state's higher education system is massive: during my sample period, the state's higher education system represents over 7 percent of U.S. undergraduate enrollment, and over a tenth of all U.S. students in public college institutions (U.S. Department of Education 2021). The system also provides an ideal setting for studying the impact of grant aid on college choice: in a typical year, roughly 1.6 million undergraduate students enroll at one of 35 public four-year colleges, 60 public two-year colleges or technical schools, or 40 private institutions. The setting therefore allows me to directly study how grant-based financial aid affects students' decisions over where to go to college.

2.1 The TEXAS Grant

The main focus of my paper is the TEXAS Grant, the state's flagship need-based financial aid program. The program is the largest financial aid program administered by the state, with yearly disbursements ranging from $200 Billion to roughly $400 billion over my sample period. To be eligible for a TEXAS Grant, students must be enrolled at least three-quarters time in a public Texas four-year institution.[11] I limit my focus to TEXAS Grants allocated through the grant's "high school graduation pathway," which limits eligibility to students who enroll in college within sixteen months of graduating from a Texas high school. This pathway generally represents over 97 percent of TEXAS Grants.[12] Students who receive an "initial-year" grant in their first year of college can renew the TEXAS Grant for up to five years, subject to requirements on satisfactory academic performance.

Eligibility for the TEXAS Grant in the student's initial year of college is determined by each student's Expected Family Contribution (EFC). To be assigned an EFC, students must complete the Free Application for Federal Student Aid (FAFSA). The EFC is centrally calculated from information reported on the FAFSA using a complicated nonlinear formula that takes household income, assets, family size, government benefits, and other factors as inputs. Every year, the Texas Higher Education Coordinating Board (THECB) determines a "priority EFC" value that is distributed to colleges and universities for the purpose of allocating TEXAS Grants.[13] Institutions are instructed to give priority to students whose EFCs fall below this priority value, providing the basis for my research design.[14]

Importantly, because the TEXAS Grant is oversubscribed, not all EFC-eligible students receive grants. The program rules give substantial discretion to schools in choosing which students are awarded Grants, subject to the eligibility rules set by the state. These eligibility rules have changed

[11] A related grant program, the Texas Educational Opportunity Grant (TEOG), provides grant aid to students at two-year community and technical colleges. The TEXAS Grant provided funding to two-year community college students until the 2014 academic year, when part of the TEXAS Grant was reappropriated to fund the TEOG. In practice, the programs are similar but target different academic sectors; the TEXAS Grant funds students in four-year B.A.-granting institutions, and the TEOG funds students in two-year community colleges. Both programs are allocated using a cutoff rule based on students' Expected Family Contribution (EFC), and the cutoff is the same for both programs. Both programs provide funding covering the entirety of tuition and fees for students receiving grants. The results in Section 4 strongly indicate that crossing the TEXAS Grant cutoff does not influence students' enrollment decisions, limiting concerns that the confluence of these grant cutoffs cause selection bias in my estimates of the TEXAS Grant's effects. I primarily limit my focus in this paper to the TEXAS Grant, which funds a larger share of eligible students and provides larger increases in grant aid than the TEOG (Baum and Blagg 2021).

[12] See https://reportcenter.highered.texas.gov/meeting/board-supporting-documents/vii-d-texas-grant-program-report/. Three other "pathways" exist towards receiving the TEXAS Grant. They are the Associate Degree pathway, which targets recent recipients of associate degrees from Texas institutions; the Honorable Military Discharge pathway, which targets recently discharged military service members, and the Transfer Pathway, which funds students who receive a Texas Educational Opportunity Grant (TEOG) in a Texas two-year institution and transfer to a four-year school. The Associate Degree and Honorable Military Discharge pathways both require that recipients be graduates of Texas high schools.

[13] The priority EFC was constant at $4,000 until 2014. A policy change in 2014 pinned the priority EFC value to 60 percent of the average statewide amount of tuition and required fees for general academic teaching institutions in each academic year.

[14] Institutions are permitted by THECB to make exceptions to this guidance and give some grant awards to students with EFCs above the threshold. I document in Section 5 that while institutions frequently award grants to students with EFCs above the threshold, in practice the probability of receiving a TEXAS Grant jumps significantly at the priority EFC cutoff.

over time. For cohorts entering college prior to 2013, the main guidance issued to was to target grants based on financial need, prioritizing students with EFCs below the TEXAS Grant threshold. In contrast, for cohorts entering college in 2014 or later, the state added a set of merit-based criteria that gave priority to students with high levels of pre-college achievement.[15]

2.2 The Automatic Zero EFC Threshold

My second source of variation in grant aid is a federal cutoff rule that determines the likelihood that a student is assigned a $0 Expected Family Contribution. This "automatic zero EFC" cutoff is the subject of prior work by Denning et al. (2019), who use THECB data to evaluate the impact of additional Pell grant aid on student outcomes. I revisit their results as a point of comparison to the TEXAS Grant, comparing grant impacts at the TEXAS Grant threshold to the impacts of providing grant aid to lower-income students at the margin of $0 EFC assignment.

Variation in grant aid at the automatic zero EFC threshold stems from an income-based cutoff in the federal government's formula for calculating the EFC. Specifically, students with family Adjusted Gross Income (AGI) below a year-specific threshold become automatically eligible to be assigned an EFC of $0, making them eligible to receive the maximum federal Pell grant. In addition, because the EFC is used as an input to determine other grant awards (including the TEXAS Grant), receiving an automatic zero EFC can substantially increase grant aid from other sources besides the Pell grant. Because the EFC depends on factors other than family AGI, it is possible both for families to have a $0 EFC despite being on the ineligible side of the threshold, and to have a nonzero EFC despite being on the eligible side (Denning et al. 2019). Finally, in contrast to the TEXAS Grant, the auto-zero EFC threshold affects grant awards at both four-year and two-year institutions, both of which are included in my analysis.

2.3 Data Sources and Sample Construction

To conduct my empirical analysis, I combine several restricted administrative registries from Texas which together describe students' trajectories from high school into college and the labor market in early adulthood.[16] The starting point for the dataset is student-level data from the Texas Education Agency (TEA) on the full population of graduates from Texas public high schools. After restricting to students with valid student identifiers, demographics, and test scores, I link this population of high school graduates to enrollment, graduation, and financial aid records from the Texas Higher Education Coordinating Board (THECB). Finally, I link students to earnings records from the Texas Workforce Commission, covering job-level quarterly earnings for all work subject

[15] I return to this distinction in Section 6, analyzing the impact of this policy change on the composition of grant recipients and the impacts of grant aid on their outcomes.

[16] I access the data through a data-sharing agreement with the UT Austin Education Research Center (https://texaserc.utexas.edu/).

to the state's unemployment insurance (UI) system. Finally, I supplement these adminsitrative records with several school- and neighborhood-level data sources: tract-level measures of economic disadvantage measured by students' high school locations from the American Community Survey,[17] and institution-level measures of college tuition and instructional costs from the Integrated Postsecondary Education Data System (IPEDS).

I classify students' college enrollment status (two-year, four-year, or no college) by their first observed college enrollment within two years of high school graduation. The choice of a two-year window for measuring enrollments is motivated by the TEXAS Grant's allocation rules, which require that grant recipients enroll within sixteen months of completing high school to be eligible. In practice, this two-year timeline for defining enrollments captures roughly 95 percent of all high school graduates for whom I ever observe any college enrollment.

The main analysis sample consists of eleven cohorts of Texas public high school graduates whose initial college enrollment is between 2007 and 2017. 2007 is the oldest cohort for which complete financial aid data are available for all FAFSA submitters, and 2017 is the latest cohort for which I observe five-year college graduation outcomes. Observing outcomes at longer time horizons requires dropping later-entering cohorts from the sample; I observe six-year graduation rates, enrollments, and earnings for students entering college between 2007 and 2016, and eight years of these outcomes for students entering between 2007 and 2014.

Several additional restrictions are required when studying the Auto-Zero cutoff. First, I restrict to cohorts entering college after 2008, as this is the first year that family AGI (the running variable) is available in the data. Following Denning et al. (2019), I exclude students whose family Adjusted Gross Income (AGI) are multiples of $1,000 to avoid bias caused by bunching in the running variable (Barreca et al. 2015). Finally, I make an additional sample restriction related to the strength of the first-stage in grant aid awards at the Auto-Zero cutoff. A policy change in 2013 significantly lowered the family income threshold for receiving a Zero EFC, targeting aid to students who already received significant financial aid packages at baseline. Appendix Figure A1 shows that as a result, the first-stage increase in first-year grant aid associated with crossing the threshold becomes statistically indistinguishable from zero from 2013 onwards. For this reason, I primarily limit attention to the 2008-2012 entry cohorts when studying the automatic zero EFC threshold, when crossing the threshold corresponded to a large increase in aid.

2.4 Key Data Challenge

I encounter an important data limitation which is common in empirical studies of higher education (Rothstein 2004). The econometric challenge is sample selection: I only observe the running

[17] I use 2009-2013 five-year American Community Survey estimates, matching students' high school location to census tracts.

variables for my RD analyses in the select sample of enrolled students who appear in the financial aid data, not in the full population of college applicants. A consequence is that if grants cause students to enroll in college, then the samples of students to either side of the threshold will not be directly comparable. I show in Section 3 that failing to account for this sample selection problem will result in biased estimates whenever grants have causal effects on enrollment. This missing-data problem motivates my bounding analysis in Section 7, which uses the available data to place bounds on treatment effects for intensive-margin and extensive-margin compliers in the presence of the grant's enrollment effects.

2.5 Descriptive Statistics

Table 1 reports summary statistics for students near the three sources of variation in grant aid. Students near the TEXAS Grant threshold are substantially economically advantaged compared to students near the margin of receiving an automatic zero EFC. Near the TEXAS Grant threshold, students are roughly half as likely to have received a free or reduced-price lunch in high school, attended high schools with substantially lower poverty rates, and are substantially less likely to be nonwhite. Students near the TEXAS Grant threshold also have higher measures of pre-college preparedness compared to students near the margin of receiving an Automatic Zero EFC, with high school test scores that are six percentiles higher on average. The samples also differ considerably in terms of their baseline financial support; students on the ineligible side of the Auto-Zero cutoff receive substantially higher baseline grant awards than students on the ineligible side of the TEXAS Grant cutoff, while the latter group takes out substantially higher amounts in first-year loans.

3 Empirical Framework

This section lays the groundwork for my empirical analysis, describing the empirical challenge of identifying grant treatment effects in selected data when grants affect enrollment.

3.1 Setup, Notation, and Assumptions

Let $S \in \{0, 1\}$ be an indicator for college attendance, and assume that data is available only for students who attend college ($S = 1$).[18] Let $D \in \{0, 1\}$ denote a treatment that increases

[18]Throughout the paper, I conceptualize S as a particular college sector; for example, when studying the effect of the TEXAS Grant on the outcomes of four-year students, $S = 1$ denotes four-year schooling. In contrast, when studying the impact of receiving a Zero EFC on the outcomes of two-year community college students, $S = 1$ denotes two-year college attendance. With this definition, $S = 0$ denotes students' outside option, which can itself consist of several alternative schooling choices, including foregoing college entirely or attending college out of state. These choices are motivated by results in Section 5, which show that neither the TEXAS Grant nor the Auto-Zero threshold cause students to significantly change their enrollment choices conditional on a particular college sector (e.g., four-year schools vs. two-year schools).

Table 1: Summary Statistics

	TEXAS Grant (4-Year Enrollees)		Auto-Zero EFC (4-Year Enrollees)		Auto-Zero EFC (2-Year Enrollees)	
	Ineligible (1)	Eligible (2)	Ineligible (3)	Eligible (4)	Ineligible (5)	Eligible (6)
Panel A. Demographics						
Female	0.57	0.56	0.57	0.56	0.54	0.54
Asian	0.06	0.06	0.06	0.07	0.02	0.02
Black	0.18	0.16	0.22	0.23	0.17	0.16
Hispanic	0.39	0.34	0.50	0.44	0.60	0.55
White	0.35	0.42	0.21	0.26	0.20	0.25
Free/reduced price lunch	0.34	0.22	0.63	0.51	0.67	0.57
Test score percentile	67.61	69.43	61.89	62.95	38.35	40.72
Poverty Rate	0.17 (0.12)	0.16 (0.12)	0.21 (0.13)	0.20 (0.13)	0.21 (0.13)	0.20 (0.13)
Unemployment Rate	0.07 (0.04)	0.07 (0.04)	0.08 (0.05)	0.07 (0.05)	0.07 (0.04)	0.07 (0.04)
SNAP Receipt Rate	0.14 (0.11)	0.13 (0.11)	0.18 (0.13)	0.16 (0.13)	0.19 (0.13)	0.17 (0.12)
Panel B. Binary Treatments						
Any TEXAS Grant	0.62	0.06	0.72	0.64	0.38	0.36
Zero EFC	0	0	0.87	0.18	0.90	0.29
Panel C. Financial Aid						
Total Grant Aid	9,262	4,746	13,031	11,330	5,837	4,855
	(5,314)	(4,807)	(5,449)	(5,662)	(2,910)	(2,905)
TEXAS Grant	3,624	349	4,328	3,911	694	660
	(3,046)	(1,463)	(3,012)	(3,137)	(971)	(962)
Pell Grant	2,108	269	5,469	4,077	4,546	3,558
	(1,047)	(515)	(1,537)	(1,807)	(1,931)	(1,930)
HB3015 Set-Asides	523	1,005	480	467		
	(1,268)	(1,706)	(1,273)	(1,229)		
Total Loans	4,702	6,597	3,076	3,877	502	604
	(4,657)	(6,106)	(3,682)	(4,174)	(1,453)	(1,622)
Observations	39,123	28,107	23,704	20,841	34,645	27,992
	67,230		44,545		62,637	
Cohorts	2007-2017		2008-2012		2008-2012	

Notes: Table reports summary statistics for the main estimation samples. TEXAS Grant observations in columns 1 and 2 are students enrolled in four-year public TEXAS universities with incoming Expected Family Contribution (EFC) within $2,000 of the TEXAS Grant cutoff. Auto-Zero EFC observations in columns 3 through 6 are students enrolled in four-year (columns 3 and 4) and two-year (columns 5 and 6) colleges with incoming family Adjusted Gross Income (AGI) within $10,000 of the automatic zero EFC threshold. The sample is restricted to students who graduate from a public Texas high school within two years prior to first college enrollment. Poverty rate, unemployment rate, and SNAP receipt rate are measured by students' home census tract, as measured by the location of the high school from which they graduated. Dollar values are in 2019 dollars.

a student's grant aid award.[19] Treatment is influenced by a cutoff rule: each student has an exogenously-determined score k, and $Z = \mathbf{1}\{k \leq k^*\}$ denotes whether the student's score is below a cutoff k^* which affects the probability of treatment. I use D_z to denote students' potential grant status depending on whether they are on the eligible ($z = 1$) or ineligible ($z = 0$) side of the grant threshold. Potential college choices S_d then describe how students choose schooling in response to whether or not they receive grants. Finally, potential outcomes $Y_{s,d}$ describe students' outcomes, such as graduation or earnings, given schooling choice s and treatment status d.

Throughout the paper, I make two standard fuzzy regression-discontinuity assumptions: an exclusion restriction stating that crossing a grant threshold only affects students' college choices and outcomes through its impact on treatment, and a continuity restriction imposing that the potential outcome and college choice functions are continuous at the cutoff:

Assumption 1 (Exclusion and Continuity)**.**

1. *(Exclusion)* $(Y_{s,d} \perp Z) \mid (s, d, k)$ *and* $(S_d \perp Z) \mid (d, k)$.
2. *(Continuity)* $\mathbb{E}[Y_{s,d} \mid k]$ *and* $P(S_d = s \mid k)$ *are continuous at* $k = k^*$ *for* $d \in \{0, 1\}$ *and* $s \in \{0, 1\}$.

Importantly, Assumption 1 does *not* assume that grants are excludable, affecting student outcomes only through their impact on college choices. Instead, the notation $Y_{s,d}$ denotes that grants can affect student outcomes even conditional on college attendance, e.g., if $Y_{1,1} \neq Y_{1,0}$. However, the conditional expectation of these potential outcome functions is assumed to be continuous at the cutoff for every school choice s and grant status g.

In addition, I assume that grants only affect the outcomes of students who attend college.[20] This assumption collapses the outcomes of non-college-attenders, $Y_{0,d}$, into a single potential outcome Y_0 representing the non-attendance outside option:

Assumption 2 (Partial Exclusion)**.** $Y_{0,1} = Y_{0,0} \equiv Y_0$.

Finally, I impose two monotonicity restrictions for empirical tractability. The first is a standard Imbens and Angrist (1994) monotonicity assumption which establishes that crossing the grant threshold does not make it less likely that an individual receives a grant:

Assumption 3 (Grant Monotonicity)**.** $D_1 \geq D_0$ *for all individuals.*

The second monotonicity condition restricts how grant offers affect college choices. I require that treatment does not make it less likely that a student attends college:

[19] In the case of the TEXAS Grant, D represents whether a student is awarded an "initial-year" TEXAS Grant for use in the first year of college; In the case of the Auto-Zero EFC cutoff, D is an indicator for whether the students are assigned a zero EFC.

[20] This assumption is motivated by the structure of almost all college grant programs, which are offered as an in-kind subsidy that lowers the price of college attendance but which cannot be used for other purposes.

Assumption 4 (Enrollment Monotonicity)**.** $(S_1 \geq S_0) \mid k = k^*$ *for all individuals.*

Assumption 4 requires that any student at the grant cutoff ($k = k^*$) who would have attended college without treatment remains in the sample if they receive a grant. While this assumption is likely to hold for grant programs that are targeted at a particular college sector or institution,[21] it may not hold in general for all grant programs. In particular, if D represents a source of grant aid that can be used by students at institutions outside the scope of the data, then treatment may cause some students to leave the sample as a result of crossing to the eligible side of the threshold. This is a particular concern for the Auto-Zero EFC cutoff, which increases students' federal Pell Grant aid at any institution (including out-of-state schools) and therefore may cause students to leave the sample. I return to this concern in Section 4, showing that a small number of community college students (namely, those with high test scores) are "enrollment defiers" who leave the sample as a result of crossing the Auto-Zero EFC threshold and therefore must be excluded from the analysis.

Given Assumptions 1 through 4, the link between observed and potential outcomes for students at the cutoff is as follows:

$$
\begin{aligned}
Y &= S(Y_{1,1}D + Y_{1,0}(1-D)) + (1-S)Y_0 \\
S &= DS_1 + (1-D)S_0 \qquad\qquad (1)\\
D &= ZD_1 + (1-Z)D_0
\end{aligned}
$$

The first equation in (1) states that observed outcomes depend on college attendance S, and conditional on college attendance also depend on grant status D. The second equation states that college attendance itself depends on grant status, allowing for the possibility that some students enter the sample as a result of treatment. Finally, the third equation states that treatment depends on whether the student falls on the eligible or ineligible side of the cutoff. Together, these equations permit unobserved heterogeneity in treatment effects between students of different complier types. The next section defines these types and states target parameters given these potentially heterogeneous effects.

3.2 Target Parameters

3.2.1 The Impact of Grant Aid on Compliers

To clearly state target parameters, it is first useful to consider what would be identified if data was available on all applicants (including students with $S = 0$ as well as $S = 1$) rather than the selected sample of college enrollees with $S = 1$. Appendix C.1 shows that with unselected data, the limiting

[21]Examples of such programs include UT-Austin's Longhorn Opportunity Scholars program and Texas A&M's Century Scholars Program, studied by (Andrews et al. 2020).

Wald (1940) estimand at the cutoff identifies the following local average treatment effect (LATE):

$$\begin{aligned}\frac{\lim_{r\uparrow k^*}\mathbb{E}[Y \mid k]-\lim_{r\downarrow k^*}\mathbb{E}[Y \mid k]}{\lim_{r\uparrow k^*}\mathbb{E}[D \mid k]-\lim_{r\downarrow k^*}\mathbb{E}[D \mid k]} &= \mathbb{E}\left[Y_{S_1,1}-Y_{S_0,0} \mid D_1 > D_0, k=k^*\right] \\ &\equiv LATE\end{aligned} \tag{2}$$

The LATE in (2) describes the grant's impacts on the outcomes of compliers: students who receive a grant if and only if they cross to the eligible side of the grant threshold ($D_1 > D_0$). However, this overall effect summarizes the grant's impacts through two channels: its extensive-margin impact on college enrollment, and its intensive-margin impact on student outcomes holding enrollment fixed. This is seen in the expression $Y_{1,1} - Y_{S_0,0}$, which allows differences in outcomes both due to enrollment effects ($S_1 > S_0$) and and direct grant impacts for enrolled students ($Y_{1,1} \neq Y_{1,0}$).

3.2.2 Impacts for Extensive-Margin and Intensive-Margin Compliers

Equation (2) shows that a fuzzy regression discontinuity design recovers a valid LATE for compliers, summarizing the grant's impact on their outcomes through both enrollment and intensive-margin treatment effects. Appendix C.2 shows that this overall LATE can be disaggregated into the grant's impact on two sets of groups: *"intensive-margin"* compliers who attend college regardless of whether they receive a grant at the cutoff, and *"extensive-margin"* compliers who are induced by grant offers to enroll in college:

$$LATE = LATE_{IM} \times \pi_{IM|Complier} + LATE_{EM} \times \pi_{EM|Complier} \tag{3}$$

where $LATE_{IM}$ and $LATE_{EM}$ are given by:

$$\begin{aligned} LATE_{IM} &= \mathbb{E}\left[Y_{1,1}-Y_{1,0} \mid S_1 = S_0 = 1, D_1 > D_0, k=k^*\right] \\ LATE_{EM} &= \mathbb{E}\left[Y_{1,1}-Y_0 \mid S_1 > S_0, D_1 > D_0, k=k^*\right]\end{aligned}$$

and where $\pi_{IM|Complier}$ and $\pi_{EM|Complier}$ describe the shares of students who fall into the two groups:[22]

$$\begin{aligned} \pi_{IM|Complier} &= P(S_1 = S_0 = 1 \mid D_1 > D_0, k=k^*) \\ \pi_{EM|Complier} &= P(S_1 > S_0 \mid D_1 > D_0, k=k^*)\end{aligned}$$

[22]Note that $\pi_{IM|Complier} + \pi_{EM|Complier}$ need not add to 1 in this setting. The reason is that there may be grant compliers who have $D_1 > D_0$ but do not have $S_1 = 1$. An example is a student who is offered a grant as a result of crossing the threshold, but who does not attend college regardless of grant receipt ($S_1 = S_0 = 0$). The LATE for these students is zero by Assumption 2.

For intensive-margin compliers, $LATE_{IM}$ measures the impact of additional grant aid holding enrollment fixed ($Y_{1,1} - Y_{1,0}$). In contrast, for extensive-margin compliers, $LATE_{EM}$ measures the combined impact of grant aid *and* college enrollment, compared to these students' no-enrollment outside option ($Y_{1,1} - Y_0$).

Many possible combinations of $LATE_{IM}$ and $LATE_{EM}$ could produce the overall $LATE$ identified in (2). Moreover, the difference between these component "subLATEs" has important policy implications for how grants should be targeted. On the one hand, if $LATE_{EM}$ is larger than $LATE_{IM}$, then the overall impact of grant aid is larger for students who attend college as a result of receiving funds. This case would suggest large benefits from policies that target grant aid towards college students who are unlikely to attend college otherwise (Dynarski et al. 2021). On the other hand, if extensive-margin compliers forego labor-market earnings to attend college or are unlikely to complete college at high rates, then $LATE_{IM}$ may exceed $LATE_{EM}$, justifying grant programs that are primarily targeted at intensive-margin students.

Unfortunately, however, $LATE_{IM}$ and $LATE_{EM}$ are not separately point identified even with data on the full population of college applicants. The reason is a fundamental underidentification problem that arises in any setting with multiple treatment margins but only one instrument: though it is possible to identify the overall $LATE$ and the shares of both complier types in (3), the data do not reveal who among the compliers is an intensive-margin complier and who is an extensive-margin complier.[23] To address this problem, I develop a bounding approach in Section 7 that delivers nonparametric bounds on $LATE_{IM}$ and $LATE_{EM}$. However, the overall $LATE$ is itself a parameter of substantial policy interest. It describes the overall effect of grant aid on the outcomes of students, summarizing both the grant's impact on enrollments and its direct impact on student outcomes, and is the main target parameter in grant aid studies with full data on the population of applicants (Angrist et al. 2022; Bettinger et al. 2019). I show in Section 9 that a related parameter — the LATE for enrolled compliers — provides the basis for cost-benefit analysis that facilitates comparisons of the welfare impacts of different marginal grant aid expansions.

3.3 The Empirical Challenge: With Selected Data, Standard IV is Biased

The previous section lays out how the target parameters — the combined $LATE$, the component sublates $LATE_{IM}$ and $LATE_{EM}$, and the population shares of intensive-margin and extensive compliers — capture estimates of grant aid's causal impacts on student outcomes. However, a challenge arises when data is only available on the population of enrolled students. In this case, extensive-margin compliers will appear in the dataset *only* on the eligible side of the cutoff, as grant

[23] In fact, when there are multiple treatment margins, standard IV methods such as two-stage least squares do not recover margin-specific treatment effects even when there are as many instruments as treatment margins (2016). However, Mountjoy (2022) shows that with two treatment margins and two continuous instruments, it is possible to recover margin-specific treatment effects using a nonparametric separate-identification approach.

awards resulting from crossing the grant threshold cause this population to enroll. As a result, the populations to either side of the cutoff are no longer directly comparable, biasing standard IV approaches to estimating the impact of grant aid on outcomes. I show in Appendix C.3 that failing to account for this compositional bias will result in IV estimates that fail to recover a LATE for any group. When grants cause some students to enroll, the limiting Wald estimand at the cutoff in enrollment-selected data becomes:

$$\frac{\lim_{k\uparrow k^*}\mathbb{E}[Y \mid S=1,k]-\lim_{r\downarrow k}\mathbb{E}[Y \mid S=1,k]}{\lim_{k\uparrow k^*}\mathbb{E}[D \mid S=1,k]-\lim_{k\downarrow k^*}\mathbb{E}[D \mid S=1,k]}$$

$$= LATE_{IM}\times\underbrace{\left(\frac{\pi^-_{IM}}{\pi^-_{IM}+\pi^-_{EM}(1+\bar{D}^+)}\right)}_{\text{Sample Size Bias}} + \underbrace{\left(\mathbb{E}[Y_{1,1}\mid \text{EM Complier},k^*]-\lim_{k\downarrow k^*}\mathbb{E}[Y \mid S=1,k]\right)\times\left(\frac{\pi^-_{EM}}{\pi^-_{IM}+\pi^-_{EM}(1+\bar{D}^+)}\right)}_{\text{Compositional Bias}} \tag{4}$$

where $\pi^-_{IM}=P(S_1=S_0=1,D_1>D_0 \mid S_1=1,k=k^{*-})$ and $\pi^-_{EM}=P(S_1>S_0,D_1>D_0 \mid S_1=1,k=k^{*-})$ denote the share of intensive- and extensive-margin compliers directly to the left (barely-eligible) side of the cutoff in the enrollment-selected sample. The term $\bar{D}^+=\lim_{k\downarrow k^*}\mathbb{E}[D \mid S=4,k]$ denotes the share of students in the enrollment-selected sample who are so-called "always-takers," receiving grants despite being on the right (barely-ineligible) side of the threshold.[24]

Equation (4) shows that when grants cause some students to enroll in college, a naïve fuzzy RD across the threshold recovers a biased estimate of $LATE_{IM}$, the treatment effect for intensive-margin compliers. This bias enters the limiting Wald estimand in two ways. The first term is the sample size bias, which arises because the sample on the eligible side of the cutoff becomes larger due to the entry of extensive-margin compliers. The second term is the compositional bias, which depends on $\mathbb{E}[Y_{4,1} \mid \text{EM Complier},k^*]$, the mean treated potential outcome of extensive-margin compliers. The sign of this bias term depends on how extensive-margin compliers' treated potential outcomes differ from the mean potential outcome across all students on the ineligible side of the cutoff. Unfortunately, even if the second term is zero, the presence of *any* enrollment effect will cause the sample size bias to be nonzero, leading the selected-sample Wald estimand to be biased downwards relative to $LATE_{IM}$. Equation (4) recovers $LATE_{IM}$ if and only if there are no enrollment effects so that π^-_{EM} equals zero.

Of course, one solution to this identification challenge is to restrict attention to grant aid

[24]Note that because the entry of extensive-margin compliers increases the sample size on the eligible side of the cutoff, it is no longer the case that the sample shares of various complier types are equal approaching the cutoff from both sides. For this reason, I use the notation k^{*-} and k^{*+} to denote limits of the cutoff from the left (eligible) and right (ineligible) sides.

programs that have no impact on student enrollment. Indeed, this is the approach taken by many RD studies of grant aid using selected data, which first establish the absence of enrollment effects by showing continuity of the density of the running variable at the grant threshold (2016; 2019; 2019). Because all compliers are intensive-margin compliers in this case, both bias terms in (4) disappear, due to the fact that $\pi_{EM}^{-} = 0$. As a result, the limiting Wald estimand is a valid estimator of $LATE_{IM}$ whenever grant programs do not have enrollment effects.

However, the fact that enrollment-selected data is ubiquitous in higher education (Rothstein 2004), together with the fact that many grant aid programs have impacts on enrollment (Dynarski 2000), suggests the need for an alternative approach that delivers valid estimates of treatment effects even in the presence of the selection bias caused by enrollment effects. Before developing such an approach in Section 7, I first evaluate the extent of this enrollment bias in my empirical settings.

4 How Large are Enrollment Effects?

Equation (4) shows that when grant aid programs have enrollment effects, standard RD estimation in a selected sample of enrolled students will be biased. In this section, I evaluate the potential for this bias in my setting by estimating the enrollment effects of the two grant programs I study: the TEXAS Grant and the Automatic Zero EFC Cutoff.

4.1 Measurement and Testing

With data on all college applicants, it would be possible to directly observe changes in enrollments across the cutoff, making it possible to measure grants' enrollment effects using an RD estimator with an indicator for college enrollment as the outcome variable.[25] Such an estimator is not possible in my setting because my dataset is itself conditioned on college enrollment. However, if the increase in grant aid at the cutoff causes some students to enter the sample, then the density of the running variable will jump discontinuously at the cutoff. As a result, a standard regression discontinuity manipulation test (e.g., McCrary 2008) provides a useful diagnostic test for evaluating whether standard IV is appropriate for estimating grant impacts.

[25] In fact, with data on all college applicants and a grant system targeted towards a single schooling option (e.g., four-year schooling), variants of Abadie (2002)-style regressions identify not only the overall share of extensive-margin compliers but the distribution of their origin locations. This distribution is described in Angrist et al. (2022) and Abdulkadiroğlu et al. (2017) as the "distribution of counterfactual destinies."

4.2 Estimates of Enrollment Effects

4.2.1 Enrollment Effects among Students in Four-Year Colleges

Figures 1a and 1b plot estimates of the density of the running variable at the TEXAS Grant and Auto Zero thresholds for students in four-year colleges. Panel (a) shows the density of the EFC for students near the TEXAS Grant threshold, and Panel (b) shows the density of AGI for students near the Auto-Zero threshold. To make the density estimates comparable, I normalize the y-axis in each figure so that 1 equals the point estimate of the density on the right (barely-ineligible) side of the cutoff. Each figure plots a scaled histogram of the running variable for each cutoff at regularly-spaced bins,[26] together with estimated densities and bias-corrected confidence intervals constructed using the method of Cattaneo et al. (2018).

The estimates in Figures 1a and 1b indicate that neither the TEXAS Grant nor the Automatic Zero EFC threshold has a significant effect on student enrollments in four-year schools. While an enrollment effect would cause the density on the left (barely-eligible) side of the cutoff to be higher than the density on the right, there is no evidence of such a jump in either sample. Appendix Tables B1 and B2 confirm this visual evidence, showing results from formal McCrary (2008) and Cattaneo et al. (2018) tests of the null hypothesis of equality at the cutoff. I cannot reject that the density is continuous across a wide variety of estimation bandwidths and subsamples; for example, the estimated (McCrary 2008) difference in densities at the TEXAS Grant cutoff is exactly zero with a standard error of (0.23), and the estimated difference at the Auto-Zero cutoff is -0.018 with a standard error of (0.028).

A second heuristic test for enrollment effects is to examine whether the bivariate relationship between students' predetermined characteristics and the running variable jumps discontinuously at the cutoff. Appendix Figures A2 and A3 plot these relationships and report RD estimates at the threshold for a wide variety of predetermined covariates at the TEXAS Grant and Auto-Zero cutoffs in four-year schools. I find no evidence of a statistically significant discontinuity in any student characteristic at either the TEXAS Grant or the Auto-Zero thresholds for students attending four-year colleges, suggesting that the composition of students does not meaningfully change across either cutoff.

I also find no evidence that the increase in grant funding at the TEXAS Grant threshold causes four-year students to "upgrade" their schooling choices, e.g., by attending more expensive or higher-quality schools. Appendix Table B4 reports regression discontinuity estimates where the outcomes are means of institutional characteristics, measured at the entry cohort-by-institution level. If TEXAS Grant recipients change their enrollment choices, then cohort-by-school characteristics should be discontinuous at the grant threshold. I find no evidence of any such discontinuity in

[26] The binwidth for Panel (a) is \$100 of EFC, and the binwidth for Panels (b) is \$1,000 of AGI.

the mean demographic characteristics, standardized test scores, or mean financial aid awards of students' peers across the threshold. These results hold even when restricting the sample to students whose application records indicate acceptance to more than one college (columns (3) and (4)), arguably the population most likely to respond to the TEXAS Grant by changing their enrollment behavior. Overall, the results suggest that at the margins of receiving a TEXAS Grant or an Automatic Zero EFC, the increase in grant aid at the threshold does not cause students to enroll in four-year colleges or change their enrollments conditional on this sector.

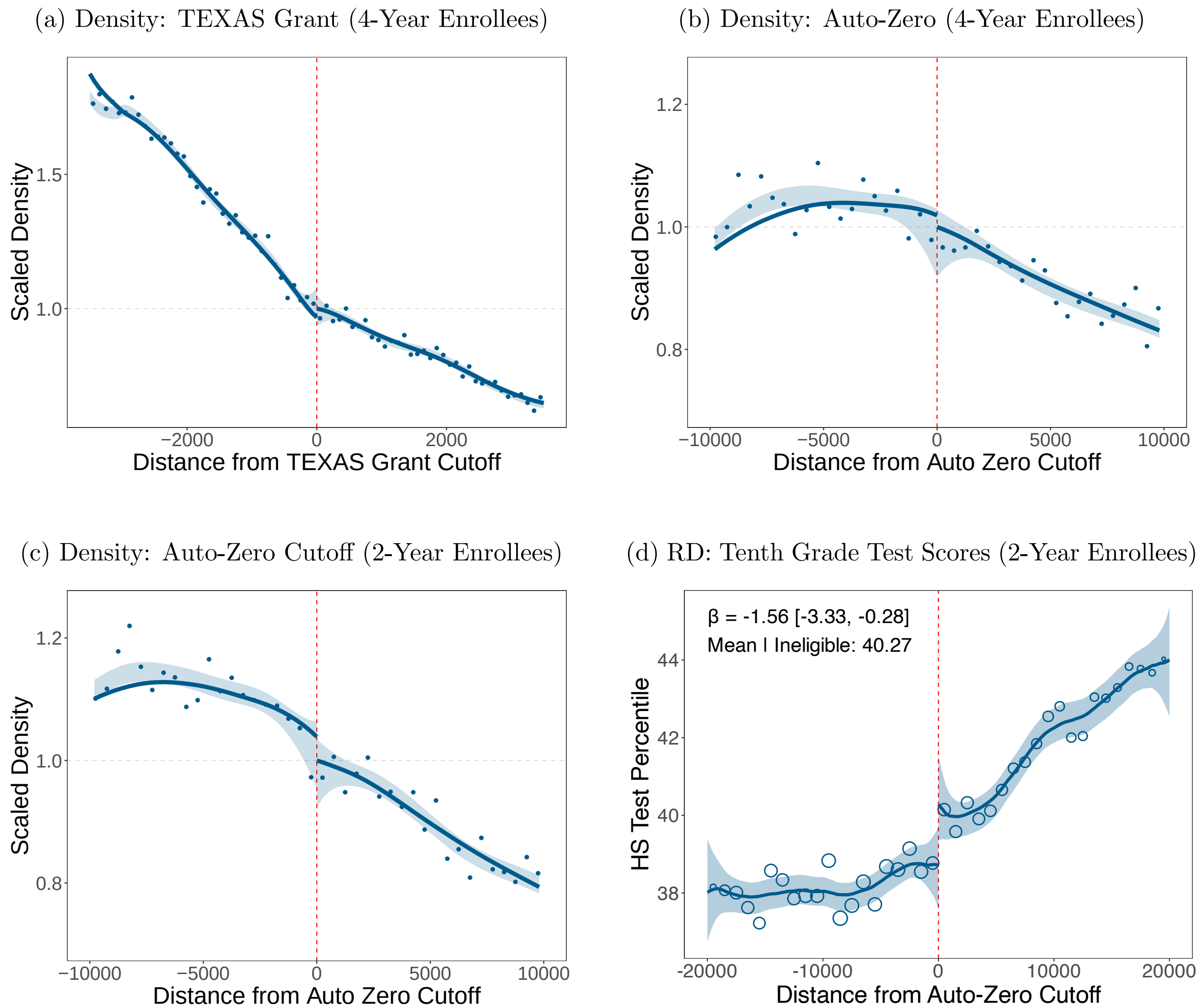
4.3 Enrollment Effects at the Auto-Zero Threshold in Two-Year Colleges

Figures 1c and 1d investigate whether crossing the Auto-Zero EFC threshold causes students to enroll in two-year community colleges. In contrast to the results in four-year schools, and consistent with the findings of Denning et al. (2019), I find evidence that students respond to receiving a zero EFC by changing their enrollment behavior. Figure 1c shows visual evidence of a jump in the number of students attending community colleges on the left (barely-eligible) side of the threshold, consistent with an effect of zero-EFC eligibility on students' enrollment behavior. Figure 1d reinforces this conclusion, plotting students' average tenth-grade test scores in a window of the Auto-Zero cutoff. Students on the barely-eligible side of the Auto-Zero threshold have mean test scores that are roughly 1.5 percentiles lower than students on the barely-ineligible side, and the difference is significant at the 95 percent level. Appendix Figure A4 shows that similar discontinuities appear in many predetermined covariates: two-year community college students on the barely-eligible side of the threshold are less likely to be white and have higher home-tract poverty and SNAP benefit receipt rates than students on the barely-ineligible side, as measured by the location of their high schools. Taken together, the results suggest that standard fuzzy RD methods applied across the Auto-Zero threshold in two-year schools will not yield valid estimates of treatment effects.

4.3.1 Testing for the Presence of Defiers

Importantly, the patterns documented in Figures 1c and 1d do not by themselves prove that the compositional changes at the cutoff are entirely due to the entry of students who otherwise would not have enrolled in college. Indeed, because a $0 EFC qualifies students for the maximum federal Pell Grant, it is possible that some students may *leave* the sample as a result of crossing the cutoff, for example by enrolling in four-year or private institutions. Students of this behavior type are "enrollment defiers" in terms of the enrollment monotonicity assumption (Assumption 4) and represent a threat to identification. If some students leave the sample as a result of crossing to the eligible side of the threshold, then regression-discontinuity estimates across the cutoff will conflate the causal impacts of grant aid programs with compositional changes caused by the departure of

Figure 1: Enrollment Effects at the TEXAS Grant and Auto-Zero Cutoffs



Notes: Panels (a) through (c) show the density of the running variable in a neighborhood of the assignment threshold for all three estimation samples. Panel (a) shows the density of four-year college enrollees' Expected Family Contribution (EFC) in a neighborhood of the TEXAS Grant cutoff. Panels (b) and (c) show the density of students' incoming family Adjusted Gross Income (AGI) in a neighborhood of the Auto-Zero EFC threshold, with Panel (b) showing four-year enrollees and Panel (c) showing two-year enrollees. Note that the running variables differ across the programs due to different assignment rules. The y-axis in Panels (a) through (c) is normalized so that 1 equals the density point estimate on the right (ineligible) side of the cutoff. The points show histogram estimates of the running variable. The smoothed fits to either side of the cutoff are local-linear estimates of the density and the blue shaded regions are 95 percent bias-corrected confidence intervals produced using the method of Cattaneo et al. (2018). Panel (d) plots the bivariate relationship between two-year community college students' entering family AGI and their tenth grade test score percentile. The circles show raw means by $1,000 AGI bins. The lines and shaded regions on either side of the threshold are local-linear fits with 95 percent bias-corrected confidence intervals produced using the method of Calonico et al. (2014). The annotation reports the estimated discontinuity at the cutoff together with the estimated fit approaching the threshold from the right (ineligible) side. The plots are shown for the main estimation samples: the 2007-2017 entering cohorts for the TEXAS Grant, and the 2008-2012 entering cohorts for the Auto-Zero threshold.

enrollment defiers from the sample. Moreover, even the number of choice defiers is not identified due to the fact that other students (the extensive-margin compliers) enter the sample as a result of treatment. The overall positive enrollment effect in Figure 1c could therefore be caused by either small or large numbers of defiers leaving the sample, as long as the offsetting number of extensive-margin compliers is slightly larger.

However, if defiers exist, then the compositional changes in Figure 1d offer a clue as to their likely composition. The patterns at the cutoff are consistent with defiers being positively selected, attending community colleges despite having high tenth-grade test scores but changing their enrollment behavior (for example, by attending four-year schools) if offered additional grant funding. Appendix Figure A5 finds support for this hypothesis, separately plotting the density of the running variable for two sub-populations of community college students: students in the first four quartiles of the high school test score distribution in Panel (a), and students in the top quintile in Panel (b). Figure A5a shows positive enrollment effects among the roughly 90 percent of community college students who score below the 80th percentile on standardized tests among their high school graduating class, indicating that for this population an automatic zero EFC causes some extensive-margin compliers to enroll in two-year schooling. In contrast, Figure A5b shows a significant negative enrollment effect among *ex-ante* high achieving community college students at the threshold. This discontinuity is statistically significant; Appendix Table B3 shows that the p-value associated with the McCrary (2008) test of equality at the cutoff is 0.028. Moreover, as Appendix Figure A6 shows, this defier behavior appears limited to top-scoring students; the discontinuities among all other quintiles are consistent with at least weakly positive enrollment effects.

As a result of these patterns, the effect of financial aid on the outcomes of top-scoring community college students at the Auto-Zero cutoff is not identified. I therefore remove these students from the sample, dropping roughly 10 percent of community college students within my estimation window of the Auto-Zero cutoff. I assume that defiers do not exist among the remaining students, so that the enrollment monotonicity assumption (Assumption 4) holds in the remaining sample.

Taking Stock

Taken together, the results in this section suggest that standard IV methods are sufficient to study the impacts of the TEXAS Grant and Auto-Zero cutoffs among four-year enrollees. The results in Figure 1 imply that at the margins of these grants, *all* compliers are intensive-margin compliers who would attend four-year colleges even without the additional grant aid furnished by crossing each threshold. As a result, neither bias terms in (4) applies, so that a standard fuzzy RD recovers valid estiamtes of $LATE_{IM}$. I therefore proceed in Section 5 with standard fuzzy regression-discontinuity methods in order to estimate these grants' impacts on intensive-margin compliers.

In contrast, the results in Figures 1c and 1d suggest that standard IV methods applied to the 2-year Auto-Zero cutoff will fail to recover a valid LATE for any group of compliers. The reason is that there still exist extensive-margin compliers who enter the sample as a result of crossing the Auto-Zero EFC threshold, as shown in Figure A5a. If these students also have lower probabilities of graduating compared to intensive-margin compliers, then the compositional bias in (4) will be negative at the threshold, producing an overall downwards-biased estimate of $LATE_{IM}$ when combined with the (unambiguously negative) sample selection bias term. The bias from standard methods suggests the need for an alternative framework which recovers valid estimates of treatment effects even in the presence of sample selection. I return to this issue in Section 7, developing a nonparametric identification argument that separately develops bounds on treatment effects for extensive-margin and intensive-margin compliers.

4.4 Implementation

The previous sections indicate that for programs without enrollment effects —- namely, the TEXAS Grant and the auto zero cutoff in four-year schools —- the populations to either side of the cutoff consist entirely of intensive-margin students and are therefore directly comparable using a standard fuzzy RD design. The estimand of interest is:

$$\beta_{IV} = \frac{\lim_{r\uparrow k^*} \mathbb{E}[Y \mid k = r] - \lim_{r\downarrow k^*} \mathbb{E}[Y \mid k = r]}{\lim_{r\uparrow k^*} \mathbb{E}[D \mid k = r] - \lim_{r\downarrow k^*} \mathbb{E}[D \mid k = r]}$$

where Y is an outcome such as graduation or earnings, and where G is an indicator for the treatment in question.[27] I follow standard practice to nonparametrically estimate β_{IV} by taking the ratio of two separate RD estimators:

$$\hat{\beta}_{IV} = \frac{\hat{\mu}_{Y-} - \hat{\mu}_{Y+}}{\hat{\mu}_{D-} - \hat{\mu}_{D+}}$$

where $(\hat{\mu}_{Y-}, \hat{\mu}_{Y+}, \hat{\mu}_{D-}, \hat{\mu}_{D+})$ are local linear regression estimators of the conditional expectations of Y and G given k, estimated separately approaching the cutoff from the left and the right.[28]

[27] In the case of the TEXAS Grant, the treatment an indicator for whether the student receives a TEXAS Grant in their first year of college. In the case of the auto zero EFC cutoff, the treatment is an indicator for whether the student receives a zero EFC.

[28] Formally, for $W \in \{Y, D\}$, the estimators are given by $\hat{\mu}_{Y-} = \alpha_{W-}(k^*, h)$ and $\hat{\mu}_{Y+} = \alpha_{W+}(k^*, h)$, where $\alpha_{W-}(k^*, h)$ and $\alpha_{W-}(k^*, h)$ come from the solutions to the following kernel-weighted least squares problems:

$$\begin{pmatrix} \alpha_{W-}(k,h) \\ \beta_{W-}(k,h) \end{pmatrix} = \arg\min_{\alpha,\beta} \sum_{i=1}^{N} \mathbf{1}\{k \leq k^*\} \left(W_i - \alpha - \beta(k_i - k^*)\right)^2 K_h(k_i - k)$$

$$\begin{pmatrix} \alpha_{W+}(k,h) \\ \beta_{W+}(k,h) \end{pmatrix} = \arg\min_{\alpha,\beta} \sum_{i=1}^{N} \mathbf{1}\{k > k^*\} \left(W_i - \alpha - \beta(k_i - k^*)\right)^2 K_h(k_i - k)$$

I also report results from specifications that adjust the standard IV estimator by adding an additively separable linear term in a vector of covariates $\boldsymbol{X}$ (2019). Because of the locally randomized assignment of grant awards at the cutoff, potential treatment assignment and potential outcomes should be independent of covariates in a neighborhood of the threshold; moreover, there is no significant jump at the cutoff in any of these covariates at either the TEXAS Grant or the Auto-Zero threshold in four-year schools. Reassuringly, inclusion of the covariates does not significantly change the IV point estimate for any outcome I consider. However, including covariates considerably improves the precision of the estimates. I implement the estimators by choosing the MSE–optimal bandwidth separately for each outcome according to Calonico, Cattaneo and Farrell (2019). Practically, the optimal bandwidth is almost always between \$750 and \$900 of EFC for models with covariates and between \$900 and \$1200 of EFC for models without covariates.

5 The Impacts of the TEXAS Grant on Student Outcomes

5.1 First Stage: The TEXAS Grant Cutoff and Grant Aid Awards

I first show that crossing the EFC threshold results in a substantial jump in the probability of receiving a TEXAS Grant. Panel (a) of Figure 2 shows the increase in TEXAS Grants at the cutoff among students entering public four-year Texas colleges during my sample period. The running variable is students' EFC in the year of college entry, which determines whether students are eligible for TEXAS Grants in their first year of college.

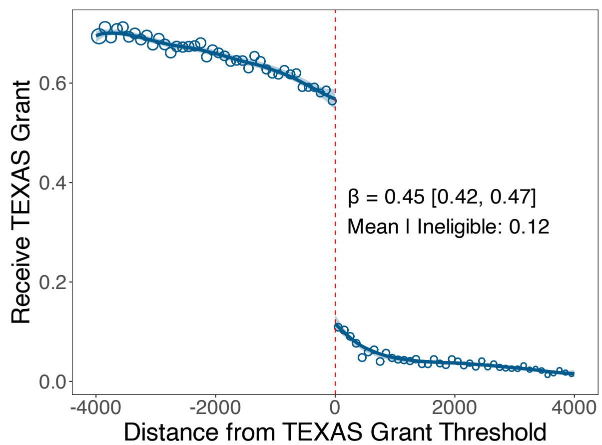
The figure shows clear evidence of a large jump in initial-year TEXAS Grants at the cutoff. The proportion of students receiving TEXAS Grants in their first year of college jumps by 45 percentage points, from 11.6 percent of barely-ineligible students receiving grants on the right of the cutoff to 56.8 percent of barely-eligible students receiving grants on the left. Importantly, compliance with the grant assignment rule is imperfect: roughly half of barely-eligible students do not receive TEXAS Grants despite being on the eligible side of the threshold, and 12 percent of financially-ineligible students receive TEXAS Grants anyway. There are two main reasons for this imperfect treatment assignment. First, the grants are rationed; there are not sufficient funds available to award TEXAS Grants to all students who are eligible on the basis of their EFC.[29] Second, schools are afforded substantial discretion in determining which students are awarded grants, and are allowed to award TEXAS Grants even to students whose EFCs fall above the grant assignment threshold. In practice, Figure 2a shows that this discretion is used relatively sparingly, and crossing the threshold still confers a large increase in the probability of receiving a TEXAS Grant.

where (W_i, k_i) denote data on W and k for observation i, and where K_h is a triangular kernel with bandwidth h.

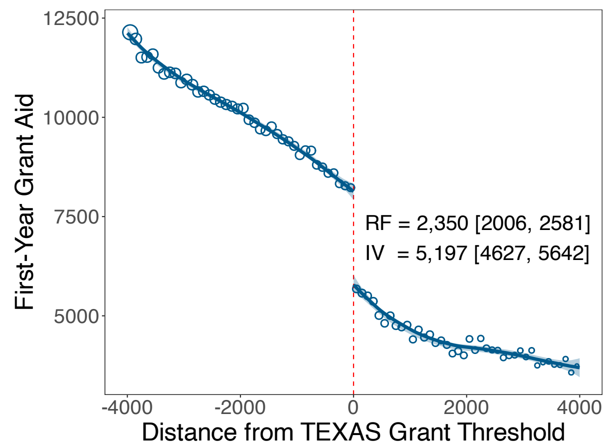
[29] This rationing stems in part from the fact that schools are required to cover the entirety of TEXAS Grant recipients' demonstrated financial need, supplanting TEXAS Grant funds with institutional or other sources of grant aid until students' full financial need is met.

Figure 2: First Stage at the TEXAS Grant Threshold

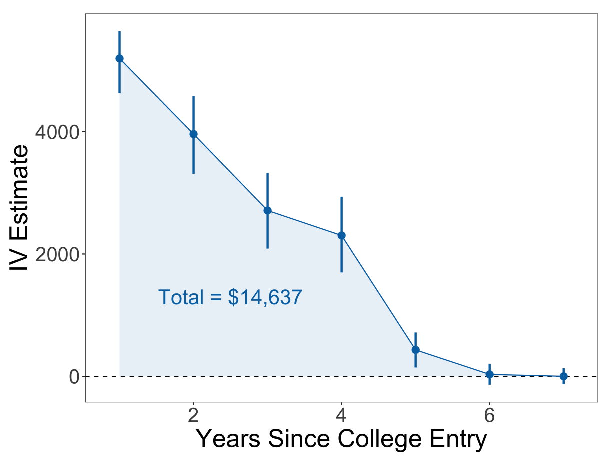
(a) Receive TEXAS Grant



(b) First-Year Grant Award



(c) Persistence of Grant Aid



Notes: This figure plots the bivariate relationship between students' grant aid awards and their Expected Family Contribution (EFC) in a window of the TEXAS Grant Cutoff. Panel (a) shows the instrumental variables first stage. The outcome is an indicator for whether the student receives a TEXAS Grant in their initial year of college. The plot annotation reports the estimated jump in initial-year TEXAS Grants at the threshold, together with the mean fraction of students on the barely-ineligible (right) side of the cutoff who do not receive grants. Panel (b) plots the relationship between total first-year grant aid awards, defined as the sum of all grant aid sources, and students' entry-year EFCs. The plot annotation reports reduced-form and IV estimates of the change in grant aid at the cutoff. Panel (c) reports IV estimates from year-by-year regressions, where the outcome is the total grant aid award received by students in each year following college entry. The error bars reflect 95 percent bias-corrected confidence intervals (Calonico et al. 2014), clustered at the institution-by-entry-cohort level. Note that because students who leave college are assigned grants of zero, the estimates in Panel (c) summarize the TEXAS Grant's impacts on both grant aid and college reenrollment.

Figures 2b and 2c show that receiving a TEXAS Grant corresponds to a substantial and persistent increase in students' grant aid packages. Figure 2b plots the bivariate relationship between students' first-year grant aid awards and their EFC in the first year of college. Students on the barely-eligible side of the threshold receive $\$2,350$ more in grant aid, on average, than barely-ineligible students. The corresponding IV estimate for the increase in first-year grant aid among compliers is $\$5,197$. Figure 2c shows that this increase in grant aid is persistent, plotting IV estimates of the increase in grant aid for compliers separately in each year after college. In constructing the figure, I assign $0 in grant awards to students who exit the college sample; as a result, a note of caution is warranted in interpreting Figure 2c, because the plotted estimates incorporate any causal effect of the TEXAS Grant on students' probability of remaining in college. Nevertheless, the impact of receiving an initial-year grant on average future aid receipts is significant; for compliers, an initial-year TEXAS Grant increases average aid received over the next 6 years of college by $\$14,637$ on average.

5.2 Impacts on Persistence and Graduation

Figure 3 reports estimates of the impact of the TEXAS Grant on students' persistence in college and graduation probability. Figure 3a displays the reduced-form graphical relationship between students' Expected Family Contribution and the probability that students re-enroll in college for a second year. The figure shows that despite the large increase in average grant aid awards at the cutoff, there is no statistically significant change in the probability of second-year reenrollment at the threshold. The reduced-form RD estimate at the cutoff is a relatively precise zero (confidence interval $[-0.02, 0.02]$), and I am able to rule out that the overall rate of reenrollment changes by more than 2 percentage points at the cutoff.

Figure 3b shows results for BA completion, plotting the graphical relationship between students' EFC and the probability of graduating within six years of college entry. I do not find evidence that the TEXAS Grant singificantly increases graduation rates at the cutoff; the reduced-form point estimate is positive at 1 percentage point, but the confidence interval rules out increases in graduation rates of more than 3 percentage points or declines of more than 2 percentage points. Importantly, this null result is not because baseline graduation rates without the TEXAS Grant are too high for any program to improve them; the figure shows that just over half of barely-ineligible students graduate. The results therefore suggest that for middle-income students at the margin of receiving a TEXAS Grant, additional grant aid is not a key determining factor in whether students complete college.

Figures 3c and 3d show examine effects on persistence and graduation at different time horizons. The figure plots instrumental variables estimates of the impact of receiving a first-year TEXAS Grant on the probability of reenrollment and graduation, with the corresponding point estimates reported in Appendix Table B6. Though the point estimates are small and positive in most years, they are not statistically significant at any time horizon. I conclude that the increase in financial support associated with crossing the TEXAS Grant threshold does not cause a significant increase in students' probability of reenrolling or graduating from college.

In addition to these results, I also find no evidence that the TEXAS Grant has a significant effect on shorter-run measures of student success in college. Appendix Figure A7 examines the impact of the TEXAS Grant on complier students' course completion. The results show that there is no significant effect of the TEXAS Grant on the number of credits completed by complier students, indicating that additional grant aid does not induce students awarded TEXAS Grants to complete significantly greater numbers of credit hours.

Taken together, the results suggest that the increase in grant aid at the threshold has no impact on the average postsecondary outcomes of TEXAS Grant recipients. Importantly, this result applies only to *marginal* recipients of the TEXAS grant who receive funds as a causal result of crossing the grant's eligibility threshold, and may not be representative of the TEXAS Grant's impacts on lower-income students with EFCs well below the grant cutoff. In Section 5.3, I investigate the possibility that reallocations of grant aid to lower-income students would produce gains in overall student outcomes through a comparison to the Auto-Zero EFC threshold in four-year schools.

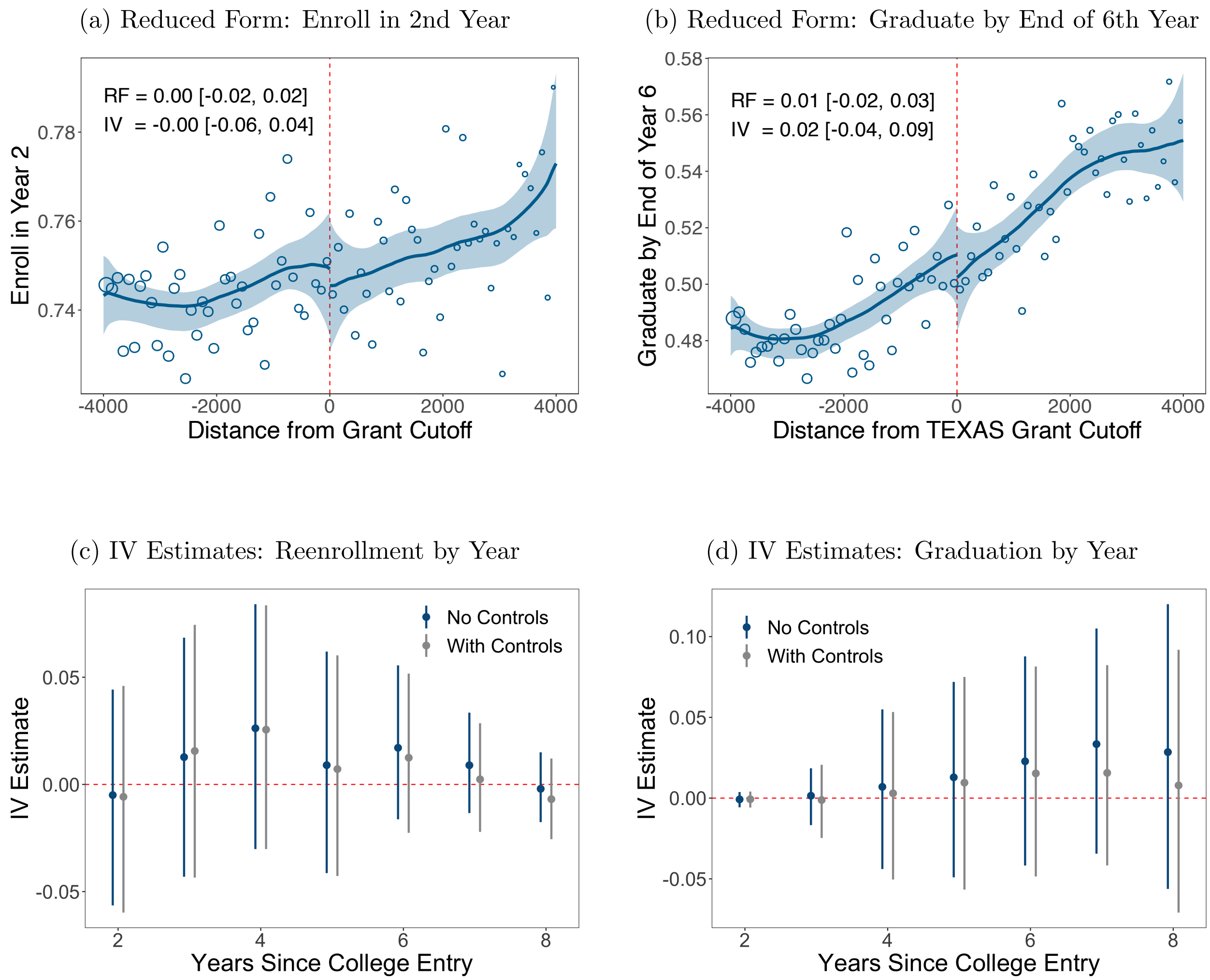
5.2.1 Impacts on Loans and Earnings

Figure 4 reports estimates of the TEXAS Grant's impacts on students' labor market earnings and student loans. The increase in TEXAS Grants at the cutoff corresponds with a sharp decrease in student loans and earnings during college, as shown in Figures 4a and 4b. Students on the barely-eligible side of the TEXAS Grant threshold take out more than $\$1,000$ less in average first-year loans than students on the barely-eligible side. Scaling this reduced-form estimate by the first stage implies that TEXAS Grant compliers at the threshold reduce their first-year loans by over \$2100 in response to receiving a TEXAS Grant award. Panel (b) shows evidence of a similar effect on complier students' earnings during college. Students on the barely-eligible side of the cutoff reduce their earnings by \$300, on average, relative to students on the barely-ineligible side. The corresponding IV estimates indicate a \$700 decrease in first-year earnings among grant compliers.

Panels (c) and (d) investigate the persistence of these effects, plotting IV estimates of the TEXAS Grant's impact on compliers' yearly borrowing and labor market earnings.[30] The grant's impact on the student loans and earnings of compliers is persistent, with point estimates remaining negative and significant through the second year of college. This persistence reflects the fact that TEXAS Grant recipients can renew their grants for additional years, subject to satisfactory

[30] To make the figures comparable, I code students who exit college as taking out zero loans; the population for each time horizon is the full set of students for whom I can observe the given year of outcomes.

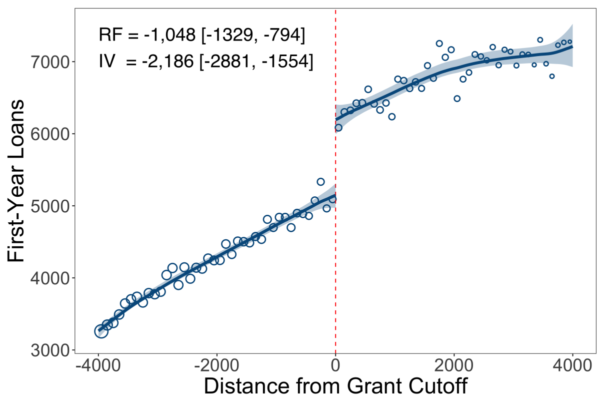
Figure 3: Effects of the TEXAS Grant on Persistence and Graduation



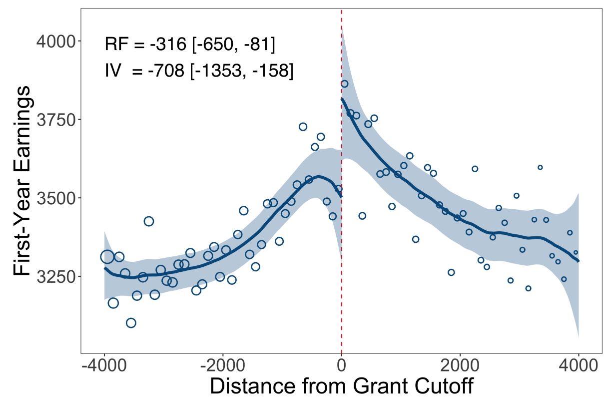
Notes: This figure shows estimates of the impacts of the TEXAS Grant on students' reenrollment in college and probability of graduation. Panel (a) shows the reduced-form relationship between students' Expected Family Contribution (EFC) and the probability of completing a BA degree at their initially-enrolled institution within 6 academic years of entry. I pool all entry cohorts and normalize the running variable so that zero equals the TEXAS Grant's EFC cutoff. The points show unconditional means by $100-wide EFC bins, and the lines to either side of the cutoff are local-linear fits with bias-corrected confidence intervals following Calonico et al. (2014). The annotation reports reduced-form and IV estimates of the discontinuity at the cutoff. Panel (b) shows IV estimates of the discontinuity at the cutoff at several estimation horizons. Results through 4 years are for the 2007-2017 entry cohorts The last entry year for results at 5, 6, 7, and 8 years is 2016, 2015, 2014, and 2013, respectively. The error bars reflect 95 percent bias-corrected confidence intervals, clustered at the institution-by-entry-cohort level.

Figure 4: Effects of the TEXAS Grant on Loans and Earnings

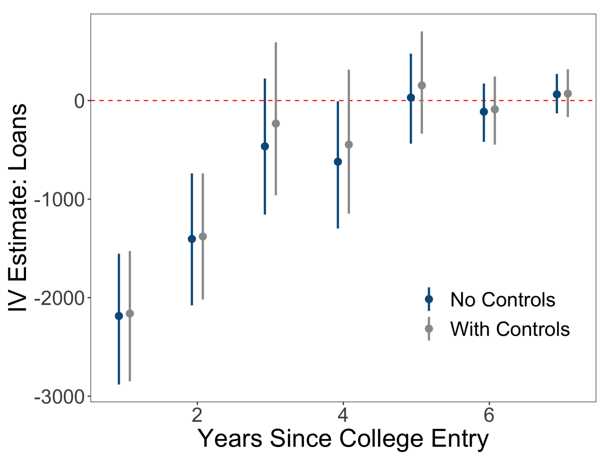
(a) Reduced Form: Loans in Year 1



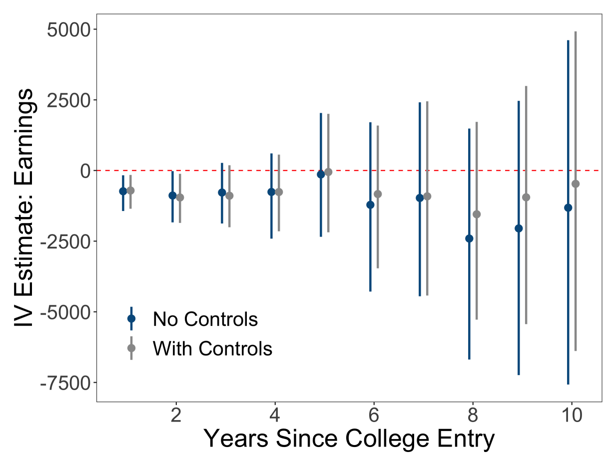
(b) Reduced Form: Earnings in Year 1



(c) IV Estimates: Loans by Year



(d) IV Estimates: Earnings by Year



Notes: This figure shows estimates of the impacts of the TEXAS Grant on enrollment persistence Panel (a) shows the reduced-form relationship between students' Expected Family Contribution (EFC) and the probability of completing a BA degree at their initially-enrolled institution within 6 academic years of entry. The points show unconditional means by $100-wide EFC bins, and the lines to either side of the cutoff are local-linear fits with bias-corrected confidence intervals following Calonico et al. (2014). The annotation reports reduced-form and IV estimates of the discontinuity at the cutoff. Panel (b) shows IV estimates of the discontinuity at the cutoff at several estimation horizons. Results through 4 years are for the 2007-2017 entry cohorts The last entry year for results at 5, 6, 7, and 8 years is 2016, 2015, 2014, and 2013, respectively.

academic progress requirements. Estimates of loan impacts fall towards zero over time, in part reflecting the fact that many students do not remain in college. Estimates on earnings appear more persistently negative, although they become imprecise in later years.

5.3 Comparison: Impacts of Automatic Zero EFC Eligibility

The results in Section 5 suggest that for middle-income students at the margin of receiving a TEXAS Grant, additional grant aid is not an influential factor in determining post secondary outcomes. In this section, I compare these null effects against estimates of the impact of additional financial support on student outcomes for four-year college students at the Automatic Zero EFC Cutoff.

5.3.1 First Stage: The Auto-Zero EFC Cutoff and Grant Aid

Figure 5 describes the impact of crossing the automatic zero EFC threshold on the grant aid awards of four-year college students. Panel (a) shows how crossing the Automatic Zero EFC threshold affects whether students receive a zero EFC in their first year of college, which I use as the definition of binary treatment to analyze the Automatic Zero EFC cutoff. The x axis plots students' family Adjusted Gross Income in \$1,000 bins, and the y axis plots the proportion of four-year college students whose EFC is zero in the entering year. The figure shows that the probability of being assigned a \$0 EFC jumps by 49 percentage points at the threshold. Because other inputs into the EFC formula affect whether a student receives a \$0 EFC, a substantial number of students receive zero EFCs despite being on the barely-ineligible side of the threshold; likewise, a small proportion of students receive nonzero EFCs despite being on the barely-eligible side. Nevertheless, the jump in \$0 EFC assignments results in a substantial increase in students' first-year grant aid awards, as shown in Panel (b). Students on the barely-eligible side of the threshold receive an additional \$800 of grant aid, on average, than students on the barely-ineligible side. The instrumental-variables estimates show that compliers who receive zero EFCs as a result of crossing the auto-zero threshold receive an additional \$1,800 in aid, on average, as a result of crossing the threshold.

Figure 5c shows that this increase in grant aid is much smaller than the increase experienced by compliers at the TEXAS Grant threshold. The figure plots yearly RD coefficients of the increase in grant aid for compliers at the TEXAS Grant and Auto-Zero thresholds, plotted separately by grant program. In constructing the figure, I code students who drop out of college as receiving \$0 in grant aid to avoid bias due to differential attrition. Estimates in later years may therefore incorporate any potential impacts of grant aid on the probability that students remain in college. The point estimates at $x = 1$ show that the \$1,800 increase in first-year aid for Auto-Zero compliers is only a third of the \$5,200 increase experienced by compliers at the TEXAS Grant threshold. This difference persists over time. Summing the yearly estimates reveals that while TEXAS Grant compliers receive an additional \$14,600 in aid over the six years following college entry, the increase for compliers at the Automatic Zero EFC threshold is only a quarter as large, amounting to roughly \$3,400 in increased financial support.

Figure 5d shows analogous yearly IV estimates where the outcome is students' yearly loans. The point estimates at $x = 0$ show that Auto-Zero EFC compliers reduce their first-year loans by roughly half as much as TEXAS Grant compliers. Notably, the offsetting reduction in loans is roughly half of the first-year change in grant aid for both programs, suggesting similar elasticities of loan originations to grant-based financial aid for both populations. However, in contrast to the persistently negative impact of the TEXAS Grant on student loans, students who receive an Auto-Zero EFC as a result of crossing the threshold take out *greater* average loan burdens in later years of college, compared to untreated students. Such an effect would arise if treated students become more likely to remain in college than untreated students. I investigate this possibility in the next subsection by comparing the impacts of the TEXAS Grant and the Zero-EFC policy on students' college persistence and graduation probability.

5.3.2 Grant Awards and Student Outcomes at the Auto-Zero EFC Threshold

Figure 6 plots IV estimates of the impact of financial support on students' academic outcomes and later-life earnings at the TEXAS Grant and Auto-Zero thresholds. To make the estimates comparable in terms of aid's impact on outcomes, I estimate models where the first stage is measured in \$1,000s of grant aid rather than a binary indicator for treatment.[31]

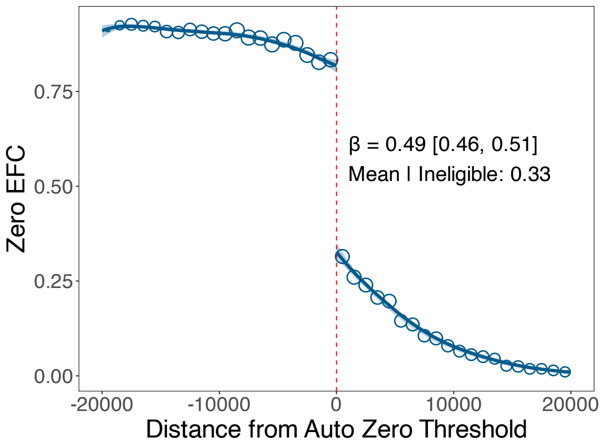
Comparing the estimates reveals that the impact of additional grant aid on student outcomes is starkly different across the two thresholds. Figure 6 shows impacts for college persistence, measured as an indicator for reenrollment. While an additional \$1,000 of grant aid at the TEXAS Grant cutoff has a precise null effect on students' probability of college reenrollment, an equivalent increase for poorer students at the margin of the Auto-Zero EFC cutoff increases the probability of reenrollment by 3.4 percentage points in year 3 and 3.9 percentage points in year 4. The point estimate for reenrollment in year 1 is positive, at 1 percentage point, but is statistically imprecise.

This increase in reenrollment probabilities corresponds to large increases in graduation for compliers at the margin of receiving an automatic zero EFC. Figure 6b shows that while TEXAS Grant compliers are not significantly more likely to complete a B.A. at any time horizon, an additional \$1,000 of grant aid for compliers at the Auto-Zero cutoff increases graduation rates by

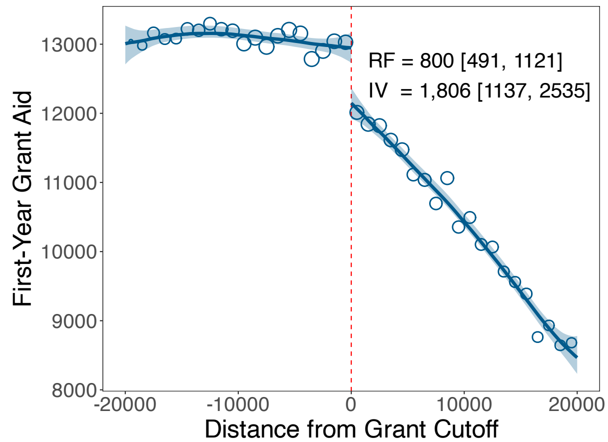
[31] This normalization effectively assumes that the TEXAS Grant and the Auto-Zero EFC cutoff affect students' grant aid awards only through the channel of increased financial support. This assumption would be violated if qualifying for either grant program gave students access to nonfinancial resources that potentially impacted their college performance. This assumption seems reasonable in the case of the Automatic Zero EFC threshold, as students' EFCs are calculated by the federal government (not by colleges) and are primarily used by colleges as an input into financial aid awards. My conversations with financial aid administrators suggest that the TEXAS Grant also acts solely as a subsidy that reduces the price of college. Note also that a similar assumption is inherent in MVPF calculations which scale the benefits of grant programs by their fiscal net costs.

Figure 5: First Stage at the Auto-Zero Threshold in Four-Year Schools

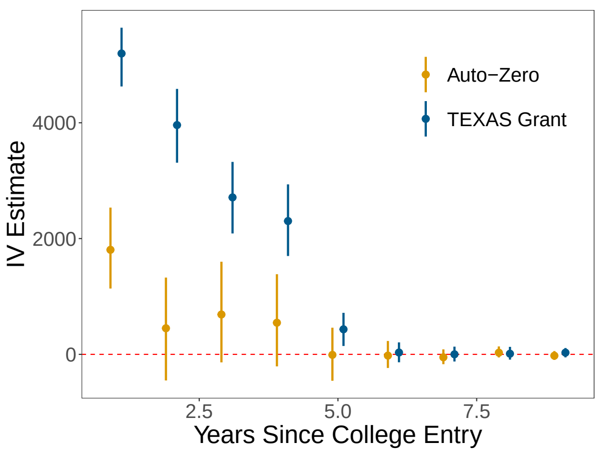
(a) Zero EFC at Auto-Zero Cutoff



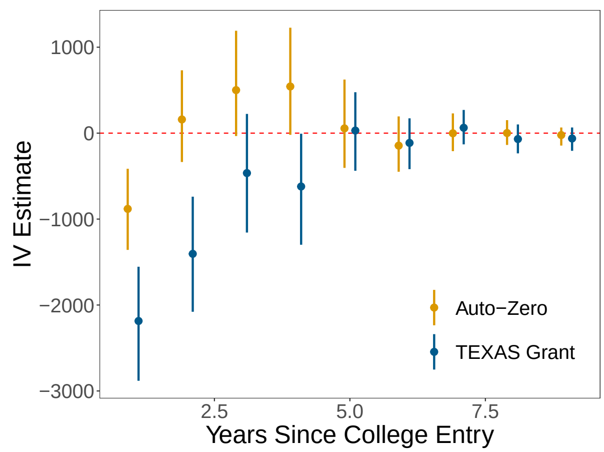
(b) First-Year Grant Aid at Auto-Zero Cutoff



(c) IV Estimates: Yearly Grants for Compliers



(d) IV Estimates: Yearly Loans for Compliers



Notes: This figure summarizes the increase in grant aid stemming from crossing the Automatic Zero EFC threshold, and compares this increase to the change in grant aid at the TEXAS Grant cutoff. Panel (a) reports the fraction of four-year college students who have a zero EFC by their family adjusted gross income (AGI). The plot pools all years of data for the 2008-2012 entering cohorts, and the x axis is normalized within each year so that zero equals the federal Auto-Zero EFC threshold. Panel (b) is similar, but instead plots students' first-year grant aid awards by their incoming AGI. I plot unrestricted means of both outcomes in \$1,000 bins and include estimated local linear regression lines and 95% bias-corrected confidence intervals on each side of the cutoff. Panels (c) and (d) plot instrumental variables estimates of the effect of receiving a TEXAS Grant or an initial-year Zero EFC on students' yearly grant aid (Panel c) and loans (Panel d). The first stage for each program is a binary indicator for whether the student is "treated" by receiving an initial-year TEXAS Grant or an initial-year zero EFC, respectively. In constructing the figures, I define grants and loans as \$0 for students who do not appear in the college sample to avoid attrition bias; note, however, that these grant and loan impacts therefore incorporate the effects of each grant program on students' college persistence. The error bars reflect 95 percent bias-corrected confidence intervals estimated using the method of (Calonico et al. 2014) and clustered at the institution-by-entry-cohort level.

2.9 percentage points by the end of the fourth year after college entry, 3.8 percentage points by the end of the fifth year, and 4.8 percentage points by the end of the sixth. These effects are large; for example, the six-year graduation increase of 4.8 percentage points is a 12 percent increase over the mean 6-year graduation rate of 41% on the barely-ineligible side of the cutoff.

Finally, Figure 6c provides suggestive evidence that these increases in persistence and graduation probability affect students' earnings trajectories. The figure plots the IV impact of an additional $1,000 of financial support on students' earnings, estimated separately by year. The figure has two main takeaways. I find that the increase in aid for compliers at the Auto-Zero threshold lowers labor market earnings during college. Notably, the elasticity of this earnings response to first-year grant aid is larger for Auto-Zero compliers than it is for compliers at the margin at the TEXAS Grant cutoff, although the difference between these coefficients is not statistically significant. Second, I find evidence that this initial decline in earnings is compensated by increased earnings starting six years after college entry. While the coefficients are not statistically significant, they suggest that the impact of financial support on the graduation rates of Auto-Zero compliers spills over into their labor market earnings after college.

Overall, these results align with the findings of Denning et al. (2019) (DMT), who use the same data source to study the impact of crossing the Auto-Zero threshold on the outcomes of four-year college students in Texas. However, my results differ from theirs in some notable respects. First, I find that Auto-Zero EFC compliers' earnings fall during college, while DMT do not find any evidence of this temporary earnings decline. Second, I find slight differences in the estimated impacts on persistence and graduation, finding (in contrast to DMT) that students are more likely to persist until the fourth year of college, and finding slightly smaller impacts than theirs on the probability of graduating within 5 years. These differences may stem from differences in sample composition; I use one additional year of data (2012) compared to DMT, and the starting point for constructing my sample is the set of all recent Texas public high school graduates rather than the full set of all first-time-in-college dependent students (including non-Texas residents). Overall, however, I find robust positive impacts of crossing the Auto-Zero threshold on the outcomes of compliers in four-year schools.

Additionally, my results may shed insight on the mechanism through which financial aid affects the outcomes of low-income students. A recent study by Eng and Matsudaira (2021) re-visits the Auto-Zero EFC threshold using the universe of federal aid recipients (rather than the selected sample of Texas students), finding much smaller effects of crossing the Automatic Zero EFC cutoff on students' probability of graduating from their initially-enrolled institution. For example, their IV estimates indicate that an additional $1,000 of aid at the threshold corresponds to a statistically imprecise 0.8 percentage point increase in the probability of BA degree completion within 6 years,[32] in

[32] See Eng and Matsudaira (2021), Table 3.

contrast to my and DMT's estimate of roughly 5 percentage points. In considering the discrepancy between these results and the findings of DMT, they conjecture that qualifying for an Auto-Zero EFC may cause large outcome gains for Texas students because a zero EFC also increases the probability of receiving a TEXAS Grant, which — in contrast to the federal Pell grant — provides a guarantee of financial support covering the entirety of tuition and fees for up to five years after students' initial college enrollment. The Automatic Zero EFC policy may therefore be especially effective in improving the outcomes of TEXAS students because it coresponds to the *persistent* increase in grant aid documented in Figure 5c. This persistent increase is likely entirely due to the impact of the TEXAS Grant, because an initial-year zero EFC is not automatically renewed in future years.[33] In contrast, an initial-year TEXAS Grant is renewable regardless of most changes in students' financial situations, yielding a persistent increase in grant aid over time. Overall, the results suggest that the renewal-guarantee aspect of the TEXAS Grant, and not the Pell Grant alone, is important in shaping the outcomes of students near the Auto-Zero EFC threshold.

6 Mechanisms: Who Gets a TEXAS Grant?

Because schools have discretion in allocating TEXAS Grants, the observed pattern of grant awards at the threshold provides a revealed measure of their preferences. How do schools use their discretion to award TEXAS Grants, and how do the characteristics of students that receive TEXAS Grants differ from the characteristics of students that do not? This section analyzes the characteristics of students who receive grants at the threshold, finding that schools are substantially more likely to give grants to two subpopulations of students: those who are likely to graduate, and those who would receive substantial amounts of grant-based financial aid even if not awarded TEXAS Grants.

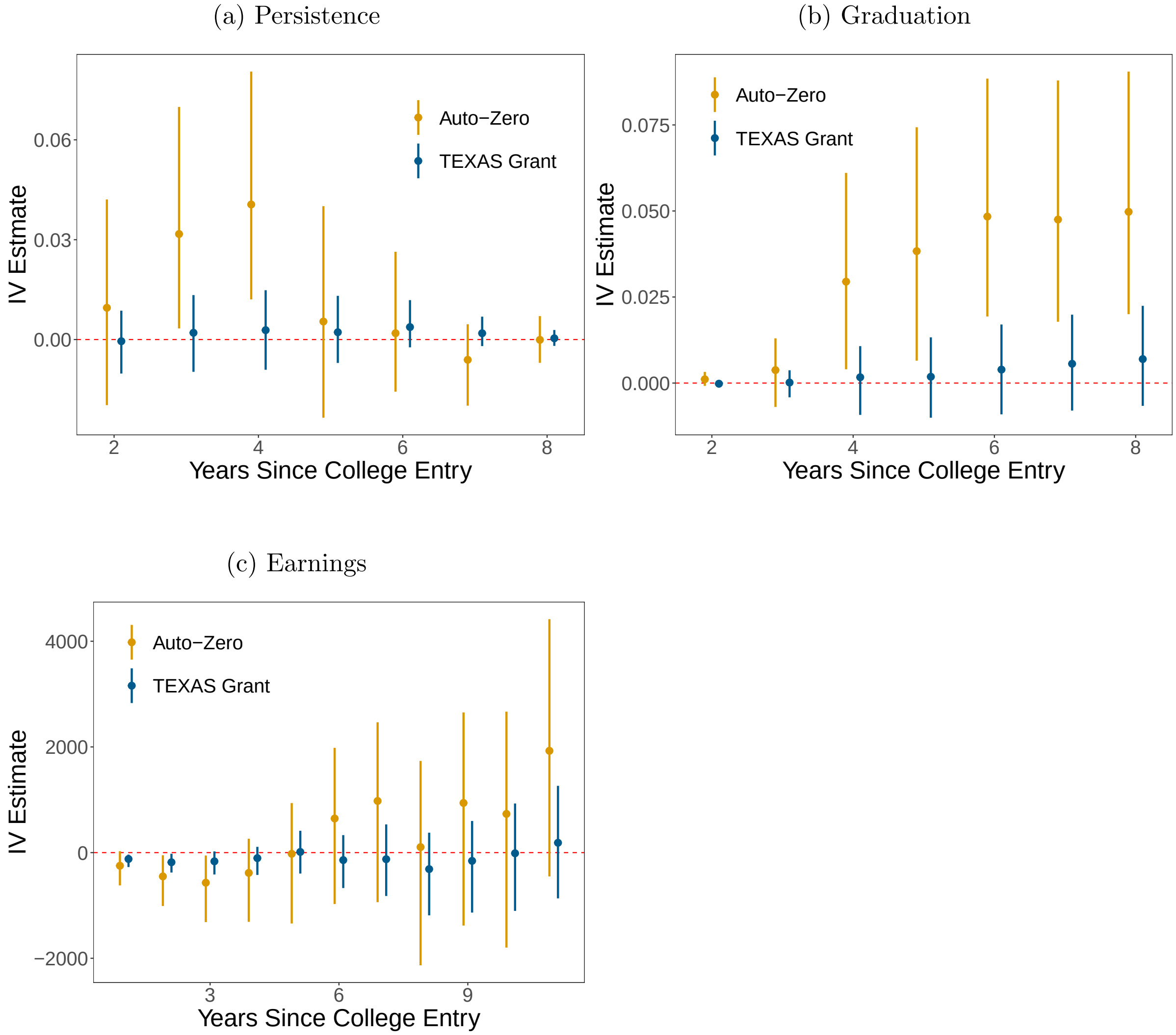
6.1 Measuring complier characteristics

Abadie (2002) shows that for any predetermined student characteristic X, the characteristics of compliers at the threshold are identified by:

$$\frac{\lim_{k\uparrow k^*} \mathbb{E}[X(1-D) \mid k] - \lim_{k\downarrow k^*} \mathbb{E}[X(1-D) \mid k]}{\lim_{k\uparrow k^*} \mathbb{E}[(1-D) \mid k] - \lim_{k\downarrow k^*} \mathbb{E}[(1-D) \mid k]} = \mathbb{E}[X \mid D_1 > D_0, k = k^*] \tag{5}$$

[33] In fact, if students' yearly family AGI incorporates an element of random noise, then students who barely receive a zero EFC in the initial year of college have probability zero of being on the barely-eligible side of the threshold in any future year. This is likely to especially hold in the samples used in my paper and DMT, which remove students who report AGI multiples of $1,000 due to bias concerns due to bunching.

Figure 6: IV Effects of $1,000 of Grant Aid at the TEXAS Grant and Auto Zero Cutoffs in Four-Year Schools



Notes: This figure compares the impact of an additional $1,000 of grant aid on two groups of students: middle-income students at the TEXAS Grant threshold, and lower-income students at the Auto-Zero AGI threshold. Each point estimate and 95% confidence interval reports the results of a separate instrumental variables regression, where the first stage is first-year grant aid measured in thousands of dollars. In all figures, the x axis is the number of years since the student entered college. The outcomes in Panels (a), (b), and (c), respectively, are graduation, persistence, and earnings at the given time horizon. The results for the TEXAS Grant are estimated on the 2007-2017 entering cohorts, and the results for the Auto-Zero cutoff are estimated on the 2008-2012 entering cohorts.

I compare the compliers' characteristics to the characteristics of "always-takers," who receive grants despite being on the barely-ineligible side of the threshold:

$$\lim_{k \downarrow k^*} \mathbb{E}[X \mid D = 1, k] = \mathbb{E}[X \mid D_0 = D_1 = 1, k = k^*] \tag{6}$$

as well as "never-takers," who do not receive grants despite being financially eligible:

$$\lim_{k \uparrow k^*} \mathbb{E}[X \mid D = 0, k] = \mathbb{E}[X \mid D_0 = D_1 = 0, k = k^*] \tag{7}$$

Estimating equations (5) through (7) for several useful comparisons. The relative characteristics of compliers and never-takers reveal how schools choose a set of TEXAS Grant awardees from the set of students without TEXAS Grants on the barely-eligible side of the threshold. Similarly, comparing always-takers to compliers reveals whether students who receive an exception to the threshold rule differ systematically from students who receive grants only if they cross the threshold.

6.2 Targeting of the TEXAS Grant

Selection on Outcome Levels Section 5 showed that the TEXAS Grant does not meaningfully impact compliers' college persistence, graduation probability, or post-college earnings trajectories. Because the TEXAS Grant does not change compliers' outcomes, differences in observed outcomes between compliers and never-takers operate entirely through selection rather than through the grant's treatment effect on compliers.[34] Moreover, schools' grant award choices are the key force that drives this selection. As a result, comparing compliers' outcomes with the outcomes of never-takers yields a measure of how schools target the TEXAS Grant on outcome levels. Do schools target TEXAS Grants to students with higher ex-ante ability?

Figure 7 provides evidence that schools target TEXAS Grants at the threshold to high-ability students. The figure separately plots the mean outcomes of compliers and never-takers at the TEXAS Grant cutoff, estimated using Equations (5) and (7). Figure 7a shows results for BA completion, showing that compliers complete BA degrees at far higher rates than never-takers. For example, six years after entering college, compliers have mean graduation rates that are twelve percentage points higher than those of never-takers. Figure 7b shows evidence of a similar pattern in students' post-college earnings. The mean earnings of compliers, measured six to eight years after college entry, are several thousand dollars higher than those of never-takers who are not awarded TEXAS Grants. Taken together, the results imply that schools target TEXAS Grants

[34] Importantly, the same is not necessarily true for always-takers who receive grants despite being on the financially ineligible side of the cutoff. It is possible that the TEXAS Grant affects always-takers' outcomes; however, because I do not observe always-takers in their no-grant counterfactual, it is not possible without additional assumptions to gauge whether the grant impacts always-takers' outcomes. For that reason, I limit outcome-based comparisons in this section to compliers and never-takers.

at the threshold to students who would graduate at higher rates (and have higher post-college labor-market earnings) than students not selected to receive grants, *even if* these grant recipients did not receive a TEXAS Grant.

Figure 7c confirms these patterns, documenting selection at the threshold by students' 10th grade test scores. Compliers' average test scores are more than three percentiles higher, on average, than the high-school test scores of never-taker students. The test scores of always-takers, who receive grants despite being ineligible for the TEXAS Grant on the basis of their EFC, fall between the complier and never-taker estimates. The results indicate that schools target on preexisting academic ability, as measured by test scores, within the set of students for whom crossing the threshold is key to receiving a grant aid award.[35]

Selection on Preexisting Aid Because schools are required to supplement TEXAS Grants with institutional aid to fully cover grant recipients' tuition and fees, they may face an incentive to disproportionately award TEXAS Grants to students with high preexisting aid awards.[36] Figure 7d examines this possibility by comparing the mean untreated potential aid awards of TEXAS Grant compliers (Equation 5) against the aid awards of never-takers at the grant threshold (Equation 7). The results show that schools disproportionately award TEXAS Grants to students with high preexisting levels of grant-based financial aid. Even if TEXAS Grant compliers at the threshold did not receive TEXAS Grants, they would still enter college with roughly \$820 more in first-year aid than never-taker students. For students at the margin of qualifying for a TEXAS Grant, schools disproportionately target aid towards *ex-ante* well-funded students.

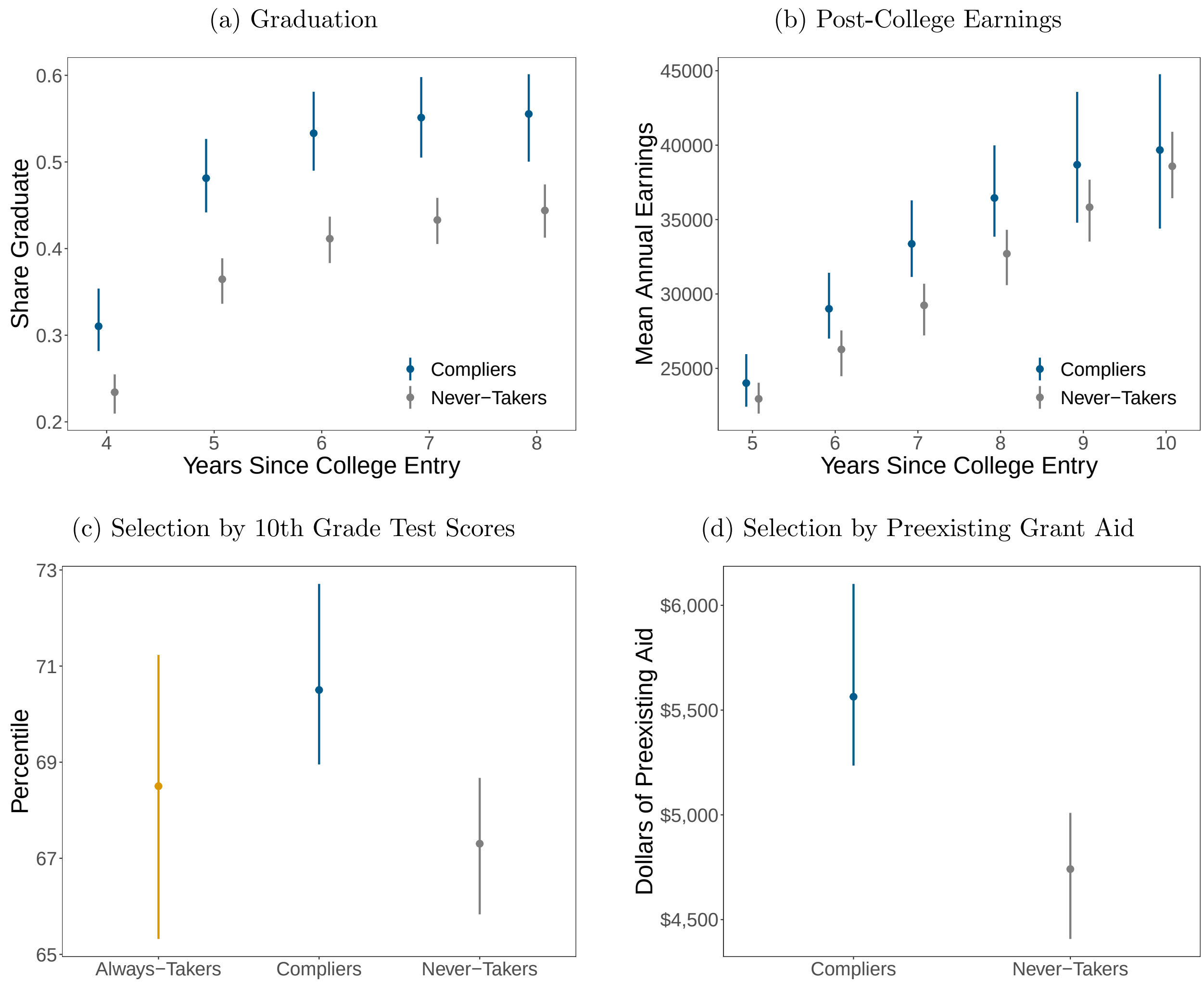
7 Nonparametric Bounds on Grant Treatment Effects

To this point, the paper has shown that the impact of grant aid on student outcomes differs substantially across the TEXAS Grant and Auto-Zero EFC cutoffs for students in four-year schools. The remaining task is to estimate the impact of grant aid on two-year students at the Auto-Zero cutoff. However, this task is made more difficult by the existence of a significant enrollment effect

[35] One possibility is that these selection patterns are due to merit standards, which since the 2014 academic year have been part of the TEXAS Grant's design. Although only four out of the eleven cohorts in my sample enter college after this policy change, it is in theory possible that the observed disproportionate allocation of grants towards high-achieving students is a reflection of state-administered merit criteria rather than a consequence of schools' preferences in distributing grant aid awards. Appendix Figure A9 provides evidence against this possibility, showing that school's revealed preference for awarding TEXAS Grants to high-achieving students predates the 2014 introduction of merit standards to the TEXAS Grant, and for some outcomes (e.g., five-year graduation rates) appears to favor ex-ante high-achieving students even more strongly.

[36] The idea that this incentive shapes schools' allocation decisions has been widely conjectured in descriptions of the TEXAS Grant (Andrews and Stange 2019; Denning et al. 2019), but to my knowledge has not been empirically documented.

Figure 7: Selection at the TEXAS Grant Threshold



Notes: This figure shows estimates of graduation probabilities (Panel (a)) and post-college earnings (Panel (b)) by the complier group of students at the TEXAS Grant cutoff. "Always-takers" are students who receive TEXAS Grants despite being on the barely-ineligible side of the threshold. "Compliers" are students who receive TEXAS Grants as a causal result of crossing the threshold. "Never-takers" are students who do not receive TEXAS Grants despite being on the barely-eligible side of the threshold. Panel (c) shows the mean tenth grade test score percentile of compliers, always-takers, and never-takers at the threshold. Panel (d) shows the mean baseline grant awards of compliers and never-takers at the TEXAS Grant threshold, defined as the amount of grant aid these students would receive if not awarded a TEXAS Grant. Counterfactual grant amounts for always-takers cannot be estimated, because all always-takers at the threshold receive TEXAS Grants. The point estimates and confidence intervals at each time horizon are constructed by separately estimating equations (5) through (7) for each outcome and time horizon; see main text for details. The error bars reflect 95 percent bias-corrected confidence intervals estimated using the method of (Calonico et al. 2014).

at the grant assignment threshold, meaning that standard IV methods applied at the threshold do not recover a valid LATE for any population.

To deal with this challenge, this section develops a bounding approach extending Lee (2009) and Gerard et al. (2020). I extend their framework by showing how an additional mean-dominance assumption motivated by the relative pretreatment characteristics of extensive-margin and intensive-margin compliers yields substantially tighter bounds than the "no-assumptions" bounds derived by Gerard et al. (2020) and considered in prior work on grant impacts (Park and Scott-Clayton 2018).

7.1 Target Parameters

The goal of the approach is to separately bound treatment effects for two groups of compliers: intensive-margin compliers who would attend two-year colleges even without the additional grant aid at the cutoff, and extensive-margin compliers who are brought into two-year college attendance as a result of additional grant aid. For notational implicitly, I express the separate LATEs for these groups as:

$$\begin{aligned}
LATE_{IM} &\equiv \mathbb{E}[Y_{1,1} - Y_{1,0} \mid \underbrace{S_1 = S_0 = 1, D_1 > D_0}_{\text{IM Complier}}, k = k^*] \\
&= \bar{Y}_{1,1}^{\text{IM}} - \bar{Y}_{1,0}^{\text{IM}} \\
LATE_{EM} &\equiv \mathbb{E}[Y_{1,1} - Y_0 \mid \underbrace{S_1 > S_0, D_1 > D_0}_{\text{EM Complier}}, k = k^*] \\
&= \bar{Y}_{1,1}^{\text{EM}} - \bar{Y}_{0}^{\text{EM}}
\end{aligned}$$

A separate target parameter is the overall effect on enrolled compliers, which is the weighted average of $LATE_{IM}$ and $LATE_{EM}$:

$$\begin{aligned}
LATE^* &= (1-\omega)LATE_{IM} + \omega LATE_{EM} \\
&= \mathbb{E}[Y_{1,1} - Y_{S_0,0} \mid S_1 = 1, D_1 > D_0, k = k^*]
\end{aligned}$$

where the weight $\omega = P(S_1 > S_0 \mid S_1 = 1, D_1 > D_0)$ denotes the proportion of enrolled compliers who are extensive-margin compliers.

7.2 Identifying the Mass of Extensive-Margin Compliers

The first step in deriving the bounds is to identify the share of extensive-margin compliers among all compliers on the eligible side of the cutoff. Appendix D shows that the mass of extensive-margin compliers on the barely-eligible side of the cutoff is identified by the discontinuity in the density of the running variable at the threshold:

$$\pi_{\text{EM}}^{-} = P(D_1 > D_0, S_1 > S_0 \mid S = 1, k = k^{*-}) = \frac{f_1(k^{*-}) - f_1(k^{*+})}{f_1(k^{*-})} \tag{8}$$

where $f_1(k)$ is the density of the running variable in the selected sample of enrolled students, and where the notation k^{*-} and k^{*+} denotes limits approaching the cutoff from the left and right sides.

It is also possible to identify the *total* mass of intensive-margin and extensive-margin compliers on the barely-eligible side of the threshold. Appendix D shows that this total complier share is recovered by a modified first-stage equation which compares the share of students receiving grants across the cutoff, but includes an adjustment term to account for the grant's enrollment effect:

$$\begin{aligned} \pi_{\text{Complier}}^{-} &= P(S_1 = 1, D_1 > D_0 \mid S = 1, k = k^{*-}) \\ &= \bar{D}^{-} - \frac{f_1(k^{*-})}{f_1(k^{*+})}\bar{D}^{+} \end{aligned} \tag{9}$$

where $\bar{D}^{-} \equiv \lim_{r \uparrow k^*} \mathbb{E}[G \mid S = 1, k = r]$ and $\bar{D}^{+} \equiv \lim_{r \downarrow k^*} \mathbb{E}[G \mid S = 1, k = r]$ denote the proportion of students with zero EFCs on the left and right of the cutoff in the sample of enrolled students. Equation (9) is analogous to the standard Wald first stage, but contains an additional sample-size correction term $(f_1(k^{*-})/f_1(k^{*+}))$ which adjusts for the fact that the left side of the cutoff includes additional students (the extensive-margin compliers) who are not present to the right of the cutoff.

Because (8) identifies the mass of extensive margin compliers and (9) identifies the mass of all compliers, the share of extensive-margin compliers among all compliers is identified by Bayes' rule:

$$\begin{aligned} \omega \equiv \frac{\pi_{\text{EM}}^{-}}{\pi_{\text{Complier}}^{-}} &= P(S_1 > S_0 \mid D_1 > D_0, S_1 = 1, k = k^{*-}) \\ &= P(\text{EM Complier} \mid \text{Complier}, k = k^{*-}) \end{aligned} \tag{10}$$

The share ω measures the proportion of the treated complier population which enters the sample as a result of treatment.

Table 2 reports estimates of complier shares obtained by estimating Equations (8) through (10). The first row shows the estimate of π_{EM}^{-}, the overall mass of extensive-margin compliers on the barely-eligible side of cutoff. The point estimate indicates that 5.2 percent of the population on the barely-eligible side of the cutoff consists of extensive-margin compliers. The second row reports estimates from the modified first-stage regression in (9), estimating that intensive and extensive margin compliers make up 39.8 percent of the population on the barely-eligible side of the cutoff. Dividing these two numbers yields the estimate that 13.1% of the complier population consists of extensive-margin compliers. As I show in the next section, this share is a crucial input into

Table 2: Estimates of Complier Shares

Quantity	Symbol	Derivation	Estimate
EM Compliers on Eligible Side $P(D_1 > D_0, S_1 > S_0 \mid S_1 = 1, k = k^{*-})$	π^-_{EM}	Eqn. (8)	0.052 (0.028)
Compliers on Eligible Side $P(D_1 > D_0 \mid S_1 = 1, k = k^{*-})$	π^-_{Complier}	Eqn. (9)	0.398 (0.019)
EM Share of Compliers on Eligible Side $P(S_1 > S_0 \mid D_1 > D_0, S_1 = 1, k = k^{*-})$	ω	$\left(\pi^-_{\text{EM}}/\pi^-_{\text{Complier}}\right)$	0.131 (0.066)

Notes: This table reports estimates of complier shares for the population of students on the eligible side of the Automatic Zero EFC threshold in two-year community colleges. The first row reports the share of all students on the barely-eligible side of the cutoff who are extensive-margin compliers. The second row reports the share of all students on the barely-eligible side of the cutoff who are compliers (either extensive-margin or intensive-margin). The third row takes the ratio of these two terms to report the share of all compliers who are extensive-margin compliers. Bootstrap standard errors are in parentheses, where the optimal bandwidth for density estimation is chosen separately within each of 500 bootstrap replications.

constructing bounds on treatment effects.

7.3 Step 2: Identifying Potential Outcome Means

With the complier shares in hand, it is possible to start constructing bounds on treatment effects. I start by considering bounds on $LATE_{IM}$, the treatment effect on intensive-margin compliers who would attend college even without receiving grants. Equation (4) shows that the standard limiting Wald estimand at the cutoff fails to recover $LATE_{IM} = \bar{Y}^{\text{IM}}_{1,1} - \bar{Y}^{\text{IM}}_{1,0}$. However, one component of $LATE_{IM}$ — the mean untreated potential outcome of intensive-margin compliers in their no-grant counterfactual, $\bar{Y}^{IM}_{1,0}$ — is still nonparametrically point identified. Appendix D.2 shows that this counterfactual mean potential outcome is given by:

$$\bar{Y}^{\text{IM}}_{1,0} = \mathbb{E}[Y_{1,0} \mid \text{IM Complier}, k^*] = \frac{\bar{Y}^+_{1,0} - \bar{Y}^-_{1,0}\kappa_0}{1 - \kappa_0} \tag{11}$$

where $\bar{Y}^-_{1,0} = \lim_{k \uparrow k^*} \mathbb{E}[Y \mid S = 1, D = 0, k]$ and $\bar{Y}^+_{1,0} = \lim_{k \downarrow k^*} \mathbb{E}[Y \mid S = 1, D = 0, k]$ are limits of the mean outcomes of untreated students approaching the cutoff from the left and right, and where $\kappa_0 = \frac{f_1(k^{*-})}{f_1(k^{*+})}\frac{(1-\bar{D}^-)}{(1-\bar{D}^+)}$ is a modified first stage that corrects for the fact that there is a greater mass of students on the eligible side of the cutoff. The adjustment term κ_0 adapts the complier-describing logic of Abadie (2002) to a case where treatment leads some individuals to enter the sample. The intuition is the same: as intensive-margin compliers become treated as a result of crossing the threshold, the change in outcomes among *un*treated units identifies the intensive-margin compliers' counterfactual potential outcomes in the untreated state.[37] It is straightforward to verify that if there are no enrollment effects, then $f_1(k^{*-}) = f_1(k^{*+})$ and (11) collapses to a standard Abadie (2002)-style Wald estimand where the numerator is a treatment-by-outcome interaction.

A similar argument identifies the *pooled* mean potential outcomes for intensive-margin and extensive-margin compliers:

$$\bar{Y}_{1,1}^{\text{Complier}} = \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*] = \frac{\bar{Y}_{1,1}^{-} - \bar{Y}_{1,1}^{+}\kappa_1}{1 - \kappa_1} \tag{12}$$

where $\bar{Y}_{1,1}^{-}$ and $\bar{Y}_{1,1}^{+}$ are limits of the mean outcomes of treated students approaching the cutoff from the left and right, and where $\kappa_1 = \frac{f_1(k^{*+})}{f_1(k^{*-})}\frac{\bar{D}^{+}}{\bar{D}^{-}}$. The pooled treated potential outcome in (12) is a weighted average of the treated potential outcomes of intensive-margin and extensive-margin compliers:

$$\bar{Y}_{1,1}^{\text{Complier}} = (1-\omega)\bar{Y}_{1,1}^{\text{IM}} + \omega\bar{Y}_{1,1}^{\text{EM}} \tag{13}$$

where $\omega = P(S_1 > S_0 \mid S_1 = 1, D_1 > D_0, k = k^{*-})$ is the extensive-margin share among compliers, identified by (10).

Equation (13) is at the heart of the identification challenge. Comparisons across the grant threshold are only able to identify $\bar{Y}_{1,1}^{\text{Complier}}$, which is a weighted average of the mean treated potential outcomes of intensive-margin and extensive-margin compliers. However, point-identification of $LATE_{IM}$ and $LATE_{EM}$ instead requires estimates of $\bar{Y}_{1,1}^{IM}$ and $\bar{Y}_{1,1}^{EM}$, which average to $\bar{Y}_{1,1}^{\text{Complier}}$ by (13) but are not separately point identified. I show in the next section that despite this identification challenge, it is possible to obtain meaningful bounds on these LATEs given our knowledge of the complier shares.

7.3.1 Complier Characteristics are Separately Point Identified by Treatment Margin

Importantly, even though the treated outcomes of intensive-margin and extensive-margin compliers are not separately point identified, the pretreatment *characteristics* of students at both complier margins are separately identified by Equation (13). To see this, replace Y in (13) with a predetermined characteristic X to obtain:

$$\bar{X}_{1}^{\text{Complier}} = (1-\omega)\bar{X}_{1}^{\text{IM}} + \omega\bar{X}_{1}^{\text{EM}} \tag{14}$$

[37] Equation (11) is a special case of a more general result proven in Gerard et al. (2020). Their paper derives bounds on treatment effects of "non-manipulated" units (in this paper, intensive-margin compliers) in RD designs when there are also "manipulated" units on one side of the cutoff (in this case, extensive-margin compliers). In contrast to their paper, which assumes that manipulated units can be either treated or untreated, the exclusion restriction in Assumption 1 implies that all extensive-margin compliers are treated, entering the 2-year sample if and only if they receive a zero EFC as a result of crossing the grant assignment threshold. This requirement is key to the point-identification results in this subsection.

where the loss of the d subscript reflects the fact that predetermined characteristics do not depend on treatment. In this case, because $\bar{X}_1^{\text{Complier}}$ is identified by (12) and $\bar{X}_1^{\text{IM}}$ is identified by (11), it is possible to solve (14) for $\bar{X}_1^{\text{EM}}$, obtaining the mean covariate values of extensive-margin compliers. Appendix Figure A8 provides an example, showing estimates of mean tenth-grade test scores for intensive-margin and extensive-margin compliers at the Auto-Zero EFC Cutoff among students enrolled in two-year schools. The point estimates indicate that extensive-margin compliers are heavily negatively selected, with high school test scores more than 20 percentiles lower than those of intensive-margin compliers. However, the estimates are imprecise, owing to the fact that extensive-margin compliers represent a small fraction of the complier population.[38] These patterns show that extensive-margin compliers enter college with far lower average levels of ability, which becomes the basis for the mean-dominance assumption introduced to tighten nonparametric bounds on $\bar{Y}_{1,1}^{IM}$ and $\bar{Y}_{1,1}^{\text{EM}}$ in the next section.

7.4 Step 3: Nonparametric Bounds on Treatment Effects

7.4.1 Worst-Case Bounds

The starting point for the partial identification approach is to consider "worst-case" bounds (2020; 2009). These bounds are built from two identified quantities: the number of extensive margin compliers ω, together with the *distribution* of outcomes among the pooled set of all compliers in four-year schools just to the left of the cutoff.[39] The key intuition, first developed by Lee (2009), is that no sample of size $(1-\omega)$ drawn from the overall complier population can have a lower mean treated outcome than the lowest $(1-\omega)$ fraction of this distribution. Calculating the mean outcome among this lowest fraction therefore yields a "worst-case" bound for $\bar{Y}_{1,1}^{IM}$. A similar argument obtains an upper bound for $\bar{Y}_{1,1}^{IM}$ by assuming that intensive-margin compliers have the highest $(1-\omega)$ outcomes in the distibution of complier outcomes.

[38] This imprecision arises because $\bar{X}_1^{\text{EM}}$ is backed out of (14) by:

$$\bar{X}_1^{\text{EM}} = \frac{\bar{X}_1^{\text{Complier}} - (1-\omega)\bar{X}_1^{\text{IM}}}{\omega}$$

In some bootstrap replications, the estimate of ω is small, leading to the conclusion that the (large) overall decline in test scores is caused by the entry of a very small proportion of the population. This fact leads to standard errors for $\bar{X}_1^{\text{EM}}$ with a lower bound below zero, which is lower than permissible for a variable defined as a 0-100 percentile. The upper end of the 90 percent confidence interval for the test scores of EM compliers includes the level for IM compliers; nevertheless, the results suggest that extensive-margin compliers are negatively selected and provide motivation for the mean-dominance assumption.

[39] To see that this distribution is identified, note that we can replace the outcome in Equation (12) with $\mathbf{1}\{Y \leq y\}$, where y is a value in the support of the distribution of Y:

$$\begin{aligned}\frac{\lim_{k\uparrow k^*} \mathbb{E}[Y \mid G=1,k] - \kappa_1 \lim_{k\downarrow k^*} \mathbb{E}[Y \mid G=1,k]}{1-\kappa_1} &= \mathbb{E}[\mathbf{1}\{Y_{1,1} \leq y\} \mid S_1=1, D_1 > D_0, k=k^*] \\ &= H(y \mid S_1=1, D_1 > D_0, k=k^*)\end{aligned}$$

where $H(y \mid S_1 = 1, D_1 > D_0, k = k^*)$ is the CDF of Y among units with $(S_1 = 1, D_1 > D_0, k = k^*)$, evaluated at y.

To formalize this intuition, let $H_{S_1=1,D_1>D_0,k=k^*}(y)$ denote the distribution of outcomes among grant compliers at the cutoff who attend two-year schools in the treated state. Let y_ω and $y_{(1-\omega)}$ denote the ω-quantile and $(1-\omega)$-quantile of $H_{S_1=1,D_1>D_0,k=k^*}(y)$. Then we can obtain the following bounds on $\bar{Y}_{1,1}^{IM}$ without adding further assumptions:

$$\mathbb{E}[Y_{1,1} \mid \text{Complier}, k=k^*, Y_{1,1} \leq y_{1-\omega}] \leq \bar{Y}_{1,1}^{IM} \leq \mathbb{E}[Y_{1,1} \mid \text{Complier}, k=k^*, Y_{1,1} \geq y_\omega] \qquad (15)$$

Of course, each bound on the mean treated outcomes of intensive-margin compliers $\left(\bar{Y}_{1,1}^{IM}\right)$ also implies a bound on mean treated outcomes of extensive-margin compliers $\left(\bar{Y}_{1,1}^{EM}\right)$ because the weighted average of these two terms must add up to the overall complier mean $\left(\bar{Y}_{1,1}^{\text{Complier}}\right)$ by Equation (13). Intuitively, if the intensive-margin compliers have the lowest (highest) $(1-\omega)$ outcomes among all compliers, then the extensive-margin compliers must have the remaining highest (lowest) ω share of outcomes. The bounds in (15) therefore yield the following no-assumptions bounds on $\bar{Y}_{1,1}^{EM}$:

$$\mathbb{E}[Y_{1,1} \mid \text{Complier}, k=k^*, Y_{1,1} \leq y_\omega] \leq \bar{Y}_{1,1}^{EM} \leq \mathbb{E}[Y_{1,1} \mid \text{Complier}, k=k^*, Y_{1,1} \geq y_{(1-\omega)}]$$

7.4.2 Mean Dominance

We can start to tighten these bounds by progressively adding assumptions on the relationship between potential outcomes and students' selection patterns. The first assumption is that mean treated outcomes among intensive-margin compliers are weakly larger than mean treated outcome among extensive-margin compliers:

Assumption 5 (Mean Dominance)**.** $\bar{Y}_{1,1}^{IM} \geq \bar{Y}_{1,1}^{EM}$.

Assumption (5) states that intensive-margin compliers, who would enroll in college even without reciving a zero EFC, have weakly greater outcomes on average than extensive-margin compliers who attend college if and only if they receive grants.[40] It states that if all grant compliers were forced to attend college, then students who would have attended college even without additional grant

[40]This assumption is a related to the Monotone Treatment Selection (MTS) restriction first introduced by Manski and Pepper (2000), but is not identical. Standard applications of MTS make assumptions about the potential outcomes of units who take up treatment versus the potential outcomes of units who do not. For example, considering the "treatment" of attending a community college, a standard MTS assumption in my setting might be:

$$\mathbb{E}[Y_1 \mid S=1] \geq \mathbb{E}[Y_1 \mid S=0]$$

which would state that the potential outcomes of students who select into two-year schooling are different from the potential outcomes of students who do not select into two-year schooling. Assumption (5) is slightly different from such a condition because the populations on both sides of the equation $Y_{1,1}^{\text{IM}} \geq Y_{1,1}^{\text{EM}}$ enroll in two-year schooling if they receive a zero EFC ($S_1 = 1$). Rather than restricting expected values of treated outcomes $Y_{1,1}$ based on the observed college choice under grant assignment (S_1), Assumption (5) instead restricts potential outcomes based on the *counterfactual* college choice S_0 that students would make if they did not receive grants.

aid would perform at least as well as extensive-margin compliers brought into college attendance as a result of treatment. Such a restriction is consistent with the notion that students sort into college options based on their comparative advantage (2016), and is motivated by the fact that extensive-margin compliers have far lower pre-college test scores than intensive-margin compliers, as shown in Appendix Figure A9.

Assumption (5) tightens bounds on treatment effects by raising the lower bound on $\bar{Y}_{1,1}^{\text{IM}}$ and lowering the upper bound on $\bar{Y}_{1,1}^{\text{EM}}$. Specifically, because $Y_{1,1}^{\text{IM}}$ and $\bar{Y}_{1,1}^{\text{EM}}$ must have $\bar{Y}_{1,1}^{\text{Complier}}$ as their weighted average, the constraint that $\bar{Y}_{1,1}^{\text{IM}} \geq \bar{Y}_{1,1}^{\text{EM}}$ implies that the lower bound on $\bar{Y}_{1,1}^{\text{IM}}$ and the upper bound on $\bar{Y}_{1,1}^{\text{EM}}$ both become $\bar{Y}_{1,1}^{\text{Complier}}$, the pooled treated mean potential outcome among all compliers.[41]

7.4.3 Monotone Treatment Response

Finally, the treatment effect bounds can be further tightened by assuming that the grant weakly improves student outcomes within each complier group:

Assumption 6 (Monotone Treatment Response)**.**

$$\begin{aligned} \mathbb{E}[Y_{1,1} \mid \textit{IM Complier}, k = k^*] &\geq \mathbb{E}[Y_{1,0} \mid \textit{IM Complier}, k = k^*] \\ \mathbb{E}[Y_{1,1} \mid \textit{EM Complier}, k = k^*] &\geq \mathbb{E}[Y_0 \mid \textit{EM Complier}, k = k^*] \end{aligned} \tag{16}$$

The "Monotone Treatment Response" (MTR) assumption (1997) restricts the sign of grant aid's impact on average outcomes, conditional on a student's complier margin. For example, if Y is an indicator for ever graduating from a four-year institution, Equation (16) assumes that the expected graduation probabilities of both intensive-margin and extensive-margin compliers are helped, on average, by the additional grant funding associated with receiving a zero EFC.[42] Although it may appear at first glance that this restriction does not "spill over" between complier types, imposing MTR for students at one complier margin also impacts bounds on treatment effects for students at the other complier margin. This spillover occurs because of the linear relationship between the mean treated outcomes of students of the two complier types required by Equation (13). Because the weighted average of the two complier groups' mean treated outcomes must equal the overall mean treated outcome for grant compliers, the lower bound imposed by MTR for one complier type tightens the upper bound for the mean treated outcome of the other complier type.

[41] To see this, note that because $\bar{Y}_{1,1}^{\text{IM}}$ and $\bar{Y}_{1,1}^{\text{EM}}$ must average to $\bar{Y}_{1,1}^{\text{Complier}}$ by (13), it cannot be the case that $Y_{1,1}^{\text{IM}} < \bar{Y}_{1,1}^{\text{Complier}}$ and $\bar{Y}_{1,1}^{\text{IM}} > \bar{Y}_{1,1}^{\text{EM}}$, because the average of a number ($Y_{1,1}^{\text{IM}}$) with a smaller number ($Y_{1,1}^{\text{EM}}$) cannot produce a larger number ($\bar{Y}_{1,1}^{\text{Complier}}$).

[42] Importantly, this restriction does not impose that *all* students' responses to receiving a grant move in the same direction; it only imposes a sign on students' average response to grant receipt within a complier type.

7.4.4 From Bounds on Treated Outcomes to Bounds on LATEs

Finally, to convert bounds on $\bar{Y}_{1,1}^{\text{IM}}$ and $\bar{Y}_{1,1}^{\text{EM}}$ into bounds on $LATE_{IM}$ and $LATE_{EM}$, I subtract the relevant counterfactual mean from each quantity. For intensive-margin compliers, this counterfactual mean is point identified by (11). However, because I do not observe the running variable for any students who do not attend college, it is not possible to precisely estimate counterfactual outcomes $\left(\bar{Y}_0^{EM}\right)$ for extensive-margin compliers on the barely-ineligible side of the cutoff.

I address this problem by constructing a noisy measure of counterfactual outcomes for extensive-margin compliers, taking means of outcomes over the full population of non-college-attending students. Because I observe outcomes for non-college-attenders, but do not precisely observe their running variable, this measure will be biased if the mean counterfactual outcomes of extensive-margin compliers (whose running variables are exactly at the threshold) differ systematically from those of the full population of non-college-attenders. Importantly, because students at the margin of receiving and Automaic Zero EFC are poorer than the overall population of non-college-attending high school graduates, my noisy estimate of $\bar{Y}_0^{\text{EM}}$ is likely to be higher than the true counterfactual mean outcomes of extensive-margin compliers at the cutoff. This, in turn, leads to conservative estimates of $LATE_{EM}$ that are likely smaller than the true treatment effect. In addition, note that this assumption does not affect estimates of $LATE_{IM}$ unless also imposing MTR assumptions that depend on $\bar{Y}_0^{\text{EM}}$.

7.5 Implementation

Estimating the bounds for discrete outcomes requires six ingredients: estimates of the density to the left and right side of the cutoff $(f_1(k^{*-}))$ and $(f_1(k^{*+}))$, as well as limits of the proportion of treated students approaching the cutoff from the left and right $(\bar{D}^-, \bar{D}^+)$ and of four limits of treatment-by-outcome interactions $(\bar{Y}_{1,1}^+, \bar{Y}_{1,1}^-, \bar{Y}_{1,0}^+, \bar{Y}_{1,0}^-)$. I estimate the densities using the CDF-based estimator of Cattaneo et al. (2018) and estimate limiting quantities of outcomes at the cutoff using local-linear regressions with optimal bandwidth set using the method of Calonico et al. (2014).

8 Results: Nonparametric Bounds on Grant Treatment Effects

8.1 Effects of a Zero EFC on Eventual B.A. Completion

I begin by estimating the impact of receiving an automatic Zero EFC on the eventual B.A. completion graduation rates of intensive-margin and extensive-margin compliers.

Figure 8 reports estimates of treatment effect bounds, where the outcome is an indicator for whether a student completes a BA degree within 6 years of initial college entry. Figure 8a shows no-assumptions bounds on the mean treated potential outcomes of the two complier types, $\bar{Y}_{1,1}^{\text{IM}}$ and

$\bar{Y}_{1,1}^{\text{EM}}$. The black downward-sloping line plots the linear relationship between $\bar{Y}_{1,1}^{\text{IM}}$ and $\bar{Y}_{1,1}^{\text{EM}}$ specified by Equation (13), showing all possible combinations of these two unknown quantities that are consistent with the share of extensive-margin compliers ω and the pooled treated potential outcome among all compliers $\bar{Y}_{1,1}^{\text{Complier}}$.[43] The extrema of the line, and their projections onto the x and y axes, correspond to the best-case and worst-case bounds described in (15); for example, the lower-right point of the line describes the case where no intensive-margin compliers graduate, which by (13) implies a 44 percent graduation rate among extensive-margin compliers. The upper-left point describes the opposite scenario in which no extensive-margin compliers graduate, corresponding to a 7 percent graduation rate among intensive-margin compliers. These limits describe Lee (2009) bounds, and are obtained by trimming the distribution of compliers' treated potential outcomes. Importantly, the bounds contain the counterfactual mean potential outcome for intensive-margin compliers, shown by the blue horizontal dotted line. As a result, it is impossible to rule out without further assumptions whether treatment effects for intensive-margin compliers are positive or negative.

Figure 8b adds the Mean-Dominance assumption, imposing the condition that IM compliers fare weakly better than EM compliers in terms of their graduation rates ($\bar{Y}_{1,1}^{\text{IM}} \geq \bar{Y}_{1,1}^{\text{EM}}$). The assumption is represented by the 45-degree line, which represents equality between $\bar{Y}_{1,1}^{\text{IM}}$ and $\bar{Y}_{1,1}^{\text{EM}}$. The remaining permissible combinations of $(\bar{Y}_{1,1}^{\text{EM}}, \bar{Y}_{1,1}^{\text{IM}})$ lie above the 45-degree line's intersection with the black line. The assumption considerably narrows the bounds for both groups; The lower bound for IM compliers and the upper bound for EM compliers becomes 6 percent, which is the mean treated graduation rate among all compliers $\bar{Y}_{1,1}^{\text{Complier}}$. Notably, imposing this assumption places the lower bound for $\bar{Y}_{1,1}^{\text{IM}}$ above the counterfactual mean potential outcome for intensive-margin compliers, meaning that the lower bound for the LATE among intensive margin compliers becomes positive.

Figure 8c presents the bounds on $LATE_{IM}$ and $LATE_{EM}$ that are implied by the patterns in Panels (a) and (b). The blue bars report bounds on $LATE_{IM}$ and the yellow bars report bounds on $LATE_{EM}$. To obtain $LATE_{IM}$, I subtract the point-identified counterfactual mean potential outcome for IM compliers, $\mathbb{E}[Y_{1,0} \mid \text{IM Complier}]$, from the upper and lower bounds for $\mathbb{E}[Y_{1,1} \mid \text{IM Complier}]$ shown on the y-axis in Panels (a) and (b). Similarly, I obtain $LATE_{EM}$ by subtracting my estimate of $\mathbb{E}[Y_0 \mid \text{EM Complier}]$, keeping in mind the preceding discussion about how this estimate is a noisy measure of the true counterfactual graduation rate among EM compliers. Each step across the x-axis from left to right adds an additional assumption that tightens

[43] Specifically, the equation for the black downward-sloping line is obtained by rearranging (13):

$$\bar{Y}_{1,1}^{\text{IM}} = \frac{\bar{Y}_{1,1}^{\text{Complier}}}{1-\omega} - \frac{\omega}{1-\omega}\bar{Y}_{1,1}^{EM}$$

The line has slope $(-\omega/(1-\omega))$ and y-intercept $\bar{Y}_{1,1}^{\text{Complier}}/(1-\omega)$. Any permissible combination of $(\bar{Y}_{1,1}^{\text{EM}}, \bar{Y}_{1,1}^{\text{IM}})$ must lie somewhere on this line.

the bounds further. The figure shows that the bounds obtained via imposing the mean-dominance assumption almost achieve the tightness of bounds that also impose monotone treatment response, imposing that the grant weakly improves student outcomes.

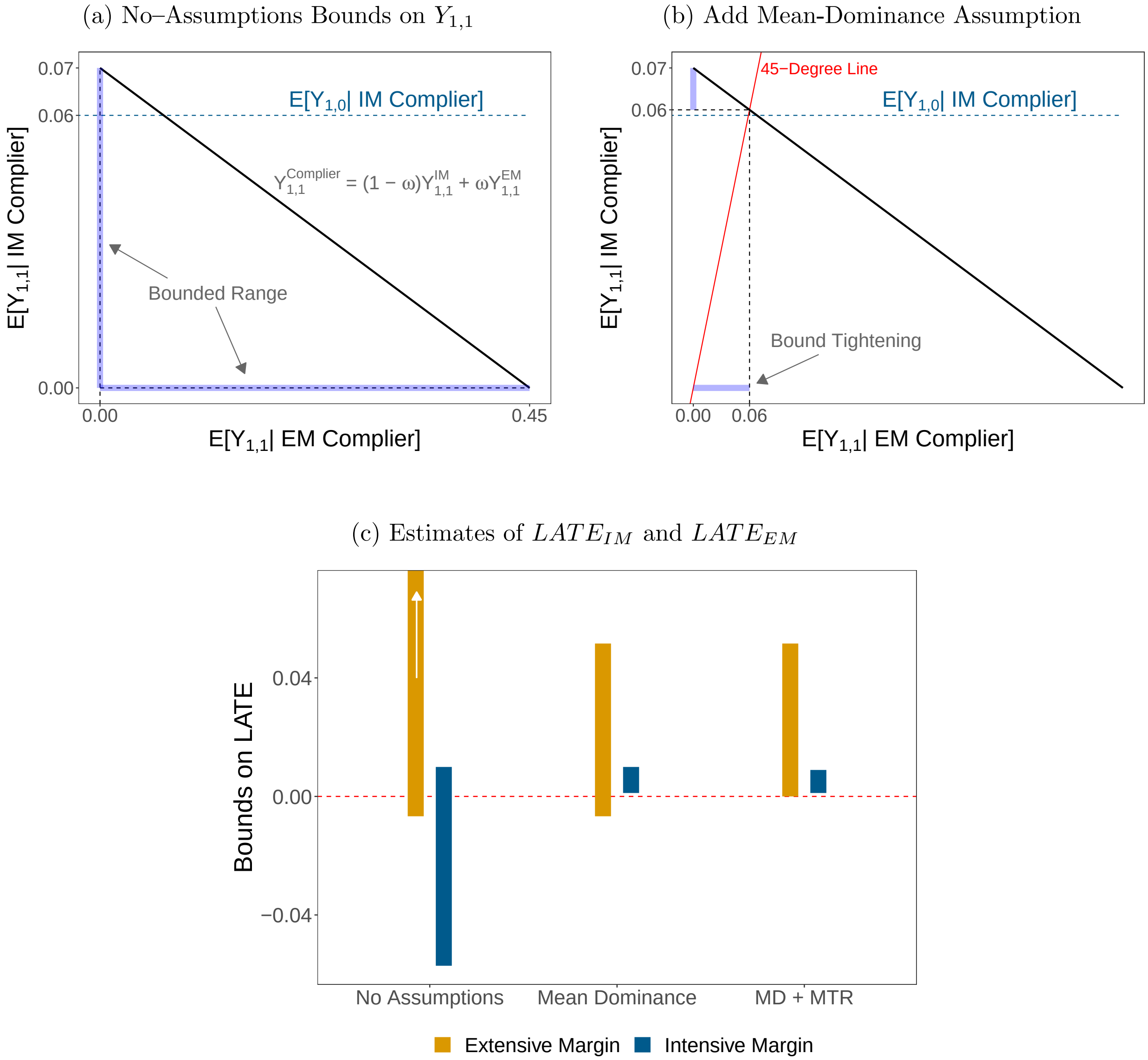
The intuition for these results can be summarized as follows. A naïve reduced-form RD estimation across the cutoff shows that 6-year graduation rates fall slightly across the threshold; Appendix Figure A10 shows that the overall graduation rate at the cutoff exhibits a small and statistically insignificant negative jump, corresponding to an reduced-form point estimate of -0.01. However, this negative jump combines two forces: treatment effects of the grant, and compositional changes due to the entry of extensive-margin compliers on the left of the cutoff. These forces may act in opposite directions, e.g. if the grant improves B.A. completion rates rates but new students entering the sample on the eligible side of the cutoff are less likely to graduate. Indeed, Appendix Figure A8 shows that extensive-margin compliers are heavily negatively selected, meaning they may be less likely to complete BA degrees, on average, than intensive margin compliers. If one imposes this belief upon the data, as in Figure 8b, then one must also believe that the treatment of a zero EFC causes small and statistically insignificant increases in BA completion rates among intensive margin compliers, ruling out that a zero EFC harms intensive-margin students' outcomes. Such a result is obfuscated by the no assumptions bounds in Figure 8a, which do not make any assumptions about the relative potential outcomes of different complier groups and thus cannot rule out that a Zero EFC is so harmful to intensive-margin compliers that it reduces their rates of B.A. completion to zero.

Effects on Graduation by Time Horizon

Figure 9 plots bounds on treatment effects separately by graduation horizon for intensive-margin and extensive-margin compliers. The bounds in the figure are constructed using the Mean Dominance assumption (Assumption 5), meaning that they assume that intensive-margin compliers perform weakly as well as extensive-margin compliers in college. Appendix Table B8 reports the estimates that correspond to the bounds plotted in the figure.

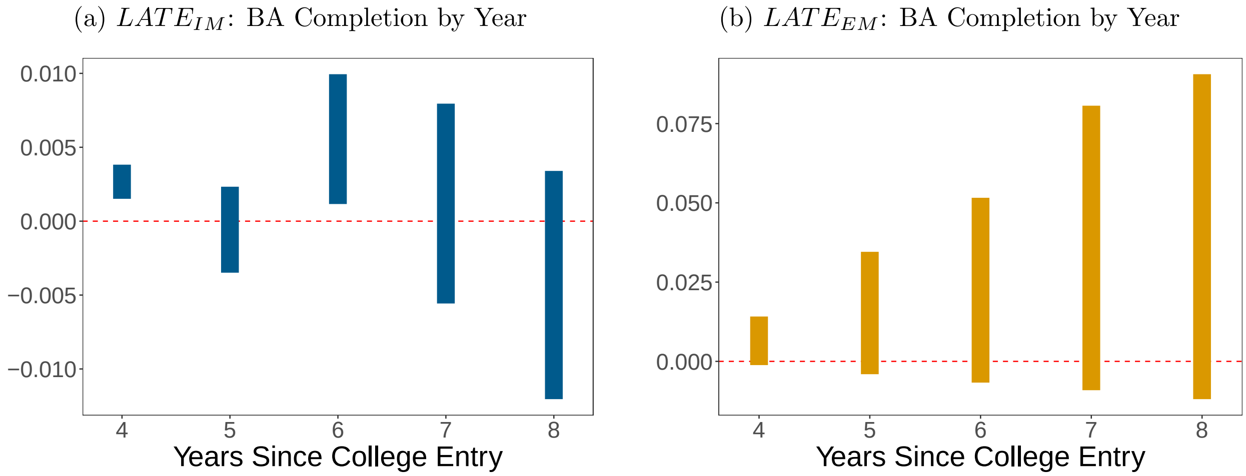
The figure indicates that receiving a Zero EFC has limited effects on the graduation probabilities of intensive-margin compliers. The bounded ranges for $LATE_{IM}$ are positive at the four year and six year horizons, but are economically small; at the six-year horizon, the estimates suggest that recieiving a zero EFC improves intensive-margin compliers' rates of 6-year B.A. completion by no more than 1 percentage point. The bounds include zero at longer time horizons, but can also rule out large negative treatment effects of a zero EFC on students' probability of B.A. completion. Even though these effects are not economically large, they are substantially more precise than no-assumptions bounds; furthermore, they suggest that the negative overall graduation effects across

Figure 8: Bounds on $LATE_{IM}$ and $LATE_{EM}$: 6-Year BA Completion



Notes: This figure shows how the Mean-Dominance and Monotone Treatment Response assumptions tighten non-parametric bounds on treatment effects. The outcome in all panels is an indicator for completing a BA degree within 6 years of community college entry. Panels (a) and (b) show bounds on treated potential outcomes for extensive-margin compliers (on the x-axis) and intensive-margin compliers (on the y-axis). The blue shaded regions show the bounded area for each complier group. The blue dotted line shows the counterfactual mean outcome for intensive-margin compliers, $\mathbb{E}[Y_{1,1} \mid \text{Intensive-Margin Complier}, k^*]$. Panel (a) shows "no-assumptions" (e.g., (Lee 2009)) bounds. Panel (b) shows how the Mean Dominance assumption (Assumption 5) tightens bounds by assuming that intensive-margin compliers have weakly higher outcomes than extensive-margin compliers (represented by the red 45-degree line). Panel (c) shows the bounds on Local Average Treatment Effects (LATEs) corresponding to these assumptions, subtracting the counterfactual mean outcome from each group from the bounded ranges in Panels (a) and (b). See main text for details.

Figure 9: Bounds for $LATE_{IM}$ and $LATE_{IM}$ under Mean Dominance Assumption



Notes: This figure shows bounds on treatment effects of receiving a Zero EFC on the graduation probability of intensive-margin compliers (Panel A) and extensive-margin compliers (Panel B) at the Auto-Zero EFC cutoff in 2-year community colleges. The bounds are constructed using the Mean Dominance assumption (5); see main text for details. The point estimates corresponding to the bounds shown in the figures are reported in Appendix Table B8.

the cutoff documented in Appendix Figure A10 are due to the selection effect of extensive-margin compliers leaving the sample, not a negative treatment effect on intensive-margin compliers.

The bounds for extensive-margin compliers, shown in Figure 9b, are large and include zero at every time horizon. However, note that the linear relationship between $\bar{Y}_{1,1}^{\mathrm{IM}}$ and $\bar{Y}_{1,1}^{\mathrm{EM}}$ in (13) implies that the bounds cannot be considered independently of each other. If the intensive-margin compliers are at the lower bound for $LATE_{IM}$, then Equation (13) implies that extensive-margin compliers must be at their upper bound for $LATE_{EM}$. Moreover, the vast majority of the bounded range for $LATE_{EM}$ is above zero, implying that the only case in which treatment effects for this population are not positive is one where intensive-margin compliers achieve nearly the upper bound of $LATE_{IM}$. Overall, the results are consistent with null effects on the graduation rates of intensive-margin compliers, together with plausible evidence that being brought into college attendance by receiving an automatic zero EFC improves extensive-margin compliers' graduation outcomes.

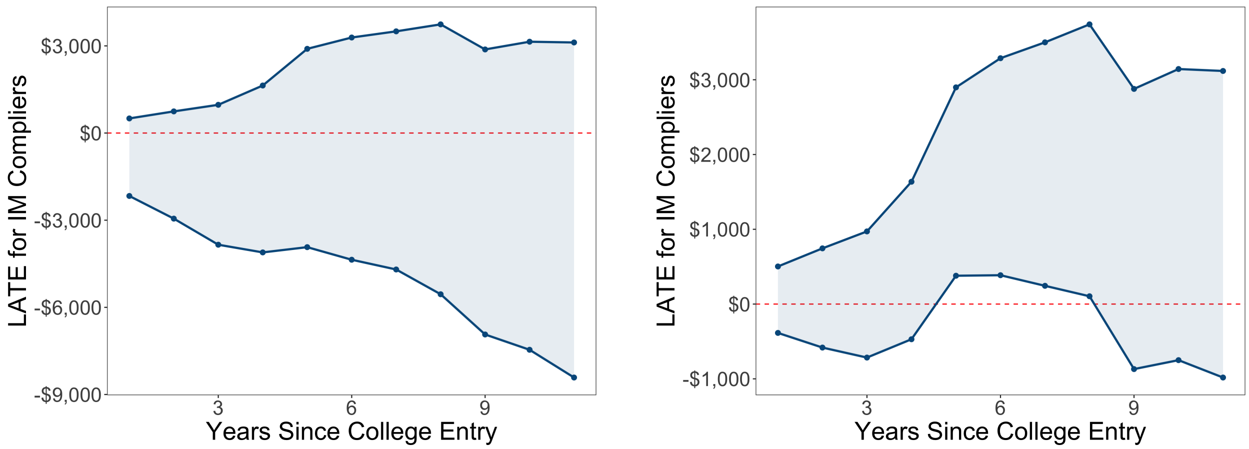
8.2 Effects of a Zero EFC on Yearly Earnings

Figure 10a reports no-assumptions bounds on the impact of receiving a Zero EFC on the yearly earnings of intensive-margin compliers. The no-assumptions bounds are too wide to be informative; for example, at six years after college entry, the bounds cannot rule out that the earnings of intensive-margin compliers increase by \$3,200 or fall by \$4360. The width of these bounds stems from the fact that extensive-margin compliers make up roughly 13 percent of all compliers on the

Figure 10: Bounds on Earnings Effects: Intensive-Margin Compliers

(a) No–Assumptions Bounds

(b) Add Mean-Dominance Assumption



Notes: This figure reports bounds on the causal effect of receiving a zero EFC on the earnings trajectories of intensive-margin compliers. Panel (a) shows no-assumptions bounds. Panel (b) introduces the mean-dominance assumption, which raises the lower bound on treatment effects for intensive-margin compliers. Panel (c) adds monotone treatment response assumptions which restrict the grant's impact on earnings to be negative in the first four years after college entry and positive in subsequent years. Panel (d) adds positive monotone treatment response in later years for extensive-margin compliers, which results in tighter upper bounds on treatment effects for intensive-margin compliers.

barely-ineligible side of the cutoff; trimming the earnings distribution by this proportion therefore produces large no-assumptions bounds.

Figure 10b imposes the mean-dominance assumption, imposing that the earnings of intensive-margin compliers weakly exceed those of extensive-margin compliers at every time horizon. This assumption raises the lower bound on $LATE_{IM}$, revealing positive effects on the earnings of intensive-margin compliers from five to eight years after college entry. Six years after college entry, the lower bound on earnings impacts for intensive-margin compliers becomes \$385. While estimates at longer time horizons cannot rule out null effects of receiving a Zero EFC on students' labor market earnings, the lower bound on $LATE_{IM}$ under the mean-dominance assumption is still substantially tighter than the no-assumptions bound in Figure 10a.

Taken together, the results suggest that the treatment of a zero EFC in the entering year of college has either null or slightly positive impacts on the B.A. completion probability and medium-term earnings of intensive-margin compliers. Under the mean dominance assumption, raw differences in these outcomes across the threshold conflate this treatment effect with the negative selection effect caused by extensive-margin compliers' entry into the sample. While no-assumptions bounds are unable to distinguish between these two effects, bounds constructed under the Mean-Dominance assumption are substantially tighter, and are motivated by the fact that extensive-margin compliers have pre-treatment covariates indicating much lower levels of incoming college preparedness.

9 Costs and Benefits of Grant Aid

In this section, I unify the analyses of the TEXAS Grant and the Automatic Zero EFC policies by comparing the relative costs and benefits of providing additional grant aid at each threshold.

9.1 Marginal Benefits and Costs

I model average per-student costs and benefits as a function of discounted yearly grants, loans, earnings, and costs per enrolled student:

$$\begin{aligned} B &= \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}[\text{Grants}_t + \text{Loans}_t + (1-\tau)\text{Earnings}_t - \mathbf{1}\{\text{Enrolled}\}_t \times \delta_{\text{tuition}}] \\ C &= \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}[\text{Grants}_t + \text{Loans}_t - \tau\text{Earnings}_t + \mathbf{1}\{\text{Enrolled}\}_t \times \delta_{\text{instruction}}] \end{aligned} \tag{17}$$

where β is the discount factor, τ is the tax rate on earnings, and δ_{tuition} and $\delta_{\text{instruction}}$ denote per-student tuition expenses and instructional costs.[44] Equation (17) states that benefits per enrolled student are the present discounted value of grants, loans, and after-tax earnings, net of tuition costs incurred in every enrolled year. Costs to the government are the discounted sum of grants, loans, and instructional expenditures per enrolled student, net of tax receipts.[45]

Appendix C.4 shows that the marginal benefits and costs of slightly expanding a grant program by raising the grant threshold k^* can be expressed in terms of LATEs for enrolled students:

$$\begin{aligned} \frac{\partial B}{\partial k^*} = \Bigg[& \sum_{t=1}^{T} \beta^{t-1} \left(LATE^*(\text{Grants}_t) + LATE^*(\text{Loans}_t) + (1-\tau) LATE^*(\text{Earnings}_t) \right) \\ & - \underbrace{\omega \times \delta_{\text{tuition}}}_{\substack{\text{First-Year} \\ \text{Enrollment Effect} \\ \text{(EM Compliers)}}} \underbrace{- \delta_{\text{tuition}} \times \sum_{t=2}^{T} LATE^*(\text{Enrolled}_t)}_{\substack{\text{Later-Year Enrollment Effects} \\ \text{(All Compliers)}}} \Bigg] \times \underbrace{P(S_1 = 1, D_1 > D_0 \mid k^*) f(k^*)}_{\text{Mass of Enrolled Compliers}} \end{aligned} \tag{18}$$

where $LATE^*(Y_t) = \mathbb{E}[Y_{1,1} - Y_{S_0,0} \mid S_1 = 1, D_1 > D_0, k = k^*]$ is the pooled effect of the grant treatment on both intensive-margin and extensive-margin compliers, and where $\omega = P(S_1 > S_0 \mid S_1 = 1, D_1 > D_0, k = k^*)$ is the fraction of enrolled compliers at the threshold who are extensive-margin compliers. Equation (18) shows that the marginal benefits of a grant expansion depend on the grant's causal effects on students' grants, loans, earnings, and enrollment trajectories. The

[44] I use $\beta = 0.95$ and a constant tax rate of $\tau = 0.33$ to calculate marginal costs and benefits.

[45] I abstract away from transfers between federal, state, and local governments and model the costs to a unitary public sector, which provides public education and provides subsidies to students through grants and loans.

term ($\omega \times \delta_{\text{tuition}}$) denotes first-year tuition for extensive-margin compliers who enroll in college as a result of receiving a grant; intensive-margin compliers do not incur this negative benefit, because they enroll at $t = 1$ regardless of whether they are treated. In later years, the negative benefits from tuition incorporate the effect of treatment on both extensive-margin and intensive-margin compliers.

Similarly, the marginal costs of raising the grant threshold are:

$$\begin{aligned}\frac{\partial C}{\partial k^*} = \Bigg[& \sum_{t=1}^{T} \beta^{t-1} \left(LATE^*(\text{Grants}_t) + LATE^*(\text{Loans}_t) - \tau LATE^*(\text{Earnings}_t) \right) \\ & + \delta_{\text{instruction}} \left(\omega + \sum_{t=2}^{T} LATE^*(\text{Enrolled}_t) \right) \Bigg] \times \underbrace{P(S_1 = 1, D_1 > D_0 \mid k^*) f(k^*)}_{\text{Mass of Enrolled Compliers}}\end{aligned} \tag{19}$$

where again, the enrollment term incorporates the first-year enrollment response of extensive-margin compliers.

Importantly, the marginal cost and benefit terms do not require separately estimating $LATE_{IM}$ and $LATE_{EM}$. Instead, they depend on $LATE^* = (1-\omega)LATE_{IM} + \omega LATE_{EM}$, which pools the effects on both complier groups. As a result, while the bounds in Section 7 are useful for separately considering the grant's impacts on the outcomes of intensive-margin and extensive-margin compliers, they are not necessary for evaluating the overall costs and benefits of the grant expansion. Instead, the point-identified ingredients $\bar{Y}_{1,1}^{\text{Complier}}$, $\bar{Y}_{1,0}^{IM}$, and $\bar{Y}_{0}^{\text{EM}}$, together with the extensive-margin complier share ω, are sufficient to characterize the $LATE^*$ terms:

$$LATE^* = \bar{Y}_{1,1}^{\text{Complier}} - \left[(1-\omega)\bar{Y}_{1,0}^{\text{IM}} + \omega \bar{Y}_{0}^{\text{EM}} \right] \tag{20}$$

this result underscores the usefulness of the selection-corrected potential outcomes identified in Section 7. Even if the analyst only observes data on enrolled students, it is possible to estimate the costs and benefits of grant expansions given that a noisy measure of $\bar{Y}_{0}^{\text{EM}}$, the counterfactual mean outcomes among extensive-margin compliers, is available.

9.2 The MVPF of a Grant Expansion

Taking the ratio of (18) and (19) yields the MVPF (Hendren and Sprung-Keyser 2020) of a marginal expansion of the grant threshold k^*:

$$\text{MVPF}_{k^*} = \frac{\sum_{t=1}^{T} \beta^{t-1} \left(\Delta_t^*(\text{Grant}) + \Delta_t^*(\text{Loan}) + (1-\tau)\Delta_t^*(\text{Earn})\right) + \delta_{\text{tuition}} \times \left(1 + \sum_{t=2}^{T} \Delta_t^*(\text{Enroll})\right)}{\sum_{t=1}^{T} \beta^{t-1} \left(\Delta_t^*(\text{Grant}) + \Delta_t^*(\text{Loan}) - \tau\Delta_t^*(\text{Earn})\right) + \delta_{\text{instructional}} \times \left(1 + \sum_{t=2}^{T} \Delta_t^*(\text{Enroll})\right)} \tag{21}$$

Equation (21) measures the value of the grant program net of the fiscal externality from potential increases in government tax revenue owing to the grant's treatment effect on students' lifetime earnings. An advantage of this MVPF criterion is that it can be used to infer the relative welfare weights required to allocate public funds to students at the margin of one grant program versus another.

9.3 Implementation

For students enrolled in four-year colleges, neither the TEXAS Grant nor the increase in grant aid at the Auto-Zero threshold impacts student enrollment. In this case, there are no extensive-margin compliers, so that $\omega = 0$ and $LATE^* = LATE_{IM}$ for these grant programs, simplifying the marginal cost and benefit formulas. Furthermore, I show in Section 5 that the TEXAS Grant has no meaningful impact on students' reenrollment probabilities, meaning the grant's costs and benefits boil down to its effects on grants, earnings, and loans. In contrast, the increase in aid at the Auto-Zero threshold has positive impacts on initial enrollment among students in two-year community colleges. For students at this treatment margin, I calculate yearly treatment effects in terms of the combined $LATE^*$ identified by Equation (20).

9.4 Estimates of Costs and Benefits

Table 3 reports estimates of the overall costs and benefits of grant aid for students at the three program margins: the TEXAS Grant and the Auto-Zero cutoff among four-year enrollees, and the Auto-Zero cutoff among two-year enrollees. There are several key takeaways. First, for the TEXAS Grant and the Automatic Zero EFC cutoff in two-year schools, the benefits of grant aid do not exceed the costs. For the TEXAS Grant, this effect arises because the grant provides substantial amounts of public funds to students at the grant margin, but fails to produce meaningful gains in later-life earnings that make up for this public expenditure. Column (1) shows that TEXAS Grant recipients receive a total discounted sum of roughly $13,700 in additional grant aid in the years following receiivng a TEXAS Grant, causing a $\$2,000$ decline in earnings and a roughly $\$4,400$ decline in loans. As a result, the grant represents a net transfer of $\$7,700$ to TEXAS Grant compliers. The government saves money on the loans that would have been taken out by compliers, but loses roughly \$1,000 in tax revenues that would have otherwise been collected from the compliers' labor market earnings, corresponding to a net cost of roughly $\$11,000$. Put together, I estimate that the MVPF of a marginal increase in the TEXAS Grant cutoff is 0.7, indicating benefits to compliers that are less than the grant's costs.

Column (3) shows that for the Automatic Zero EFC cutoff in two-year schools, the comparison of costs and benefits is even less favorable. The entry of extensive-margin compliers into the sample lowers benefits and increases costs due to increases in tuition and institutional expenditures. In addition, because the grant increase at the cutoff does not produce overall earnings gains for two-year students, the government does not recoup the costs of providing grant aid to students at the margin. Overall, while the benefits are positive for students at the margin of grant reciept, they are overwhelmed by the cost increases, producing a MVPF of roughly 0.3.

In contrast, Column (2) shows that for the Automatic Zero EFC cutoff in four-year schools, the benefits of providing grant aid are more than double the costs. The main reason for this difference is the grant's positive impact on students' lifetime earnings, which far outweigh increases in students' tuition costs stemming from the grant's impact on college persistence. The goverment recoups some of these earnings increases as tax revenue, lowering the overall costs. While this estimate suggests a positive MVPF from expenditures of public funds on four-year college students at the margin of the Auto-Zero cutoff, it is smaller than the infinite MVPF estimated by Denning et al. (2019), who find that the increase in grant aid to students is likely to be fully recouped by the government within 10 years. The reason for this discrepancy is my inclusion of costs to students and the government stemming from students' enrollment responses.

Overall, the results suggest that budget-neutral reallocations of funding to students at the Auto-Zero cutoff in four-year schools, financed by reductions in funding at the other two treatment margins, are likely to be welfare enhancing. Among four-year college students, such a reallocation would imply aggregate benefits from lowering the TEXAS Grant threshold, targeting grant aid towards lower-income students at the margin of receiving a Zero EFC rather than funding comparatively higher-income students at the margin of receiving a TEXAS Grant. Among the overall population of lower-income students at the margin of receiving an Automatic Zero EFC, the results suggest benefits from reallocating grant aid from the two-year to the four-year sector. The results suggest that despite the fact that while receiving Zero EFC has substantial impacts on two-year college enrollment, the overall gains in earnings and B.A. completion among compliers are not sufficiently large to justify spending public funds on this population instead of on four-year students.

Table 3: Costs and Benefits of Grant Programs

	TEXAS Grant (1)	Auto-Zero (4-Year) (2)	Auto-Zero (2-Year) (3)
Panel A. Benefits			
After-Tax Earnings	-2,088	2,653	-451
Grants	13,733	3,270	2,748
Loans	-4,447	105	81
Tuition and Fees	0	-409	-739
Total Discounted Benefits	7,700	7,779	1,761
Panel B. Costs			
Tax Receipts	1,029	-1,307	222
Grants	13,733	3,270	2,748
Loans	-4,447	105	93
Institutional Expenditures	0	1,642	2,958
Total Discounted Costs	11,070	3,255	6,237
MVPF	0.70	2.39	0.28
BCR	0.30	2.92	1.70

Notes: This table reports estimates of the costs and benefits of the TEXAS Grant and the increase in grant aid associated with crossing the Automatic Zero EFC threshold. Benefits to students are defined as the sum of after-tax earnings, grant receipts, loan awards, and changes in tuition, defined as the change in the probability of reenrollment multiplied by statewide average tuition and fees. Costs are defined as grants, loans, and expenditures net of tax revenues, where changes in expenditures are calculated as enrollment effects multiplied by statewide average institutional costs per full-time-enrolled student. Expenditure is calculated as average statewide four-year institutional expenditures per full-time enrolled student, calculated separately for 2-year and 4-year students using 2008-2012 IPEDS data and weighted by effects on enrollment. Tuition is calculated similarly using average statewide tuition and fees in 2-year and 4-year schools. I use a constant tax rate of 0.33 to calculate after-tax earnings and tax receipts. I use a deadweight loss of $\phi = 0.5$ to calculate the benefit-cost ratio.

10 Conclusion

The United States spends over $240 Billion on grant aid every year. While a substantial body of research indicates that grant aid enables some students to enroll in college, a substantial portion of funds may go to inframarginal students who would attend college even if not offered additional aid. This paper has explored the relative benefits and costs of targeting additional grant aid dollars at students of different incomes and pre-college preparedness, highlighting the heterogeneity in treatment effects that may stem from differential responses across students in the enrollment impact to grant aid.

Methodologically, separately identifying the impact of grant aid on marginal and inframarginal students presents a challenge. Among the population of enrolled students, it is not possible to tell from the data alone who appears in the sample because they received a grant aid award, and who would remain in the sample even without an additional aid award. I develop a bounding approach to separately estimate the impacts of grant aid on these two groups of students, and apply the using linked administrative data from Texas that links students' high school performance, college entry and graduation records, and later-life earnings. I find that the costs and benefits of grant aid heavily favor spending on low-income students in four-year colleges, who experience large earnings gains and increases in graduation and reenrollment.

References

Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, *97*(457), 284–292. Retrieved August 31, 2023, from http://www.jstor.org/stable/3085782

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., & Pathak, P. A. (2017). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica*, *85*(5), 1373–1432. https://doi.org/https://doi.org/10.3982/ECTA13925

Anderson, D. M., Broton, K. M., Goldrick-Rab, S., & Kelchen, R. (2020). Experimental evidence on the impacts of need-based financial aid: Longitudinal assessment of the wisconsin scholars grant. *Journal of Policy Analysis and Management*, *39*(3).

Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, *74*.

Andrews, R. J., & Stange, K. M. (2019). Price regulation, price discrimination, and equality of opportunity in higher education: Evidence from texas. *American Economic Journal: Economic Policy*, *11*(4), 31–65.

Angrist, J., Autor, D., & Pallais, A. (2022). Marginal effects of merit aid for low-income students. *Quarterly Journal of Economics*, *137*(2).

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the other 99 percent. *Science*, *344*(6186), 843–851.

Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2015). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*.

Baum, S., & Blagg, K. (2021). *A review of state grant aid in texas* (tech. rep.). Urban Institute Research Report. https://files.eric.ed.gov/fulltext/ED613254.pdf

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., & Stevens, M. (2019). The long-run impacts of financial aid: Evidence from california's cal grant. *American Economic Journal: Economic Policy*.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2019). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, *23*(2), 192–210. https://doi.org/10.1093/ectj/utz022

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

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326. Retrieved July 28, 2023, from http://www.jstor.org/stable/43616914

Castleman, B. L., & Long, B. T. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, *34*(4).

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261. https://doi.org/10.1177/1536867X1801800115

Chen, X., Flores, C. A., & Flores-Lagunes, A. (2018). Going beyond late: Bounding average treatment effects of job corps training. *Journal of Human Resources*, *53*(4).

Cohodes, S. R., & Goodman, J. S. (2014). Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, *6*(4), 251–85.

Deming, D., & Dynarski, S. (2010). Targeting investments in children: Fighting poverty when resources are limited. In P. B. Levine & D. J. Zimmerman (Eds.). University of Chicago Press.

Denning, J. T., Marx, B. M., & Turner, L. J. (2019). Propelled: The effects of grants on graduation, earnings, and welfare. *American Economic Journal: Applied Economics*, *11*(3).

Denning, J. T. (2018). Born under a lucky star: Financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources*, *58*(6).

Dynarski, S. (2000). Hope for whom? financial aid for the middle class and its impact on college attendance. *National Tax Journal*, *53*(3).

Dynarski, S. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review*, *93*(1).

Dynarski, S., Libassi, C., Michelmore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, *111*(6), 1721–56.

Dynarski, S., Page, L. C., & Scott-Clayton, J. (2022). College costs, financial aid, and student decisions. *NBER Working Paper 30275*.

Dynarski, S., & Scott-Clayton, J. (2013). Financial aid policy: Lessons from research. *The Future of Children*, *23*(1).

Eng, A., & Matsudaira, J. (2021). Pell grants and student success: Evidence from the universe of federal aid recipients. *Journal of Labor Economics*, *39*(S2).

Finkelstein, A., & Notowidigdo, M. J. (2019). Take-up and targeting: Experimental evidence from snap. *Quarterly Journal of Economics*, *134*(3), 1505–1556.

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, *58*(1), 21–29. https://doi.org/https://doi.org/10.1111/j.0006-341X.2002.00021.x

Fuller, W. C., Manski, C. F., & Wise, D. A. (1982). New evidence on the economic determinants of postsecondary schooling choices. *The Journal of Human Resources*, *17*(4), 477–498. Retrieved September 30, 2023, from http://www.jstor.org/stable/145612

Gerard, F., Rokkanen, M., & Rothe, C. (2020). Bounds on treatment effects in regression discontinuity designs with a manipulated running variable. *Quantitative Economics*, *11*(3), 839–870. https://doi.org/https://doi.org/10.3982/QE1079

Hendren, N., & Sprung-Keyser, B. (2020). A Unified Welfare Analysis of Government Policies*. *The Quarterly Journal of Economics*, *135*(3), 1209–1318. https://doi.org/10.1093/qje/qjaa006

Ida, T., Ishihara, T., Ito, K., Kido, D., Kitagawa, T., Sakaguchi, S., & Sasaki, S. (2023). Choosing who chooses: Selection-driven targeting in energy rebate programs. *Working Paper*.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Ito, K., Ida, T., & Tanaka, M. (2021). Selection on welfare gains: Experimental evidence from electricity plan choice. *Working Paper*.

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of Study, Earnings, and Self-Selection*. *The Quarterly Journal of Economics*, *131*(3), 1057–1111. https://doi.org/10.1093/qje/qjw019

Kline, P., & Walters, C. R. (2016). Evaluating Public Programs with Close Substitutes: The Case of Head Start*. *The Quarterly Journal of Economics*, *131*(4), 1795–1848. https://doi.org/10.1093/qje/qjw027

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies*, *76*(3), 1071–1102. https://doi.org/10.1111/j.1467-937X.2009.00536.x

Ma, J., & Pender, M. (2022). *Trends in college pricing and student aid 2022* (tech. rep.). College Board.

Manski, C. F. (1989). Anatomy of the selection problem. *The Journal of Human Resources*, *24*(3), 343–360. Retrieved September 30, 2023, from http://www.jstor.org/stable/145818

Manski, C. F. (1990). Nonparametric bounds on treatment effects. *The American Economic Review*, *80*(2), 319–323. Retrieved August 21, 2023, from http://www.jstor.org/stable/2006592

Manski, C. F. (1997). Monotone treatment response. *Econometrica*, *65*(6), 1311–1334. Retrieved August 21, 2023, from http://www.jstor.org/stable/2171738

Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, *68*(4), 997–1010. Retrieved August 21, 2023, from http://www.jstor.org/stable/2999533

Marx, B. M., & Turner, L. J. (2018). Borrowing trouble? human capital investment with opt-in costs and implications for the effectiveness of grant aid. *American Economic Journal: Applied Economics*, *10*(2), 163–201.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test [The regression discontinuity design: Theory and applications]. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/https://doi.org/10.1016/j.jeconom.2007.05.005

Montenegro, S. (2020). Three essays on the economics of financial aid. *UT-Dallas Graduate Dissertation*.

Mountjoy, J. (2022). Community colleges and upward mobility. *American Economic Review*, *112*(8), 2580–2630. https://doi.org/10.1257/aer.20181756

Park, R. S. E., & Scott-Clayton, J. (2018). The impact of pell grant eligibility on community college students' financial aid packages, labor supply, and academic outcomes. *Educational Evaluation and Policy Analysis*, *40*(4), 557–585. https://doi.org/10.3102/0162373718783868

Rothstein, J. M. (2004). College performance predictions and the sat. *Journal of Econometrics*, *121*(1-2), 297–317.

Scott-Clayton, J. (2011). On money and motivation: A quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources*, *46*(3).

Scott-Clayton, J., & Zafar, B. (2019). Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid. *Journal of Public Economics*, *170*.

U.S. Department of Education. (2021). Digest of education statistics [Table 304.10]. https://nces.ed.gov/programs/digest/

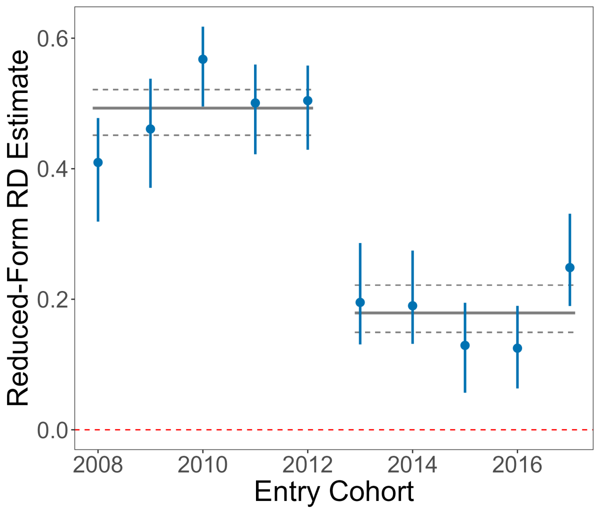
Villareal, M. U. (2018). The politics and policies of higher education in texas, 1995-2013. *UT-Austin Graduate Dissertation*.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. *The Annals of Mathematical Statistics*, *11*(3), 284–300. Retrieved September 5, 2023, from http://www.jstor.org/stable/2235677

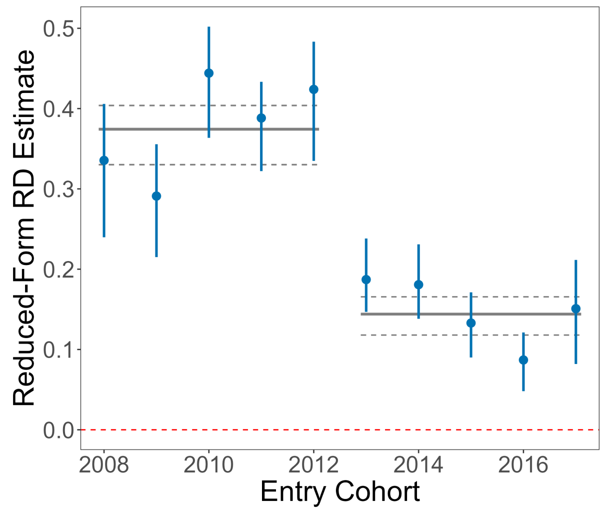
A Additional Figures

Figure A1: Strength of the First-Stage Relationship at the Auto-Zero Cutoff by Year

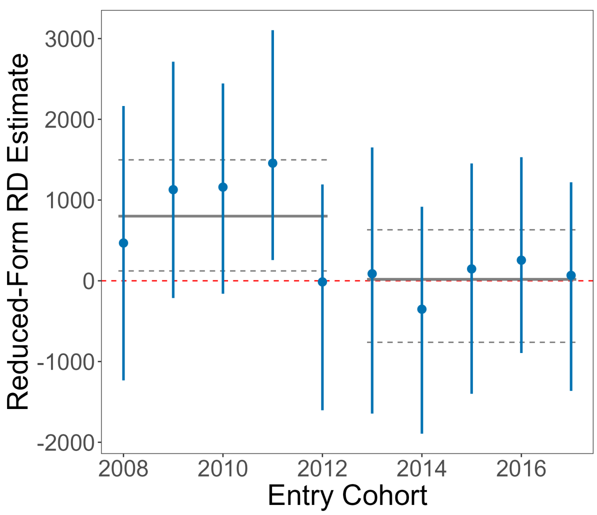
(a) Zero EFC: 4-Year Schools



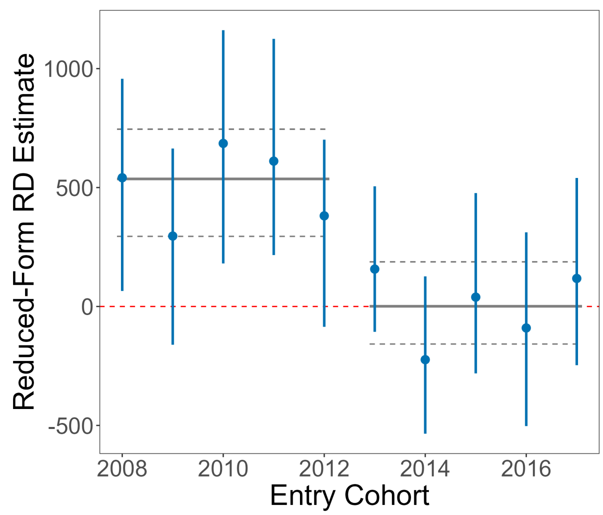
(b) Zero EFC: 2-Year Schools



(c) Total Grant Aid: 4-Year Schools



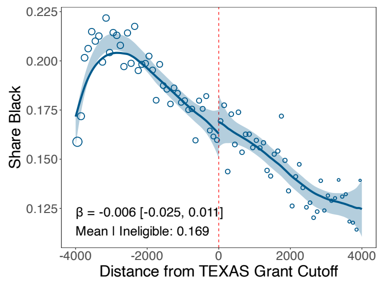
(d) Total Grant Aid: 2-Year Schools



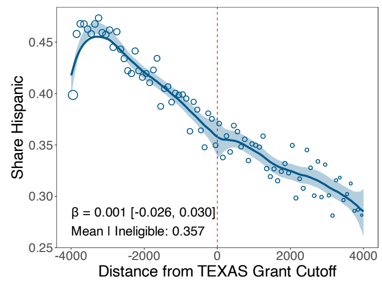
Notes: This figure shows regression discontinuity estimates of the effect of crossing the Auto-Zero EFC cutoff on receipt of a Zero EFC (Panels A and B) and students' first-year grant aid awards (Panels B and C). Each point estimate and 95 percent bias-corrected confidence interval (Calonico et al. 2014) is the result of a separate regression discontinuity estimate. The gray solid and dotted horizontal lines show estimates from regression discontinuity estimations that pool the 2008-2012 and 2013-2017 samples.

Figure A2: Covariate Balance: TEXAS Grant

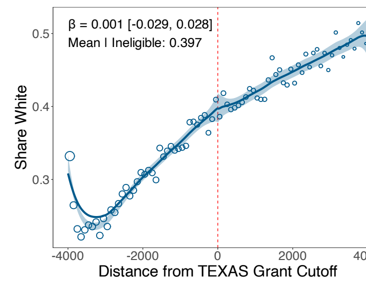
(a) Share Black



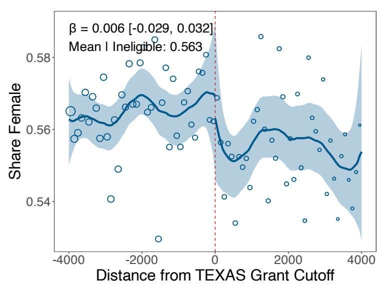
(b) Share Hispanic



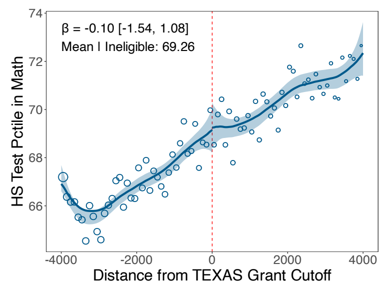
(c) Share White



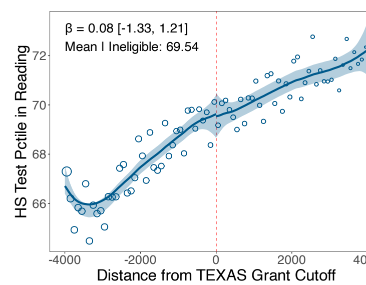
(d) Share Female



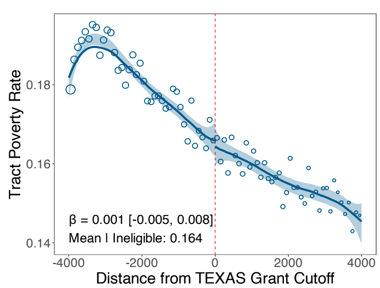
(e) HS Math Test Pctile



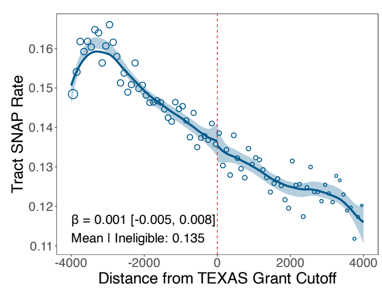
(f) HS Reading Test Pctile



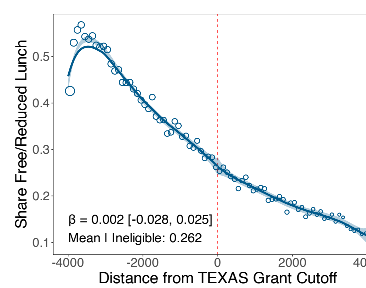
(g) Home Tract Poverty Rate



(h) Home Tract SNAP Rate

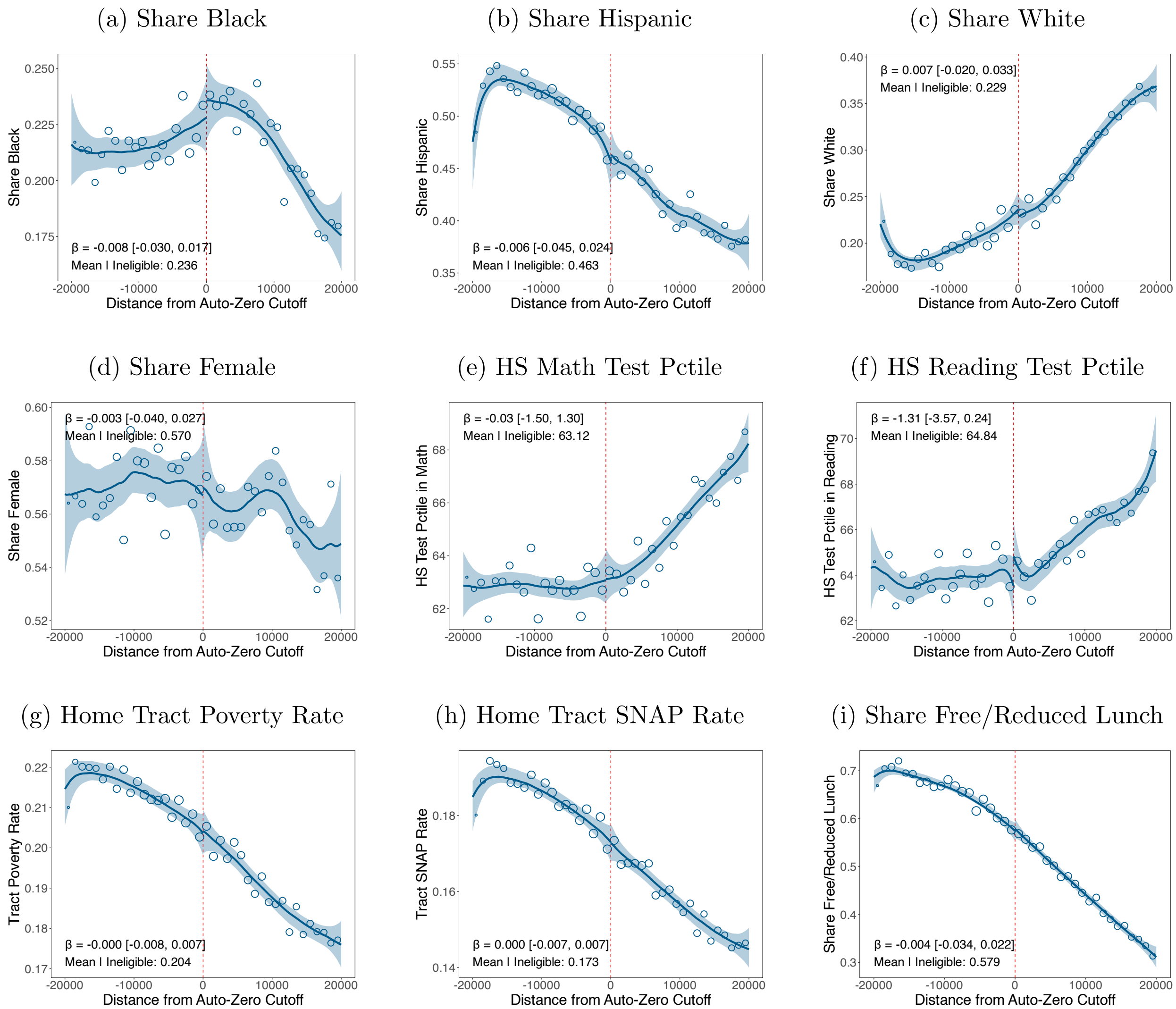


(i) Share Free/Reduced Lunch



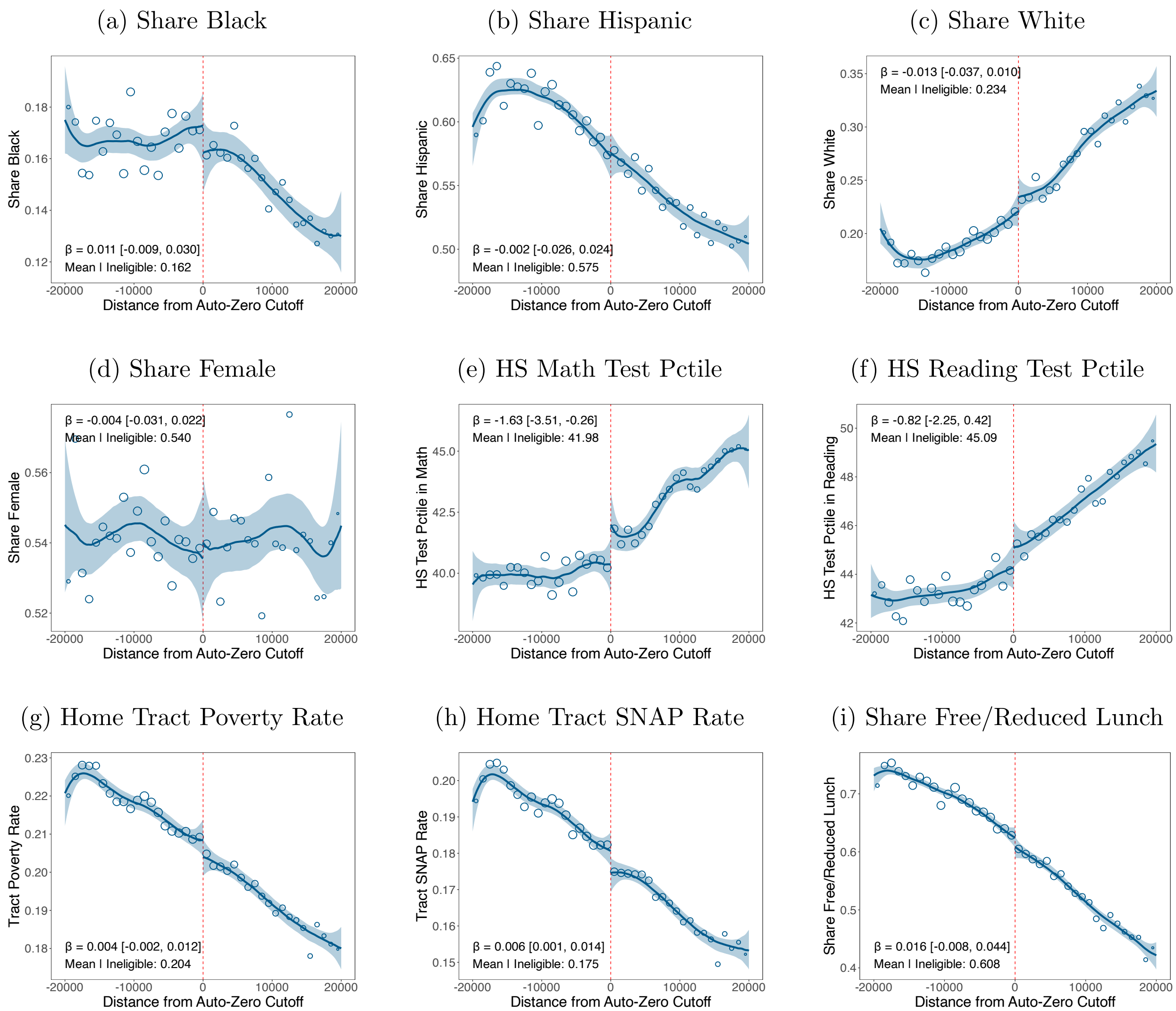
Notes: This figure shows individual characteristics by students' Expected Family Contribution (EFC) in students' entering year of college, for students within $4,000 of the TEXAS Grant threshold. All individual characteristics are obtained from high school graduation data and measured prior to college entry. The sample consists of students who enter college between the 2007 and 2017 academic years. In each graph, I plot unrestricted means within $100-wide EFC bins and predicted means from local linear regressions estimated separately on each side of the cutoff. The blue regions represent 95 percent confidence intervals of the regression lines. The annotation reports the estimated RD coefficient at the cutoff, together with the mean value of the characteristic on the right (ineligible) side of the cutoff.

Figure A3: Covariate Balance: Auto-Zero Cutoff in Four-Year Schools



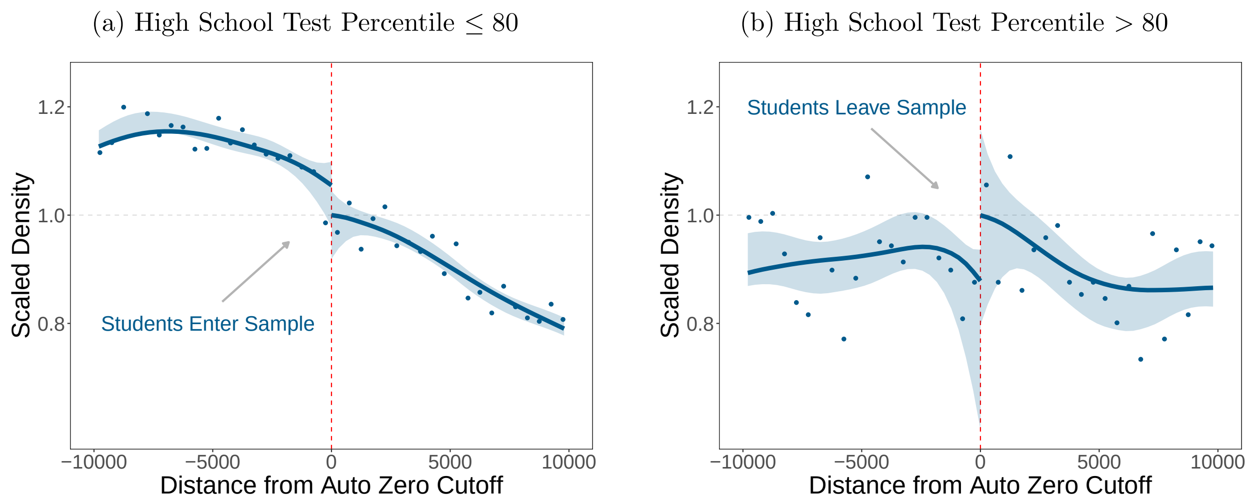
Notes: This figure shows individual characteristics by students' family Adjusted Gross Income (AGI), for students within $20,000 of the Auto Zero threshold in four-year schools. Family AGI is measured in students' entering year of college. All individual characteristics are obtained from high school graduation data and measured prior to college entry. The sample consists of students who enter college between the 2007 and 2012 academic years. In each graph, I plot unrestricted means within $1000-wide AGI bins and predicted means from local linear regressions estimated separately on each side of the cutoff. The blue regions represent 95 percent confidence intervals of the regression lines. The annotation reports the estimated RD coefficient at the cutoff, together with the mean value of the characteristic on the right (ineligible) side of the cutoff.

Figure A4: Covariate Balance: Auto-Zero Cutoff in Two-Year Schools



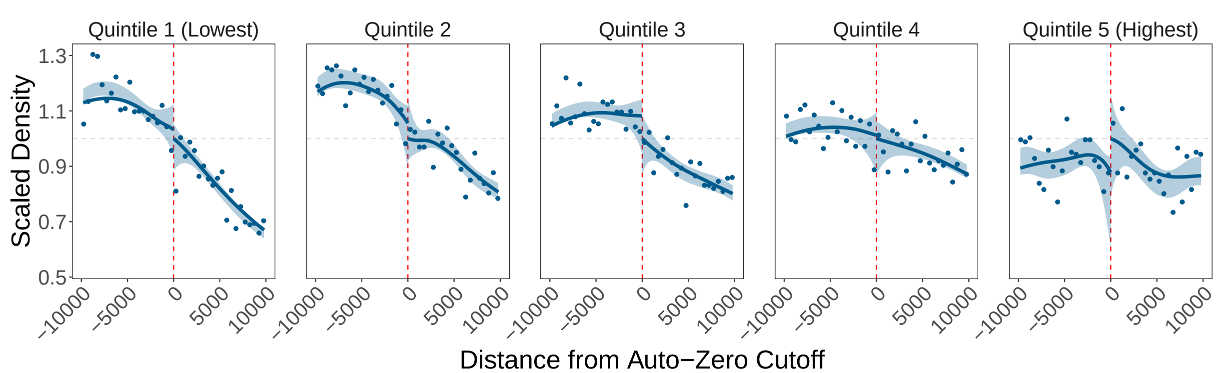
Notes: This figure shows individual characteristics by students' family Adjusted Gross Income (AGI), for students within \$20,000 of the Auto Zero threshold in two-year schools. Family AGI is measured in students' entering year of college. All individual characteristics are obtained from high school graduation data and measured prior to college entry. The sample consists of students who enter college between the 2007 and 2012 academic years. In each graph, I plot unrestricted means within \$1000-wide AGI bins and predicted means from local linear regressions estimated separately on each side of the cutoff. The blue regions represent 95 percent confidence intervals of the regression lines. The annotation reports the estimated RD coefficient at the cutoff, together with the mean value of the characteristic on the right (ineligible) side of the cutoff.

Figure A5: Defier Behavior among High-Test-Score Community College Students



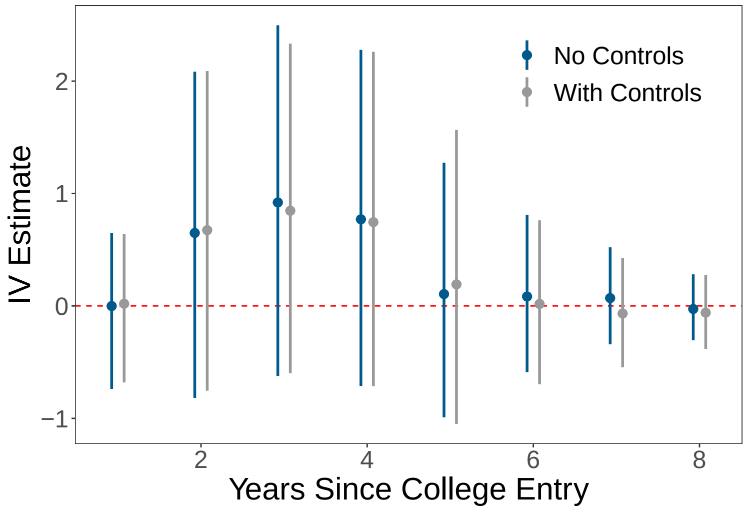
Notes: The figure plots density estimates of students in two-year community colleges in a window of the Automatic Zero EFC cutoff. Panel (a) plots the density for students whose high school test scores are below the 80th percentile of the test score distribution of their graduating class. Panel (b) plots the density for students whose test scores are in the top 20 percent of their high school graduating class. The densities are scaled so that 1 equals the point estimate of the density to the right of the threshold. The blue dots show scaled histogram estimates of the data in $1,000 AGI bins. The blue fits and 95% confidence intervals are constructed using the method of Cattaneo et al. (2018) at the MSE-optimal bandwidth.

Figure A6: Density Estimates by Quintile: Auto-Zero Cutoff in Two-Year Schools



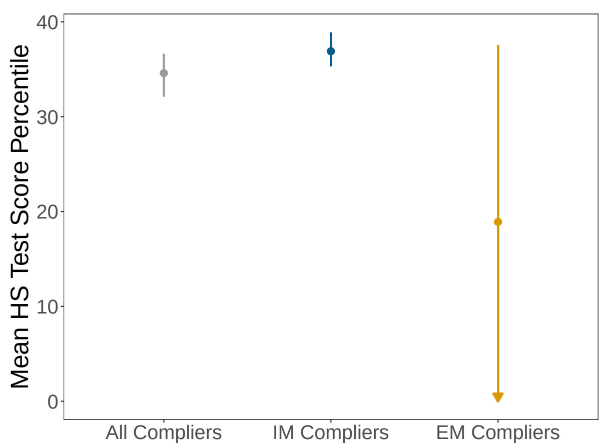
Notes: This figure plots the density of students' family Adjusted Gross Income (AGI) in a neighborhood of the Auto Zero cutoff for the sample of community college students. Each panel plots the density separately by a different quintile of students' high school test scores, with the lowest-scoring students on the left and the highest-scoring students on the right. The y axis is scaled so that 1 equals the density point estimate to the right of the cutoff. The dots plot scaled histogram estimates in $\$1,000$ AGI bins, and the solid lines and shaded regions are density estimates and bias-corrected 95% confidence intervals using the method of Cattaneo et al. (2018). The sample is all students entering between the 2008 and 2012 academic years.

Figure A7: Effects of the TEXAS Grant on Course Completion



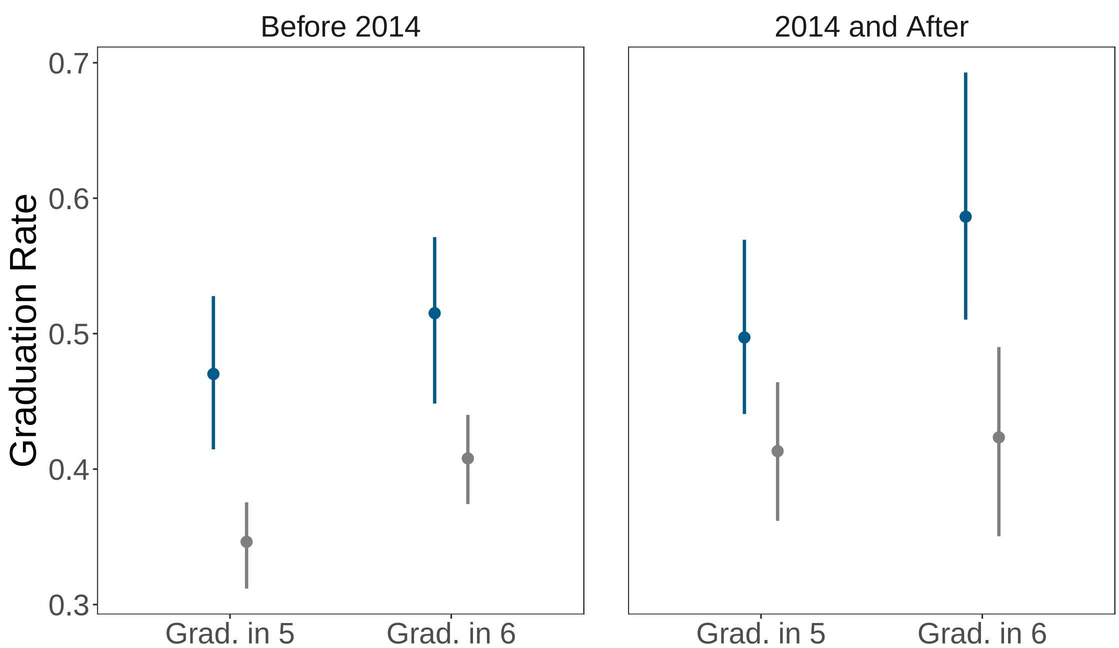
Notes: This figure shows regression discontinuity estimates of the impact of crossing the TEXAS Grant threshold on the credit completion of compliers. The x axis measures years since the student entered college, and the outcome variable at each time horizon is the number of credits completed in the given year. The figure shows point estimates and 95 percent bias-corrected confidence intervals constructed using the method of Calonico et al. (2014). The gray estimates report covariate-adjusted estimates following Calonico, Cattaneo and Farrell (2019).

Figure A8: Extensive-Margin Compliers are Negatively Selected



Notes: This figure reports estimates of the characteristics of extensive-margin and intensive-margin compliers at the threshold for Auto-Zero EFC assignment in two-year community colleges. The point estimates are constructed by estimating the analogue to (14) by the analogy principle, and the standard errors are constructed by taking the 5th and 95th percentiles of the estimated parameters from 1,000 bootstrap replications of the entire estimation procedure. The figure truncates the lower confidence interval for mean test scores among extensive-margin compliers; because the overall extensive-margin complier share is very small in some bootstrap replications, the lower bound on the confidence interval falls below zero despite the fact that test score percentile is bounded between 0 and 100.

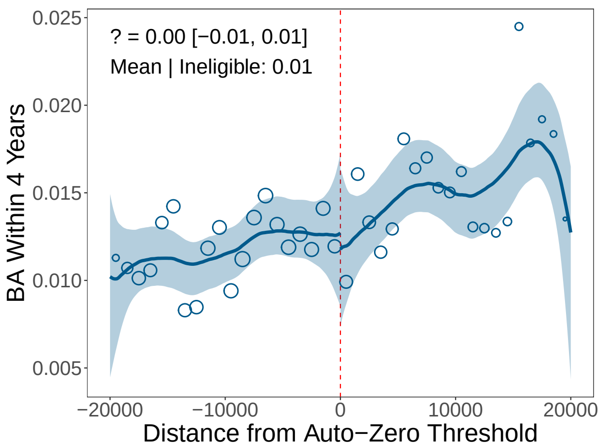
Figure A9: Grant Allocation at the Cutoff Before and After the 2014 Introduction of Merit Standards



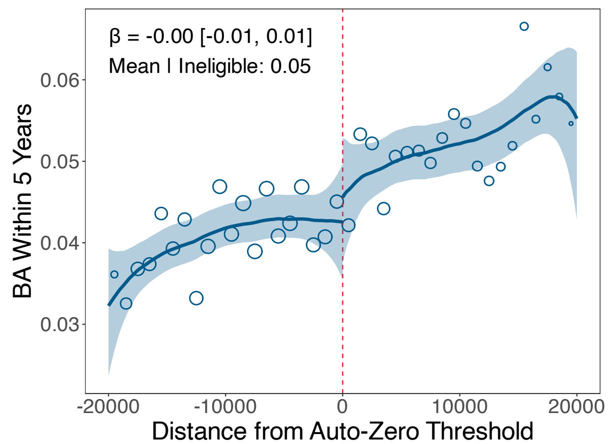
Notes: This figure shows the relative graduation rates of TEXAS Grant compliers and never-takers, separately for students who enter college in 2013 or prior (in the left panel) and for students who enter college in 2014 or later (in the right panel). The error bars are bias-corrected 95 percent confidence intervals following the method of (Calonico et al. 2014).

Figure A10: Raw Comparisons of BA Completion Rates at the Auto-Zero Cutoff in Community Colleges

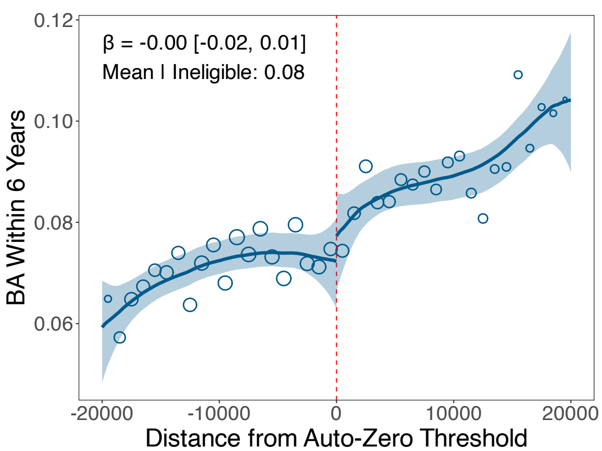
(a) BA Within 4 Years



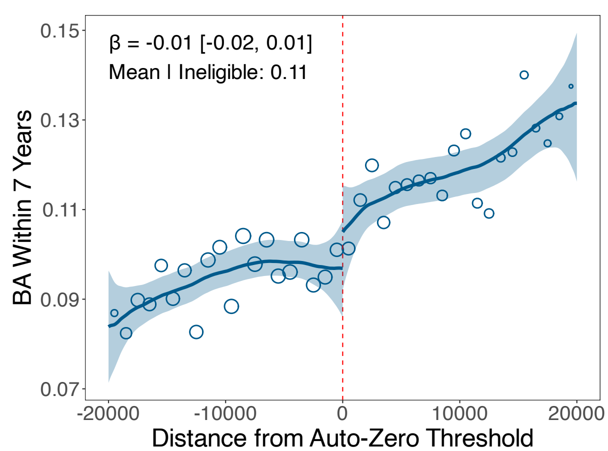
(b) BA Within 5 Years



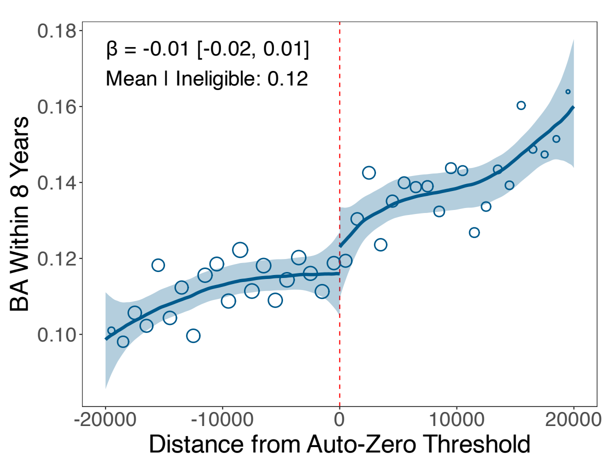
(c) BA Within 6 Years



(d) BA Within 7 Years



(e) BA Within 8 Years



B Additional Tables

Table B1: Density Tests at the TEXAS Grant Cutoff

		Cattaneo et al. (2018) Test				McCrary (2008) Test		
	N	BW_{left}	BW_{right}	Diff (s.e.)	p	BW	Diff (s.e.)	p
Optimal Bandwidth	146172	863.4	863.4	0.018	0.72	1185.9	0.000	0.98
				(0.049)			(0.023)	
By Bandwidth								
BW = 400	146172	400.0	400.0	0.038	0.17	400.0	-0.066	0.095
				(0.072)			(0.039)	
BW = 800	146172	800.0	800.0	0.008	0.96	800.0	-0.032	0.25
				(0.051)			(0.028)	
BW = 1200	146172	1200.0	1200.0	0.051	0.97	1200.0	0.001	0.97
				(0.042)			(0.023)	
Entry 2007–2012	82927	880.4	880.4	0.004	0.71	1397.3	-0.018	0.5
				(0.066)			(0.027)	
Entry 2013–2017	63245	1726.1	1726.1	0.065	0.09	1152.1	0.040	0.27
				(0.052)			(0.037)	
By Subsample								
Women	82086	1850.3	1850.3	-0.020	0.75	1558.5	-0.026	0.33
				(0.045)			(0.027)	
Men	64086	853.3	853.3	0.068	0.98	1059.2	0.019	0.6
				(0.074)			(0.037)	
Asian	10235	2862.0	2862.0	0.068	0.92	1200.9	-0.081	0.36
				(0.099)			(0.089)	
Black	25775	902.3	902.3	0.045	0.21	1226.1	0.043	0.44
				(0.112)			(0.055)	
Hispanic	57242	2921.6	2921.6	0.043	0.24	1096.7	-0.001	0.98
				(0.042)			(0.040)	
White	50755	1007.8	1007.8	0.075	0.72	1334.7	-0.007	0.83
				(0.083)			(0.034)	
Test Quintile = 1 (Lowest)	5440	1247.0	1247.0	0.091	0.87	1398.3	0.064	0.58
				(0.201)			(0.115)	
Test Quintile = 2	15753	1526.6	1526.6	0.014	0.07	1040.0	-0.117	0.14
				(0.107)			(0.079)	
Test Quintile = 3	27807	1827.1	1827.1	0.060	0.07	965.0	0.086	0.15
				(0.078)			(0.060)	
Test Quintile = 4	41310	1148.4	1148.4	0.040	0.89	1102.4	-0.024	0.58
				(0.081)			(0.044)	
Test Quintile = 5 (Highest)	55862	989.5	989.5	0.061	0.78	1343.3	-0.010	0.76
				(0.076)			(0.034)	

Notes: This table reports results from formal tests of equality of the density of Expected Family Contribution (EFC) across the TEXAS Grant cutoff. The estimation sample is students entering from 2007-2017 with EFC within $\$4,000$ of the cutoff. Columns 3 through 6 report estimates of the test of Cattaneo et al. (2018), and columns 7 through 9 report esitmates of the test of McCrary (2008). Except for tests by bandwidth, the bandwidths for each test are optimally chosen according to each test's selection procedure. Columns 5 and 8 report estimates and standard errors of the discontinuity in the estimated density at the threshold, and columns 6 and 9 report p-values from Wald tests of equality at the cutoff.

Table B2: Density Tests at the Auto-Zero Cutoff in 4-Year Schools

		Cattaneo et al. (2018) Test				McCrary (2008) Test		
	N	BW$_{left}$	BW$_{right}$	Diff (s.e.)	p	BW	Diff (s.e.)	p
Optimal Bandwidth	80077	19999.0	19999.0	-0.042 (0.007)	0.72	5292.4	-0.018 (0.028)	0.52
By Bandwidth								
BW = 1000	80077	1000.0	1000.0	-0.001 (0.031)	0.33	1000.0	-0.044 (0.066)	0.51
BW = 3000	80077	3000.0	3000.0	-0.001 (0.018)	0.70	3000.0	-0.019 (0.038)	0.61
BW = 5000	80077	5000.0	5000.0	-0.001 (0.014)	0.78	5000.0	-0.018 (0.029)	0.55
BW = 7000	80077	7000.0	7000.0	-0.005 (0.012)	0.99	7000.0	-0.018 (0.025)	0.47
BW = 9000	80077	9000.0	9000.0	-0.007 (0.010)	0.87	9000.0	-0.021 (0.022)	0.33
By Subsample								
Women	45332	5186.9	5186.9	0.002 (0.018)	0.47	6018.4	-0.016 (0.035)	0.65
Men	34745	19999.0	19999.0	-0.035 (0.011)	0.44	4747.5	-0.025 (0.046)	0.58
Asian	5137	13573.2	13573.2	0.021 (0.034)	0.36	6188.3	-0.028 (0.102)	0.78
Black	17258	6085.9	6085.9	0.020 (0.029)	0.41	6767.5	0.014 (0.052)	0.78
Hispanic	37373	18468.3	18468.3	-0.056 (0.011)	0.98	4857.7	-0.002 (0.044)	0.96
White	19607	5237.7	5237.7	-0.005 (0.027)	0.37	6525.3	-0.046 (0.053)	0.38
Test Quintile = 1 (Lowest)	4645	10432.6	10432.6	0.013 (0.042)	0.78	7195.2	-0.073 (0.101)	0.47
Test Quintile = 2	11309	5218.5	5218.5	0.015 (0.037)	0.69	5028.1	0.001 (0.076)	0.99
Test Quintile = 3	16861	6203.6	6203.6	-0.011 (0.028)	0.53	5375.2	-0.074 (0.060)	0.22
Test Quintile = 4	21933	8190.1	8190.1	0.013 (0.021)	0.24	6130.2	0.025 (0.050)	0.61
Test Quintile = 5 (Highest)	25329	6825.5	6825.5	-0.030 (0.021)	0.59	8503.2	-0.034 (0.041)	0.4

Notes: This table reports results from formal tests of equality of the density of family Adjusted Gross Income (AGI) across the Auto-Zero cutoff in 4-year schools. The estimation sample is students entering from 2008-2012 with AGI within $\$20,000$ of the cutoff. Columns 3 through 6 report estimates of the test of Cattaneo et al. (2018), and columns 7 through 9 report esitmates of the test of McCrary (2008). Except for tests by bandwidth, the bandwidths for each test are optimally chosen according to each test's selection procedure. Columns 5 and 8 report estimates and standard errors of the discontinuity in the estimated density at the threshold, and columns 6 and 9 report p-values from Wald tests of equality at the cutoff.

Table B3: Density Tests at the Auto-Zero Cutoff in 2-Year Schools

		Cattaneo et al. (2018) Test				McCrary (2008) Test		
	N	BW_{left}	BW_{right}	Diff (s.e.)	p	BW	Diff (s.e.)	p
Optimal Bandwidth	113896	5402.5	5402.5	-0.009 (0.011)	0.46	4398.5	-0.006 (0.027)	0.83
By Bandwidth								
BW = 1000	113896	1000.0	1000.0	0.013 (0.026)	0.21	1000.0	0.059 (0.057)	0.3
BW = 3000	113896	3000.0	3000.0	-0.006 (0.015)	0.71	3000.0	0.007 (0.032)	0.84
BW = 5000	113896	5000.0	5000.0	-0.009 (0.012)	0.54	5000.0	-0.008 (0.025)	0.75
BW = 7000	113896	7000.0	7000.0	-0.011 (0.010)	0.54	7000.0	-0.019 (0.021)	0.36
BW = 9000	113896	9000.0	9000.0	-0.012 (0.009)	0.29	9000.0	-0.024 (0.018)	0.2
By Subsample								
Women	61645	4714.3	4714.3	-0.010 (0.017)	0.61	6268.4	-0.008 (0.030)	0.8
Men	52251	5040.3	5040.3	-0.007 (0.017)	0.70	4393.6	-0.029 (0.039)	0.46
Asian	2759	7745.6	7745.6	-0.073 (0.061)	0.04	5354.6	-0.399 (0.156)	0.01
Black	18192	8071.6	8071.6	-0.028 (0.024)	0.67	5388.8	-0.073 (0.058)	0.21
Hispanic	66034	5680.3	5680.3	-0.009 (0.015)	0.59	4475.4	0.001 (0.035)	0.98
White	26006	4656.4	4656.4	0.003 (0.025)	0.95	5374.0	0.043 (0.050)	0.39
Test Quintile = 1 (Lowest)	29987	8399.7	8399.7	-0.009 (0.018)	0.32	4631.4	-0.084 (0.050)	0.096
Test Quintile = 2	31141	14171.5	14171.5	-0.026 (0.013)	0.74	6094.0	-0.010 (0.043)	0.82
Test Quintile = 3	25515	6301.4	6301.4	-0.016 (0.023)	0.67	6443.6	-0.047 (0.045)	0.29
Test Quintile = 4	18073	10862.2	10862.2	-0.005 (0.020)	0.78	5036.2	0.030 (0.063)	0.64
Test Quintile = 5 (Highest)	9180	5749.2	5749.2	0.040 (0.038)	0.10	5591.6	0.183 (0.083)	0.028

Notes: This table reports results from formal tests of equality of the density of family Adjusted Gross Income (AGI) across the Auto-Zero cutoff in 2-year schools. The estimation sample is students entering from 2008-2012 with AGI within $\$20,000$ of the cutoff. Columns 3 through 6 report estimates of the test of Cattaneo et al. (2018), and columns 7 through 9 report esitmates of the test of McCrary (2008). Except for tests by bandwidth, the bandwidths for each test are optimally chosen according to each test's selection procedure. Columns 5 and 8 report estimates and standard errors of the discontinuity in the estimated density at the threshold, and columns 6 and 9 report p-values from Wald tests of equality at the cutoff.

Table B4: RD Estimates: Institutional Characteristics

	All Students		> 1 Acceptance	
	RF	IV	RF	IV
	(1)	(2)	(3)	(4)
Mean Share Asian	-0.00	-0.00	-0.00	-0.00
	[-0.00, 0.00]	[-0.01, 0.01]	[-0.01, 0.01]	[-0.02, 0.01]
Mean Share Black	-0.01	-0.02	-0.00	-0.01
	[-0.02, 0.00]	[-0.04, 0.00]	[-0.02, 0.01]	[-0.04, 0.02]
Mean Share Hispanic	0.01	0.01	0.00	-0.00
	[-0.00, 0.02]	[-0.02, 0.04]	[-0.02, 0.02]	[-0.05, 0.04]
Mean Share White	0.00	0.01	0.00	0.02
	[-0.01, 0.01]	[-0.02, 0.03]	[-0.01, 0.03]	[-0.02, 0.06]
Mean Share Female	0.00	0.00	-0.00	-0.00
	[-0.00, 0.00]	[-0.01, 0.01]	[-0.01, 0.00]	[-0.02, 0.01]
Mean Incoming SAT Score	0.35	2.58	-0.12	1.58
	[-5.63, 6.07]	[-10.75, 15.42]	[-9.47, 10.19]	[-19.40, 23.65]
Mean Parent AGI	-778.66	-646.13	-654.96	-1358.60
	[-3302, 1539]	[-7261, 6186]	[-4743, 4558]	[-12358, 10719]
Mean Share w/ TEXAS Grant	0.00	0.01	0.00	0.00
	[-0.00, 0.01]	[-0.01, 0.02]	[-0.01, 0.01]	[-0.03, 0.03]
Mean Total Grant Aid	82.82	47.40	184.29	116.21
	[-227, 371]	[-731, 778]	[-246, 614]	[-982, 1149]
Mean Total Loans	12.50	149.45	257.96	722.10
	[-321, 280]	[-745, 1074]	[-254, 919]	[-628, 2306]
Mean TEXAS Grant	77.65	155.77	44.37	69.33
	[-18, 182]	[-58, 395]	[-93, 183]	[-296, 412]
Mean Pell Grant	44.42	-3.55	34.70	-132.53
	[-82, 165]	[-309, 299]	[-154, 188]	[-678, 265]

Notes: This table reports estimates of changes in mean institutional characteristics among four-year college students across the TEXAS Grant cutoff. $N = 160,270$ for regressions on all students and $N = 55,025$ for regressions on students with > 1 acceptance. The outcome in each regression is the mean of the given characteristic within a student's institution and entering cohort. Columns (1) and (2) report results for the full population of college entrants between the 2007 and 2017 academic years. Columns (3) and (4) restrict the sample to students whose application records indicate more than one acceptance to a Texas public four-year college. Each estimation uses the MSE-optimal bandwidth calculated separately for each outcome using the method of Calonico et al. (2014).

Table B5: Estimated Bounds for Potential Outcomes by Complier Group, 6-Year BA Completion

Complier Margin	Intensive	Extensive
	(1)	(2)
Panel A. Untreated Potential Outcomes		
Untreated Mean, IM Compliers	0.057 (0.015)	
Untreated Mean, EM Compliers		0.007 (Assumed)
Panel B. Treated Potential Outcomes		
Pooled Treated Mean, All Compliers	0.058 (0.017)	
Panel C. Bounds on Treated Potential Outcomes by Group		
Upper Bound, No Assumptions	0.067 (0.020)	0.445 (0.277)
Lower Bound, No Assumptions	0.000 (0.018)	0.000 (0.000)
Upper Bound, Mean Dominance	0.067 (0.020)	0.058 (0.017)
Lower Bound, Mean Dominance	0.058 (0.017)	0.000 (0.000)
Upper Bound, Mean Dominance + MTR	0.066 (0.019)	0.058 (0.031)
Lower Bound, Mean Dominance + MTR	0.058 (0.013)	0.007 (0.000)
Panel D. Bounds on LATEs by Complier Group		
Upper Bound, No Assumptions	0.010 (0.026)	0.439 (0.277)
Lower Bound, No Assumptions	-0.057 (0.023)	-0.007 (0.000)
Upper Bound, Mean Dominance	0.010 (0.026)	0.052 (0.017)
Lower Bound, Mean Dominance	0.001 (0.023)	-0.007 (0.000)
Upper Bound, Mean Dominance + MTR	0.009 (0.025)	0.052 (0.031)
Lower Bound, Mean Dominance + MTR	0.001 (0.015)	0.000 (0.000)
Share of Compliers	0.869	0.131

Notes: This table shows estimates of bounds on the effect of receiving a Zero EFC on students' probability of B.A. completion within six years of initial college entry. Panel A reports the mean B.A. completion rates of untreated intensive-margin compliers (Column (1)), estimated by Equation (11), and untreated extensive-margin compliers (Column (2)), estimated by taking the mean 6-year B.A. completion rate among all non-college-attenders in my sample. Panel B reports the pooled B.A. completion rates of treated intensive-margin and extensive-margin compliers, etsimated by Equation (12). Panel C reports upper and lower bounds on $\bar{Y}_{1,1}^{\mathrm{IM}}$ and $\bar{Y}_{1,1}^{\mathrm{EM}}$, the treated mean potential outcomes of intensive-margin and extensive-margin compliers. Panel D subtracts the counterfactual means in panel A from these estimates to arrive at bounds on $LATE_{IM}$ and $LATE_{EM}$. See main text for additional details.

Table B6: IV Estimates of the Impact of the TEXAS Grant on Persistence and Graduation

X =	2	3	4	5	6	7	8
A. Reenrollment in Year X							
TEXAS Grant	-0.01	0.01	0.01	0.01	0.02	0.01	0.00
	[-0.06, 0.05]	[-0.05, 0.06]	[-0.04, 0.07]	[-0.04, 0.06]	[-0.01, 0.05]	[-0.01, 0.03]	[-0.01, 0.01]
Mean | Ineligible	0.42	0.24	0.15	0.16	0.09	0.03	0.01
Observations	146,172	146,172	146,172	146,172	146,172	146,172	146,172
Obs. w/in Bandwidth	27,273	31,119	29,455	30,399	27,063	26,935	26,420
Panel B. Graduation by End of Year X							
TEXAS Grant	-0.00	-0.00	0.00	0.01	0.02	0.02	0.01
	[-0.01, 0.00]	[-0.02, 0.02]	[-0.05, 0.05]	[-0.06, 0.08]	[-0.05, 0.08]	[-0.04, 0.08]	[-0.07, 0.09]
Mean | Ineligible	-0.00	0.00	-0.06	0.00	0.04	0.05	0.07
Observations	146,172	146,172	146,172	133,480	126,714	112,927	99,246
Obs. w/in Bandwidth	40,982	25,981	27,580	21,910	25,463	33,234	21,766

Notes: This table reports estimates of the impacts of the TEXAS Grant on students' probability of renrollment in college (Panel A) and graduation (Panel B). Each point estimate comes from a separate fuzzy RD specification across the TEXAS Grant threshold, where the first stage is an indicator for receiving a TEXAS Grant in the first year of college. The columns show outcomes measured at different time horizons, measured from the year of college entry. Each estimation uses the MSE-optimal bandwidth calculated using the method of Calonico et al. (2014).

Table B7: IV Estimates of the Impact of the TEXAS Grant on Loans and Earnings

$X =$	1	2	3	4	5	6	7
Panel B. Loans in Year X							
TEXAS Grant	-2,186	-1,404	-465	-621	30	-114	63
	[-2881, -1554]	[-2078, -739]	[-1157, 224]	[-1299, -7]	[-438, 476]	[-419, 172]	[-130, 270]
Mean | Ineligible	6195	4693	4338	3798	1672	569	232
Observations	146,172	146,172	146,172	146,172	146,172	133,480	126,714
Obs. w/in Bandwidth	27,793	30,514	31,755	33,491	34,420	31,003	31,949
Panel B. Earnings in Year X							
TEXAS Grant	-734	-886	-779	-759	-140	-1,215	-972
	[-1434, -169]	[-1836, -22]	[-1875, 269]	[-2411, 605]	[-2347, 2035]	[-4284, 1708]	[-4450, 2413]
Mean | Ineligible	6189.16	10188.52	13256.89	14630.12	14981.33	17190.14	19380.95
Observations	146,172	146,172	146,172	146,172	146,172	146,172	133,480
Obs. w/in Bandwidth	25,523	25,604	28,474	22,868	26,081	23,118	23,405

Notes: This table reports estimates of the impacts of the TEXAS Grant on students' student loans (Panel A) and labor market earnings (Panel B). Each point estimate comes from a separate fuzzy RD specification across the TEXAS Grant threshold, where the first stage is an indicator for receiving a TEXAS Grant in the first year of college. The columns show outcomes measured at different time horizons, measured from the year of college entry. Each estimation uses the MSE-optimal bandwidth calculated using the method of Calonico et al. (2014).

Table B8: Bounds on BA Completion Effects by Horizon for Intensive-Margin and Extensive-Margin Compliers

		$LATE$ for IM Compliers			$LATE$ for EM Compliers		
BA Within:	Bound	No-Assm.	MD	MD + MTR	No-Assm.	MD	MD + MTR
4 Years	Upper	0.004 (0.009)	0.004 (0.009)	0.004 (0.009)	0.116 (0.255)	0.014 (0.007)	0.014 (0.009)
	Lower	-0.014 (0.006)	0.002 (0.008)	0.002 (0.005)	-0.001 (0.000)	-0.001 (0.000)	0.000 (0.000)
5 Years	Upper	0.002 (0.020)	0.002 (0.020)	0.002 (0.019)	0.291 (NA)	0.035 (0.013)	0.011 (0.023)
	Lower	-0.042 (0.015)	-0.003 (0.018)	0.000 (0.011)	-0.004 (0.000)	-0.004 (0.000)	0.000 (0.000)
6 Years	Upper	0.010 (0.026)	0.010 (0.026)	0.009 (0.025)	0.439 (0.277)	0.052 (0.017)	0.052 (0.031)
	Lower	-0.057 (0.023)	0.001 (0.023)	0.001 (0.015)	-0.007 (0.000)	-0.007 (0.000)	0.000 (0.000)
7 Years	Upper	0.008 (0.031)	0.008 (0.031)	0.007 (0.030)	0.676 (0.242)	0.081 (0.019)	0.044 (0.045)
	Lower	-0.095 (0.034)	-0.006 (0.028)	0.000 (0.015)	-0.009 (0.000)	-0.009 (0.000)	0.000 (0.000)
8 Years	Upper	0.003 (0.030)	0.003 (0.030)	0.002 (0.030)	0.771 (0.215)	0.091 (0.020)	0.011 (0.050)
	Lower	-0.115 (0.040)	-0.012 (0.027)	0.000 (0.014)	-0.012 (0.000)	-0.012 (0.000)	0.000 (0.000)

Notes: This table shows estimates of bounds on the effects of receiving a Zero EFC on the graduation probability of intensive-margin compliers and extensive-margin compliers at the Auto-Zero EFC threshold in two-year community colleges. Extensive-margin compliers are students who enter the sample as a result of treatment, and intensive-margin compliers are students who would remain in two-year colleges even without the treatment of receiving a zero EFC. Each row reports results for a separate outcome, defined as an indicator for completing any BA degree within the specified number of years. The "no-assumptions" columns report no-assumptions bounds, while the "MTS" and "MTS+MTR" bounds add additional assumptions; see main text for details. Standard errors are from $1,000$ bootstrap replications of the entire estimation procedure, following the bootstrap procedure for RD designs with manipulation of Gerard et al. (2020).

C Appendix

C.1 Proof of Equation (2)

Let $Y_{s,g}$ denote the potential outcomes of an individual who chooses schooling option $s \in \{0,1\}$ and receives grant status $g \in \{0,1\}$. Realized schooling choices S_g are a function of grant receipt. Finally, let $Z = \mathbf{1}\{k \leq k^*\}$ be an indicator for whether the student's EFC falls below the a grant threshold, and let $D_z \in \{0,1\}$ represent the student's potential grant status as a function of $z \in \{0,1\}$.

Equation 2 states that

$$\frac{\lim_{r\uparrow k^*} \mathbb{E}[Y \mid k=r] - \lim_{r\downarrow k^*} \mathbb{E}[Y \mid k=r]}{\lim_{r\uparrow k^*} \mathbb{E}[D \mid k=r] - \lim_{r\downarrow k^*} \mathbb{E}[D \mid k=r]} = \mathbb{E}[Y_{S_1,1} - Y_{S_0,0} \mid D_1 = 1, D_0 = 0, k = k^*]$$

To see this, note that we can write realized outcomes Y as:[46]

$$\begin{aligned} Y &= Y_{S_1,1} D + Y_{S_0,0}(1-D) \\ &= D\left(Y_{S_1,1} - Y_{S_0,0}\right) + Y_{S_0,0} \end{aligned}$$

The numerator of the limiting Wald Estimand is therefore:

$$\begin{aligned} &\lim_{r\uparrow k^*} \mathbb{E}[Y \mid k=r] - \lim_{r\downarrow k^*} \mathbb{E}[Y \mid k=r] \\ &= \lim_{r\uparrow k^*} \mathbb{E}\big[D(Y_{S_1,1} - Y_{S_0,0})\big] - \lim_{r\downarrow k^*} \mathbb{E}\big[D(Y_{S_1,1} - Y_{S_0,0})\big] \\ &\quad + \underbrace{\lim_{r\uparrow k^*} \mathbb{E}[Y_{S_0,0}] - \lim_{r\downarrow k^*} \mathbb{E}[Y_{S_0,0}]}_{=0} \end{aligned}$$

[46] To see this, start with Equation (1):

$$Y = S(Y_{1,1}D + Y_{1,0}(1-D)) + (1-S)(Y_{0,1}D + Y_{0,0}(1-D))$$

Plugging in $S = DS_1 + (1-D)S_0$ and rearranging gives:

$$Y = D(S_1 Y_{1,1} + (1-S_1)Y_{0,1}) + (1-D)(S_0 Y_{1,0} - Y_{0,0}(1-S_0))$$

which immediately yields the result.

where the canceled term comes from continuity of potential outcomes at the cutoff (Assumption 1). We can expand the first term of the numerator as:

$$
\begin{aligned}
&\lim_{r\uparrow k^*}\mathbb{E}\big[G(Y_{S_1,1}-Y_{S_0,0})\big]\\
&=\lim_{r\uparrow k^*}\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=C,k=r]P(T_g=C\mid k=r)\\
&\quad+\lim_{r\uparrow k^*}\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=AT,k=r]P(T_g=AT\mid k=r)\\
&=\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=C,k=k^*]P(T_g=C\mid k=k^*)\\
&\quad+\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=AT,k=k^*]P(T_g=AT\mid k=k^*)
\end{aligned}
$$

where $T_g=C$ is notational shorthand for $(D_1=1,D_0=0)$ (grant compliers) and $T_g=AT$ stands for $(D_1=D_0=1)$ (grant always-takers), and where the last equality again invokes the assumed continuity of the potential outcome and type composition functions at the cutoff. By a similar argument, the second term of the numerator is:

$$
\begin{aligned}
&\lim_{r\downarrow k^*}\mathbb{E}\big[D(Y_{S_1,1}-Y_{S_0,0})\big]\\
&=\lim_{r\downarrow k^*}\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=AT,k=r]P(T_g=AT\mid k=r)\\
&=\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=AT,k=k^*]P(T_g=AT,k=k^*)
\end{aligned}
$$

It follows that the numerator of the limiting Wald Estimand is:

$$
\begin{aligned}
&\lim_{r\uparrow k^*}\mathbb{E}[Y\mid k=r]-\lim_{r\downarrow k^*}\mathbb{E}[Y\mid k=r]\\
&=\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=C,k=k^*]P(T_g=C\mid k=k^*)
\end{aligned}
$$

By a similar argument, the denominator of the limiting Wald Estimand is given by:

$$
\lim_{r\uparrow k^*}\mathbb{E}[G\mid k=r]-\lim_{r\downarrow k^*}\mathbb{E}[G\mid k=r]=P(T_g=C\mid k=k^*) \tag{22}
$$

Together, these imply that:

$$
\frac{\lim_{r\uparrow k^*}\mathbb{E}[Y\mid k=r]-\lim_{r\downarrow k^*}\mathbb{E}[Y\mid k=r]}{\lim_{r\uparrow k^*}\mathbb{E}[G\mid k=r]-\lim_{r\downarrow k^*}\mathbb{E}[G\mid k=r]}=\mathbb{E}[Y_{S_1,1}-Y_{S_0,0}\mid T_g=C,k=k^*]
$$

which is the expression in Equation 2.

C.2 Decomposition into subLATEs (Equation 3)

To derive Equation 3, breaking down (2) into subLATEs, we use the law of total expectation to write:

$$
\begin{aligned}
&\mathbb{E}[Y_{S_1,1} - Y_{S_0,0} \mid T_g = C, k = k^*] \\
&= \mathbb{E}[Y_{1,1} - Y_{S_0,0} \mid S_1 = 1, T_g = C, k = k^*] P(S_1 = 1 \mid T_g = C, k = k^*) \\
&\quad + \mathbb{E}[Y_{0,1} - Y_{S_0,0} \mid S_1 = 0, T_g = C, k = k^*] P(S_1 = 0 \mid T_g = C, k = k^*) \\
&= \mathbb{E}[Y_{1,1} - Y_{S_0,0} \mid S_1 = 1, T_g = C, k = k^*] P(S_1 = 1 \mid T_g = C, k = k^*) \\
&\quad + \mathbb{E}[Y_{0,1} - Y_{0,0} \mid S_1 = 0, T_g = C, k = k^*] P(S_1 = 0 \mid T_g = C, k = k^*) \\
&= \mathbb{E}[Y_{1,1} - Y_{S_0,0} \mid S_1 = 1, T_g = C, k = k^*] P(S_1 = 1 \mid T_g = C, k = k^*) \\
&= \mathbb{E}[Y_{1,1} - Y_{1,0} \mid S_1 = S_0 = 1, T_g = C, k = k^*] P(S_1 = S_0 = 1 \mid T_g = C, k = k^*) \\
&\quad + \mathbb{E}[Y_{1,1} - Y_0 \mid S_1 > S_0, T_g = C, k = k^*] P(S_1 > S_0 \mid T_g = C, k = k^*)
\end{aligned}
$$

which is Equation 3. The first equality above comes directly from the law of total expectation. The second equality comes from Assumptions (4) (Enrollment Monotonicity). The third equality comes from Assumption (2), which states that $Y_{0,1} = Y_{0,0} = 0$. The final equality comes from the law of total probability.

C.3 Fuzzy RD in Enrolled Data is Biased (Equation 4)

Consider the selected-data Wald estimand:

$$
\frac{\lim_{k\uparrow k^*} \mathbb{E}[Y \mid S = 1, k] - \lim_{r\downarrow k} \mathbb{E}[Y \mid S = 1, k]}{\lim_{k\uparrow k^*} \mathbb{E}[D \mid S = 1, k] - \lim_{k\downarrow k^*} \mathbb{E}[D \mid S = 1, k]}
$$

First, consider the terms in the numerator. Using Assumptions 3 and 4, we can expand the limit of $\mathbb{E}[Y \mid S = 1, k]$ approaching the cutoff from the left (eligible) side into mean potential outcomes for each complier stratum:

$$
\begin{aligned}
&\lim_{k\uparrow k^*}\mathbb{E}[Y \mid S=1,k] \\
&\quad = \mathbb{E}[Y_{1,1} \mid D_1 > D_0, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 > D_0, S_1 = S_0 = 1 \mid k = k^{*-}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,1} \mid D_1 = D_0 = 1, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 = D_0 = 1, S_1 = S_0 = 1 \mid k = k^{*-}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,1} \mid D_1 > D_0, S_1 > S_0, S=1, k=k^*]P(D_1 > D_0, S_1 > S_0 \mid k = k^{*-}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,1} \mid D_1 = D_0 = 1, S_1 > S_0, S=1, k=k^*]P(D_1 = D_0 = 1, S_1 > S_0 \mid k = k^{*-}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,0} \mid D_1 = D_0 = 0, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 = D_0 = 0, S_1 = S_0 = 1 \mid k = k^{*-}, S=1)
\end{aligned} \tag{23}
$$

By a similar expansion, we can write $\lim_{k\downarrow k^*}\mathbb{E}[Y \mid S=1,k^*]$ as:

$$
\begin{aligned}
&\lim_{k\downarrow k^*}\mathbb{E}[Y \mid S=1,k] \\
&\quad = \mathbb{E}[Y_{1,0} \mid D_1 > D_0, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 > D_0, S_1 = S_0 = 1 \mid k = k^{*+}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,1} \mid D_1 = D_0 = 1, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 = D_0 = 1, S_1 = S_0 = 1 \mid k = k^{*+}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,1} \mid D_1 = D_0 = 1, S_1 > S_0, S=1, k=k^*]P(D_1 = D_0 = 1, S_1 > S_0 \mid k = k^{*+}, S=1) \\
&\qquad + \mathbb{E}[Y_{1,0} \mid D_1 = D_0 = 0, S_1 = S_0 = 1, S=1, k=k^*]P(D_1 = D_0 = 0, S_1 = S_0 = 1 \mid k = k^{*+}, S=1)
\end{aligned} \tag{24}
$$

As discussed in Appendix D, the notation k^{*+} and k^{*-} indicates that limits of the sample shares in enrollment-selected data are not equal at the cutoff. The reason is that extensive-margin compliers are present on the left (eligible) side of the cutoff but not on the right. To relate these shares to each other, note that by Bayes' rule we can write:

$$P(W \mid k = k^{*+}) = \frac{f_1(k^{*+} \mid W)P(W \mid S=1)}{f_1(k^{*+})} \qquad P(W \mid k = k^{*-}) = \frac{f_1(k^{*-} \mid W)P(W \mid S=1)}{f_1(k^{*-})}$$

where $f_1(k)$ is the density of the running variable in enrollment-selected data, and where W denotes a complier stratum defined by values of D_1, D_0, S_1, and S_0. Rearranging gives:

$$P(W \mid k = k^{*-}, S=1) = \frac{f_1(k^{*+})}{f_1(k^{*-})}P(W \mid k = k^{*+}, S=1) \tag{25}$$

Using (25) to re-write (23) and subtracting (24) yields, after some algebra:

$$
\begin{aligned}
&\lim_{k\uparrow k^*} \mathbb{E}[Y \mid S=1,k] - \lim_{k\downarrow k^*} \mathbb{E}[Y \mid S=1,k] \\
&= \underbrace{\mathbb{E}[Y_{1,1} - Y_{1,0} \mid D_1 > D_0, S_1 = S_0 = 1, S=1, k=k^*]}_{LATE_{IM}} \underbrace{P(D_1 > D_0, S_1 = S_0 = 1 \mid S=1, k=k^*)}_{\pi^-_{IM}} \\
&\quad + \mathbb{E}[Y_{1,1} \mid \underbrace{D_1 > D_0, S_1 > S_0}_{\text{EM Complier}}, S_1, k=k^*] \underbrace{P(D_1 > D_0, S_1 > S_0 \mid k = k^{*-}, S=1)}_{\pi^-_{EM}} \\
&\quad - \underbrace{\left(1 - \frac{f_1(k^{*+})}{f_1(k^{*-})}\right)}_{(*)} \left(\lim_{k\downarrow k^*} \mathbb{E}[Y \mid S=1,k]\right)
\end{aligned}
$$

In Appendix D.1.1 I prove that the starred term $(*)$ equals $\pi^-_{EM} = P(D_1 > D_0, S_1 > S_0 \mid k = k^{*-}, S=1)$. Using this result, the numerator becomes:

$$
LATE_{IM}\pi^-_{IM} + \left(\mathbb{E}[Y_{1,1} \mid \text{EM Complier}] - \lim_{k\downarrow k^*} \mathbb{E}[Y \mid S=1,k]\right) \pi^-_{EM}
$$

Now consider the denominator, $\lim_{k\uparrow k^*} \mathbb{E}[D \mid S=1,k] - \lim_{k\downarrow k^*} \mathbb{E}[D \mid S=1,k]$. Writing $\bar{D}^+$ and $\bar{D}^-$ as shorthand for these limits, note that we can write:

$$
\bar{D}^+ - \bar{D}^- = \underbrace{\left(\bar{D}^- - \frac{f_1(k^{*+})}{f_1(k^{*-})}\bar{D}^-\right)}_{(\dagger)} - \underbrace{\left(1 - \frac{f_1(k^{*+})}{f_1(k^{*-})}\right)}_{(*)} \bar{D}^+ \tag{26}
$$

Appendix D.1 shows that the term $(\dagger)$ equals π^-_{IM}, and the term $(*)$ equals π^-_{EM}. As a result, we have:

$$
\bar{D}^+ - \bar{D}^- = \pi^-_{IM} - \pi^-_{EM}(1 + \bar{D}^+) \tag{27}
$$

Dividing (26) by (27) yields Equation (4) in the main text, completing the proof.

C.4 Derivation of Marginal Cost and Benefit formulas (Equations 18 and 19)

Start with the definitions in Section 9.1:

$$
\begin{aligned}
B &= \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}[\text{Grants}_t + \text{Loans}_t + (1-\tau)\text{Earnings}_t - \mathbf{1}\{\text{Enrolled}\}_t \times \delta_{\text{tuition}}] \\
C &= \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}[\text{Grants}_t + \text{Loans}_t - \tau\text{Earnings}_t + \mathbf{1}\{\text{Enrolled}\}_t \times \delta_{\text{instruction}}]
\end{aligned}
$$

Taking $\partial B/\partial k^*$ or $\partial C/\partial k^*$ boils down to deriving $\partial \mathbb{E}[Y_t]/\partial k^*$, where Y_t is an outcome in $\{\text{Grants}_t, \text{Loans}_t, \text{Earnings}_t, \mathbf{1}\{\text{Enrolled}\}_t\}$. To take this derivative, note first that $\mathbb{E}[Y_t]$ can be written:

$$
\begin{aligned}
\mathbb{E}[Y_t] = \mathbb{E}[Y_t(S(D(Z)), D(Z))] &= \int_{\underline{k}}^{\bar{k}} \mathbb{E}[Y_t(S(D(Z)), D(Z)) \mid k] f(k) dk \\
&= \int_{\underline{k}}^{k^*} \mathbb{E}[Y_t(S(D_1), D_1) \mid k] f(k) dk + \int_{k^*}^{\bar{k}} \mathbb{E}[Y_t(S(D_0), D_0) \mid k] f(k) dk
\end{aligned}
$$

where the last equality comes from the law of iterated expectations. It follows by the Fundamental Theorem of Calculus that:

$$
\begin{aligned}
\frac{\partial \mathbb{E}[Y_t]}{\partial k^*} &= \mathbb{E}[Y_t(S(D_1), D_1) - Y_t(S(D_0), D_0) \mid k^*] f(k^*) \\
&= \mathbb{E}[Y_t(S(D_1), D_1) - Y_t(S(D_0), D_0) \mid D_1 > D_0, k^*] P(D_1 > D_0 \mid k^*) f(k^*) \\
&= \underbrace{\mathbb{E}[Y_t(S_1, 1) - Y_t(S_0, 0) \mid D_1 > D_0, k^*]}_{LATE} \cdot P(D_1 > D_0 \mid k^*) \cdot f(k^*) \\
&= \underbrace{\mathbb{E}[Y_t(S_1, 1) - Y_t(S_0, 0) \mid S_1 = 1, D_1 > D_0, k^*]}_{LATE^*} \cdot P(S_1 = 1, D_1 > D_0 \mid k^*) \cdot f(k^*)
\end{aligned}
$$

where the second equality comes from the fact that the difference inside the expectation can only be nonzero if $D_1 > D_0$, and where the third equality comes from Assumption 3 (Grant Monotonicity). The fourth equality comes from the law of total expectation and from Assumptions 2 and 4 (Partial Exclusion and Enrollment Monotonicity), which require that if $S_1 = 0$ then $S_0 = 0$, implying $Y_{0,1} = Y_{0,0} = Y_0$, so that the LATE for units with $S_1 = 0$ is zero.

Applying this result to each outcome and plugging the result back into the expression for B gives:

$$
\begin{aligned}
\frac{\partial B}{\partial k^*} = \Bigg[\sum_{t=1}^{T} \beta^{t-1} \Big(& LATE^*(\text{Grants}_t) + LATE^*(\text{Loans}_t) + (1-\tau) LATE^*(\text{Earnings}_t) \\
& + \delta_{\text{tuition}} \times LATE^*(\mathbf{1}\{\text{Enrolled}\}_t) \Big) \Bigg] \times P(S_1 = 1, D_1 > D_0 \mid k^*) f(k^*)
\end{aligned}
$$

Now consider the enrollment term. At $t = 1$, the enrollment indicator is equal to S, the student's initial enrollment choice. We can expand $LATE^*(\text{Enrolled}_t)$ into:

$$
LATE^*(\text{Enrolled}_t) = (1-\omega) LATE_{IM}(\text{Enrolled}_t) + \omega LATE_{EM}(\text{Enrolled}_t)
$$

where $\omega = P(S_1 > S_0 \mid S_1 = 1, D_1 > D_0, k = k^*)$ is the share of extensive-margin compliers among all enrolled compliers. At $t = 1$, we have by definition that $LATE_{IM}(\text{Enrolled}_1) = 0$ and $LATE_{EM}(\text{Enrolled}_1) = 1$, because intensive-margin compliers are students who would have enrolled in college regardless of treatment and extensive-margin compliers are brought into initial college enrollment by the treatment. This allows us to finally write:

$$\frac{\partial B}{\partial k^*} = \Bigg[\sum_{t=1}^{T} \beta^{t-1} \left(LATE^*(\text{Grants}_t) + LATE^*(\text{Loans}_t) + (1-\tau) LATE^*(\text{Earnings}_t) \right) \\ + \left(\omega + \sum_{t=2}^{T} \beta^{t-1} LATE^*(\mathbf{1}\{\text{Enrolled}\}_t) \right) \times \delta_{\text{tuition}} \Bigg] \times P(S_1 = 1, D_1 > D_0 \mid k^*) f(k^*)$$

The proof for $\partial C / \partial k^*$ proceeds similarly.

D Identification with Selected Data

This appendix shows how to derive important quantities with data subject to sample selection. I first show point-identification of the following sample shares:

$$P(\text{Treated Complier} \mid S = 1, k = k^{*-}) \equiv P(S_1 = 1, D_1 > D_0 \mid S = 1, k = k^{*-})$$
$$P(\text{IM Complier} \mid S = 1, k = k^{*-}) \equiv P(S_1 = S_0 = 1, D_1 > D_0 \mid S = 1, k = k^{*-})$$
$$P(\text{EM Complier} \mid S = 1, k = k^{*-}) \equiv P(S_1 > S_0, D_1 > D_0 \mid S = 1, k = k^{*-})$$

where the notation $k = k^{*-}$ denotes limits of the relevant quantity approaching the cutoff from the left. Then I show how to point-identify the following potential outcome means:

$$\mathbb{E}[Y_{1,1} \mid \text{Treated Complier}] \equiv \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*]$$
$$\mathbb{E}[Y_{1,0} \mid \text{IM Complier}] \equiv \mathbb{E}[Y_{1,0} \mid S_1 = 1, S_0 = 1, D_1 > D_0, k = k^*]$$

The selection problem is that the running variable is not available for the $S = 0$ sample. As a result, it is not possible to derive the full-sample analogues of sample shares. Instead, I show how to derive sample shares in the selected sample of four-year students immediately to the left of the cutoff, and show how the data are still sufficient to point-identify certain potential outcome means in the full sample.

The identification proofs in this section measure selection behavior using the densities of the running variable k in various samples. Throughout, I assume that the full-sample distribution of k, written $f(k)$, is continuous at the cutoff k^*. However, this full-sample distribution is not observed because data is missing for students who forego college. Instead, we observe the distribution of the running variable in selected samples:

$$f(k) = \underbrace{f_1(k)}_{\text{observed}} \times P(S=1) + \underbrace{f_0(k)}_{\text{observed}} \times P(S=0)$$

where $f_s(k) = f(k \mid S = s)$.

As a result of this selection problem, while Assumption (1) ensured continuity of potential outcomes and potential college choices in the full sample, this continuity does not necessarily hold in selected samples, as crossing the threshold leads some students to leave one sample and enter another. The proofs in this appendix leverage the fact that discontinuities in the densities of k selected samples in fact *measure* these selection responses to grant receipt. Throughout, I use the notation k^{*+} and k^{*-} to denote limits of the relevant quantity approaching the cutoff from the right and the left.

Some of these identification arguments extend proofs originally derived by Gerard et al. (2020), who derive bounds for treatment effects in regression discontinuity designs subject to manipulation. In contrast to their framework, which derives bounds for non-manipulated units in a setting where some units on the treated side of the cutoff are manipulated, my framework assumes that the full-sample distribution of the running variable is not manipulated, but uses a similar limiting-densities argument to address sample selection.[47]

D.1 Point-identification of population shares from densities

Identifying the mass of treated compliers

To identify $P(\text{Treated Complier} \mid S = 1, k = k^{*-}) \equiv P(S_1 = 1, D_1 > D_0 \mid S = 1, k = k^{*-})$, let $\bar{D}^+$ and $\bar{D}^-$ denote the share of treated students to the left and right of the cutoff in the enrolled sample:

$$\bar{D}^+ \equiv \lim_{k \downarrow k^*} \mathbb{E}[G \mid S = 1, Z = 0] = \mathbb{E}[G \mid S = 1, k = k^{*+}]$$

$$\bar{D}^- \equiv \lim_{k \uparrow k^*} \mathbb{E}[G \mid S = 1, Z = 1] = \mathbb{E}[G \mid S = 1, k = k^{*-}]$$

Let $f_1(k) = f(k \mid S = 1)$ denote the density of EFC conditional on enrollment. Let $f_1(k^{*-})$ and $f_1(k^{*+})$ denote the limits of this density approaching the cutoff from the left and right. To the right of the cutoff we have $D = D_0$, and to the left we have $D = D_1$.

[47] Here is a mapping between concepts in Gerard et al. (2020) (GRR) and this paper. The full sample in GRR is analogous to the $S = 1$ sample in my paper, and the $M = 1$ group in their paper corresponds to the extensive margin compliers ($S_1 > S_0, D_1 > D_0$) in this paper. Furthermore, in the notation of their paper, $\tau_0 = 0$ in my setting because there are no “manipulated” untreated units to the left of the cutoff; students only enter the enrolled sample as a result of being awarded a grant.

All grant recipients to the right of the cutoff are grant always-takers, allowing us to expand $\bar{D}^+$ as follows:

$$\begin{aligned} \bar{D}^+ &\equiv \lim_{k \downarrow k^*} \mathbb{E}[D \mid S = 1] \\ &= P(D_0 = 1 \mid S = 1, k = k^{*+}) \\ &= P(D_0 = D_1 = 1 \mid S = 1, k = k^{*+}) \\ &= P(D_0 = D_1 = 1, S_1 = 1 \mid S = 1, k = k^{*+}) \\ &= \frac{f_1(k^{*+} \mid S_1 = 1, D_1 = D_0 = 1) P(S_1 = 1, D_1 = D_0 = 1 \mid S = 1)}{f_1(k^{*+})} \end{aligned} \tag{28}$$

where the third equation comes from the grant monotonicity assumption, the fourth equation is by the law of total probability, and the fifth equation is by Bayes' theorem.

In contrast, grant recipients to the left of the cutoff are a mix of grant always-takers and grant compliers:

$$\begin{aligned} \bar{D}^- &\equiv \lim_{k \uparrow k^*} \mathbb{E}[G \mid S = 1] \\ &= P(D_1 = 1 \mid S = 1, k = k^{*-}) \\ &= P(D_1 > D_0 \mid S = 1, k = k^{*-}) + P(D_1 = D_0 = 1 \mid S = 1, k = k^{*-}) \\ &= P(D_1 > D_0, S_1 = 1 \mid S = 1, k = k^{*-}) + P(D_1 = D_0 = 1, S_1 = 1 \mid S = 1, k = k^{*-}) \\ &= \frac{1}{f_1(k^{*-})} \Big(f_1(k^{*-} \mid S_1 = 1, D_1 > D_0) P(S_1 = 1, D_1 > D_0 \mid S = 1) \\ &\qquad + f_1(k^{*-} \mid S_1 = 1, D_1 = D_0 = 1) P(S_1 = 1, D_1 = D_0 = 1 \mid S = 1) \Big) \end{aligned} \tag{29}$$

Combining and rearranging (28) and (29) yields:

$$\begin{aligned} \bar{D}^- - \frac{f_1(k^{*+})}{f_1(k^{*-})} \bar{D}^+ &= \frac{f_1(k^{*-} \mid S_1 = 1, D_1 > D_0) P(S_1 = 1, D_1 > D_0 \mid S = 1)}{f_1(k^{*-})} \\ &= P(S_1 = 1, D_1 > D_0 \mid S = 1, k = k^{*-}) \\ &\equiv P(\text{Treated Complier} \mid S = 1, k = k^{*-}) \end{aligned} \tag{30}$$

where the first equation uses the fact that $f_1(k^{*-} \mid S_1 = 1, D_1 = D_0 = 1) = f_1(k^{*+} \mid S_1 = 1, D_1 = D_0 = 1)$,[48] and the second equation comes from Bayes' theorem.

[48] To see this, use Bayes' theorem to write:

$$\begin{aligned} f_1(k^{*-} \mid S_1 = 1, D_1 = D_0 = 1) &= \frac{\overbrace{P(S = 1 \mid S_1 = 1, D_1 = D_0 = 1)}^{=1} \overbrace{f(k^{*-} \mid S_1 = 1, D_1 = D_0 = 1)}^{= f(k^* \mid S_1 = 1, D_1 = D_0 = 1)}}{P(S = 1 \mid S_1 = 1, D_1 = D_0 = 1)} \\ &= \frac{f(k^* \mid S_1 = 1, D_1 = D_0 = 1)}{P(S = 1 \mid S_1 = 1, D_1 = D_0 = 1)} \end{aligned}$$

Equation (30) identifies the fraction of the four-year sample directly to the left of the cutoff who are grant compliers. We can expand this population into the two complier margins:

$$\begin{aligned}
P(\text{Treated Complier} \mid S=1, k=k^{*-}) &\equiv P(S_1=1, D_1>D_0 \mid S=1, k=k^{*-}) \\
&= P(S_1=S_0=1, D_1>D_0 \mid S=1, k=k^{*-}) \\
&\quad + P(S_1>S_0, D_1>D_0 \mid S=1, k=k^{*-}) \\
&= P(\text{IM Complier} \mid S=1, k=k^{*-}) \\
&\quad + P(\text{EM Complier} \mid S=1, k=k^{*-})
\end{aligned} \tag{31}$$

D.1.1 Identifying the mass of extensive-margin compliers

To identify $P(\text{EM Complier} \mid S=1, k=k^{*-})$, note that under Assumptions 3 and 4, we can break the enrolled sample on the right (ineligible) side of the cutoff into four subpopulations defined by possible combinations of (S_0, S_1, D_0, D_1):[49]

$$\begin{aligned}
f_1(k^{*+}) = &\overbrace{f_1(k^{*+} \mid S_1>S_0, D_1>D_0)}^{=0} P(S_1>S_0, D_1>D_0 \mid S=1) \\
&+ f_1(k^{*+} \mid S_0=S_1=1, D_1>D_0) P(S_1=S_0=1, D_1>D_0 \mid S=1) \\
&+ f_1(k^{*+} \mid S_0=S_1=1, D_1=D_0=0) P(S_0=S_1=1, D_1=D_0=0 \mid S=1) \\
&+ f_1(k^{*+} \mid S_1=1, D_1=D_0=1) P(D_1=D_0=1 \mid S=1)
\end{aligned}$$

where $f_1(k^{*+} \mid S_1>S_0, D_1>D_0)=0$ because extensive-margin grant compliers only enter the four-year sample if they receive grants, which only happens if they are on the left of the cutoff.

where the second equation comes from Assumption (1), as $f(k^{*-} \mid S_1=1, D_1=D_0=1)$ is defined on the full sample, not on a selected sample. Finally, note that we would get the an equal result expanding $f_1(k^{*+} \mid S_1=1, D_1=D_0=1)$ the same way.

[49] They are:

- $D_1>D_0, S_1>S_0$: Extensive-margin grant compliers. These students receive a grant if and only if they cross the grant threshold ($D_1>D_0$), and attend four-year schools if and only if they receive a grant ($S_1>S_0$). For this reason, they appear only on the left side of the cutoff in the $S=1$ sample, not on the right.
- $D_1>D_0, S_0=S_1=1$: Intensive-margin grant compliers. These students do not receive a grant because they are on the right of the threshold, but would receive a grant if moved to the left of the threshold ($D_1>D_0$). However, they attend four-year college despite not being offered a grant ($S_0=S_1=1$).
- $D_1=D_0=0, S_0=S_1=1$: Intensive-margin grant never-takers. These students do not receive a grant regardless of whether they cross the threshold ($D_0=D_1=0$). However, they attend four-year college regardless of grant receipt ($S_1=S_0=1$).
- $D_1=D_0=1, S_1=1$: Intensive-margin grant always-takers. These students receive a grant despite being on the "wrong" side of the cutoff ($D_0=1$), and enroll in college ($S_1=1$). Note that this population can be broken down further into students with $S_0=1$ and students with $S_0 \neq 1$, but this decomposition is not necessary for the proof.

The expansion of $f_1(k^{*-})$ incorporates extensive-margin grant compliers:

$$\begin{aligned}
f_1(k^{*-}) = {} & f_1(k^{*-} \mid S_1 > S_0, D_1 > D_0) P(S_1 > S_0, D_1 > D_0 \mid S = 1) \\
& + f_1(k^{*-} \mid S_0 = S_1 = 1, D_1 > D_0) P(S_1 = S_0 = 1, D_1 > D_0 \mid S = 1) \\
& + f_1(k^{*-} \mid S_0 = S_1 = 1, D_1 = D_0 = 0) P(S_0 = S_1 = 1, D_1 = D_0 = 0 \mid S = 1) \\
& + f_1(k^{*-} \mid S_1 = 1, D_1 = D_0 = 1) P(D_1 = D_0 = 1 \mid S = 1)
\end{aligned}$$

Letting W denote groups of students, note that $f(k^{*-} \mid W, S = 1) = f(k^{*+} \mid W, S = 1)$ for each of the principal strata $W \in \{(S_0 = S_1 = 1, D_1 > D_0), (S_0 = S_1 = 1, D_1 = D_0 = 0), (S_1 = 1, D_1 = D_0 = 1)\}$. To see this, note that by Bayes' theorem, we can write both $f(k^{*+})$ and $f(k^{*-})$ as:

$$f(k^{*(\text{sgn})} \mid W, S = 1) = \frac{\overbrace{P(S = 1 \mid W, k = k^{*(\text{sgn})})}^{=1} \overbrace{f(k^{*(\text{sgn})} \mid W)}^{=f(k^*|W)}}{P(S = 1 \mid W)} = \frac{f(k^* \mid W)}{P(S = 1 \mid W)}$$

which holds for both $(\text{sgn}) \in \{+, -\}$ and shows that $f(k^{*+} \mid W, S = 1) = f(k^{*-} \mid W, S = 1)$ for all $W \in \{(S_0 = S_1 = 1, D_1 > D_0), (S_0 = S_1 = 1, D_1 = D_0 = 0), (S_1 = 1, D_1 = D_0 = 1)\}$. The same equality does not hold for $W = (S_1 > S_0, D_1 > D_0)$, because $P(S = 1 \mid S_1 > S_0, D_1 > D_0, k = k^{*-}) = 0$ while $P(S = 1 \mid S_1 > S_0, D_1 > D_0, k = k^{*-}) = 1$ due to the fact that extensive-margin compliers appear only to the left of the cutoff.

As a result, differencing the conditional densities in the enrolled sample across the cutoff yields:

$$f_1(k^{*-}) - f_1(k^{*+}) = f_1(k^{*-} \mid S_1 > S_0, D_1 > D_0) P(S_1 > S_0, D_1 > D_0 \mid S = 1)$$

It follows from Bayes' theorem that:

$$\begin{aligned}
\frac{f_1(k^{*-}) - f_1(k^{*+})}{f_1(k^{*-})} &= P(S_1 > S_0, D_1 > D_0 \mid S = 1, k = k^{*-}) \\
&= P\left(\text{EM Complier} \mid S = 1, k = k^{*-}\right)
\end{aligned} \tag{32}$$

which identifies the mass of extensive-margin compliers in the enrolled sample to the left of the cutoff.

Identifying the mass of intensive-margin compliers

Combining (31) and (32) immediately yields the mass of intensive-margin compliers:

$$P(\text{IM Complier} \mid S=1, k=k^{*-}) = \underbrace{\left(\bar{D}^{-} - \frac{f_1(k^{*+})}{f_1(k^{*-})}\bar{D}^{+}\right)}_{\text{Treated Compliers}} - \underbrace{\frac{f_1(k^{*-}) - f_1(k^{*+})}{f_1(k^{*-})}}_{\text{EM Compliers}} \tag{33}$$

With the sample shares in hand, it remains to identify the counterfactual means.

D.2 Point-Identification of Counterfactual Means

We wish to identify:

- $\mathbb{E}[Y_{1,0} \mid S_1 = 1, S_0 = 1, D_1 > D_0, k = k^*]$
- $\mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*]$

Both sides of the cutoff contain grant-receiving and non-grant-receiving students ($D \in \{0,1\}$). We therefore observe four conditional mean outcomes directly from the data:

$$\bar{Y}_{1,1}^{+} \equiv \lim_{k \downarrow k^*} \mathbb{E}[Y \mid S=1, D=1, Z=0]$$
$$\bar{Y}_{1,0}^{+} \equiv \lim_{k \downarrow k^*} \mathbb{E}[Y \mid S=1, D=0, Z=0]$$
$$\bar{Y}_{1,1}^{-} \equiv \lim_{k \uparrow k^*} \mathbb{E}[Y \mid S=1, D=1, Z=1]$$
$$\bar{Y}_{1,0}^{-} \equiv \lim_{k \uparrow k^*} \mathbb{E}[Y \mid S=1, D=0, Z=1]$$

as well as the share of treated students to the left and right of the cutoff:

$$\bar{D}^{+} \equiv \lim_{k \downarrow k^*} \mathbb{E}[D \mid S=1, Z=0] = \mathbb{E}[D \mid S=1, k=k^{*+}]$$
$$\bar{D}^{-} \equiv \lim_{k \uparrow k^*} \mathbb{E}[D \mid S=1, Z=1] = \mathbb{E}[D \mid S=1, k=k^{*-}]$$

and the limiting densities $f_1(k^{*+})$ and $f_1(k^{*+})$ of the running variable on either side of the cutoff.

First, note that grant monotonicity (Assumption 3) implies that crossing the grant threshold can only increase the probability that a student receives a grant. Thus, all of the units to the right of the threshold who receive grants ($D = 1, Z = 0$) are grant always-takers, and all of the units to the left of the threshold who do not receive grants ($D = 0, Z = 1$) are grant never-takers. Futhermore, any student who enrolls not receiving a grant ($S = 1, D = 0$) must have ($S_1 = S_0 = 1$) by Assumption 4. Therefore $\bar{Y}_{1,0}^{-}$ and $\bar{Y}_{1,1}^{+}$ identify potential outcomes for enrolled grant never-takers and enrolled grant always-takers:

$$
\begin{aligned}
\bar{Y}_{1,0}^{-} &\equiv \lim_{k\uparrow k^*} \mathbb{E}[Y \mid S=1, D=0, Z=1] \\
&= \lim_{k\uparrow k^*} \mathbb{E}[Y_{1,0} \mid S=1, D=0, Z=1] \\
&= \lim_{k\uparrow k^*} \mathbb{E}[Y_{1,0} \mid S_0=1, D_1=0] \\
&= \mathbb{E}[Y_{1,0} \mid S_0=1, D_1=D_0=0, k=k^*] && \text{(by continuity and grant monotonicity)} \\
&= \mathbb{E}[Y_{1,0} \mid S_0=S_1=1, D_1=D_0=0, k=k^*] && \text{(by enrollment monotonicity)}
\end{aligned}
$$

and:

$$
\begin{aligned}
\bar{Y}_{1,1}^{+} &\equiv \lim_{k\downarrow k^*} \mathbb{E}[Y \mid S=1, D=1, Z=0] \\
&= \lim_{k\downarrow k^*} \mathbb{E}[Y_{1,1} \mid S=1, D=1, Z=0] \\
&= \lim_{k\downarrow k^*} \mathbb{E}[Y_{1,1} \mid S_1=1, D_0=1] \\
&= \mathbb{E}[Y_{1,1} \mid S_1=1, D_1=D_0=1, k=k^*] && \text{(by continuity and grant monotonicity)}
\end{aligned}
$$

In contrast, untreated mean outcomes for four-year students to the right of the cutoff are a weighted average across grant compliers and grant never-takers:

$$
\begin{aligned}
\bar{Y}_{1,0}^{+} &\equiv \lim_{k\downarrow k^*} \mathbb{E}[Y \mid S=1, D=0, Z=0] \\
&= \lim_{k\downarrow k^*} \mathbb{E}[Y_{1,0} \mid S_0=1, D_0=0, Z=0] \\
&= \mathbb{E}[Y_{1,0} \mid S_0=S_1=1, D_0=D_1=0, k=k^*]P(D_1=0 \mid S_0=1, D_0=0, S=1, k=k^{*+}) \\
&\quad + \mathbb{E}[Y_{1,0} \mid S_0=S_1=1, D_1>D_0, k=k^*]P(D_1=1 \mid S_0=1, D_0=0, S=1, k=k^{*+}) \\
&= \bar{Y}_{1,0}^{-}P(D_1=D_0=0 \mid S_0=1, D_0=0, S=1, k=k^{*+}) \\
&\quad + \mathbb{E}[Y_{1,0} \mid S_0=S_1=1, D_1>D_0, k=k^*]P(D_1>D_0 \mid S_0=1, D_0=0, S=1, k=k^{*+}) \\
&= \bar{Y}_{1,0}^{-}(1-P(D_1>D_0 \mid S_0=1, D_0=0, S=1, k=k^{*+})) \\
&\quad + \mathbb{E}[Y_{1,0} \mid S_0=S_1=1, D_1>D_0, k=k^*]P(D_1>D_0 \mid S_0=1, D_0=0, S=1, k=k^{*+})
\end{aligned}
$$

where the fourth equation used grant monotonicity. This implies that untreated mean outcomes for intensive-margin grant compliers are:

$$\begin{aligned}
&\mathbb{E}[Y_{1,0} \mid S_0 = S_1 = 1, D_1 > D_0, k = k^*] \\
&\quad = \frac{\bar{Y}_{1,0}^{+} - \bar{Y}_{1,0}^{-}(1 - P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^*))}{P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^*)}
\end{aligned} \tag{34}$$

Similarly, treated mean outcomes to the left of the cutoff are a weighted average of grant compliers and grant always-takers:

$$\begin{aligned}
\bar{Y}_{1,1}^{-} &\equiv \lim_{k \uparrow k^*} \mathbb{E}[Y \mid S = 1, D = 1, Z = 1] \\
&= \lim_{k \uparrow k^*} \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 = 1, Z = 1] \\
&= \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 = D_0 = 1, k = k^*] P(D_0 = 1 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*) \\
&\quad + \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*] P(D_0 = 0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*) \\
&= \bar{Y}_{1,1}^{+}(1 - P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*)) \\
&\quad + \mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*] P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*)
\end{aligned}$$

Implying that treated mean outcomes for grant compliers are:

$$\begin{aligned}
&\mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*] \\
&\quad = \frac{\bar{Y}_{1,1}^{-} - \bar{Y}_{1,1}^{+}(1 - P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*))}{P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*)}
\end{aligned} \tag{35}$$

It remains to identify $P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, k = k^*)$ and $P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^*)$. These probabilities are identified by the treated shares to either side of the cutoff $\bar{D}^+$ and $\bar{D}^-$, together with the limiting EFC densities $f(k^{*+})$ and $f(k^{*-})$. To identify $P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^{*+})$, note first that we can write:

$$\begin{aligned}
(1 - \bar{D}^+) &\equiv \lim_{k \downarrow k^*} \mathbb{E}[(1 - D) \mid S = 1] \\
&= P(D_0 = 0 \mid S = 1, k = k^{*+}) \\
&= P(D_1 > D_0 \mid S = 1, k = k^{*+}) + P(D_1 = D_0 = 0 \mid S = 1, k = k^{*+}) \\
&= P(D_1 > D_0, S_1 = S_0 = 1 \mid S = 1, k = k^{*+}) \\
&\quad + P(D_1 = D_0 = 0, S_1 = S_0 = 1 \mid S = 1, k = k^{*+}) \\
&= \frac{1}{f_1(k^{*+})} \Big(f_1(k^{*+} \mid S_1 = S_0 = 1, D_1 > D_0) P(S_1 = S_0 = 1, D_1 > D_0 \mid S = 1) \\
&\qquad + f_1(k^{*+} \mid S_1 = S_0 = 1, D_1 = D_0 = 0) P(S_1 = S_0 = 1, D_1 = D_0 = 0 \mid S = 1) \Big)
\end{aligned}$$

where the third equality comes from expanding into the $D_1 = 1$ and $D_1 = 0$ cases, and the fourth equality comes from recognizing that $D_0 = 0$ and $S = 1$ imply $S_0 = S_1 = 1$.[50] The fifth equality is a direct application of Bayes' theorem.

We can use an analogous argument to write:

$$(1 - \bar{D}^-) = \frac{1}{f_1(k^{*-})} f_1(k^{*-} \mid S_1 = S_0 = 1, D_1 = D_0 = 0) P(S_1 = S_0 = 1, D_1 = D_0 = 0 \mid S = 1)$$

Re-arranging these two results gives us:

$$\begin{aligned}
(1 - \bar{D}^+) - \frac{f_1(k^{*-})}{f_1(k^{*+})}(1 - \bar{D}^-) &= \frac{f_1(k^{*+} \mid S_1 = S_0 = 1, D_1 > D_0) P(S_1 = S_0 = 1, D_1 > D_0 \mid S = 1)}{f_1(k^{*+})} \\
&= P(S_1 = S_0 = 1, D_1 > D_0 \mid S = 1, k = k^{*+}) \\
&= P(S_1 = S_0 = 1, D_1 > D_0 \mid D_0 = 0, S = 1, k = k^{*+}) P(D_0 = 0 \mid S = 1, k = k^{*+}) \\
&\quad + \underbrace{P(S_1 = S_0 = 1, D_1 > D_0 \mid D_0 = 1, S = 1, k = k^{*+})}_{=0} P(D_0 = 1 \mid S = 1, k = k^{*+})
\end{aligned}$$

where the first equation comes from the fact that $f_1(k^{*-} \mid S_1 = S_0 = 1, D_1 = D_0 = 0) = f_1(k^{*+} \mid S_1 = S_0 = 1, D_1 = D_0 = 0)$ by Assumption 1. The second equation again comes from Bayes' theorem, and the third comes from the law of total probability. Further dividing both sides by $(1 - \bar{D}^+) = P(D_0 = 0 \mid S = 1, k = k^*)$ gives:

$$\begin{aligned}
1 - \frac{f_1(k^{*-})}{f_1(k^{*+})} \frac{(1 - \bar{D}^-)}{(1 - \bar{D}^+)} &= P(S_1 = S_0 = 1, D_1 > D_0 \mid D_0 = 0, S = 1, k = k^{*+}) \\
&= P(S_0 = 1, D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^{*+}) \\
&= P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^{*+})
\end{aligned}$$

where the second equality comes from Assumption 4 and from the fact that $(D_0 = 0, S = 1, k = k^{*+})$ imply $S_0 = 1$.

[50] To see this, expand one of the terms into the $S_0 = 1$ and $S_0 \neq 1$ cases:

$$P(D_1 > D_0 \mid S = 1, k = k^{*+}) = P(D_1 > D_0, S_0 = 1 \mid S = 1, k = k^{*+}) + \underbrace{P(D_1 > D_0, S_0 \neq 1 \mid S = 1, k = k^{*+})}_{=0}$$

where the second term is zero because ($k = k^{*+}$ implies $D = D_0$), ($D_1 > D_0$ implies $D_0 = 0$), and therefore $S = S_0 = 1$. A similar argument applies to the $P(D_1 = D_0 \mid S = 1, k = k^{*+})$ term.

The derivation of $P(D_0 = 0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^{*-})$ proceeds similarly. Starting with Equation (30) we have:

$$\begin{aligned}
\bar{D}^- - \frac{f_1(k^{*+})}{f_1(k^{*-})}\bar{D}^+ &= P(D_1 > D_0, S_1 = 1 \mid S = 1, k = k^{*-}) \\
&= P(D_1 > D_0, S_1 = 1 \mid D_1 = 1, S = 1, k = k^{*-})P(D_1 = 1 \mid S = 1, k = k^{*-}) \\
&\quad + \underbrace{P(D_1 > D_0, S_1 = 1 \mid D_1 = 0, S = 1, k = k^{*-})}_{=0} P(D_1 = 0 \mid S = 1, k = k^{*-})
\end{aligned}$$

Dividing both sides by $\bar{D}^- = P(D_1 = 1 \mid S = 1, k = k^{*-})$ gives:

$$\begin{aligned}
1 - \frac{f_1(k^{*+})}{f_1(k^{*-})}\frac{\bar{D}^+}{\bar{D}^-} &= P(D_1 > D_0, S_1 = 1 \mid D_1 = 1, S = 1, k = k^{*-}) \\
&= P(D_1 > D_0, S_1 = 1 \mid S_1 = 1, D_1 = 1, S = 1, k = k^{*-}) \\
&= P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^{*-})
\end{aligned}$$

where the second equation comes from the fact that $(D_1 = 1, S = 1, k = k^{*-})$ imply $S_1 = 1$.

Armed with expressions for $P(D_1 > D_0 \mid S_0 = 1, D_0 = 0, S = 1, k = k^{*+})$ and $P(D_1 > D_0 \mid S_1 = 1, D_1 = 1, S = 1, k = k^{*-})$ in terms of observable quantities, we can now write Equations (34) and (35) as:

$$\begin{aligned}
\mathbb{E}[Y_{1,0} \mid S_0 = S_1 = 1, D_1 > D_0, k = k^*] &= \frac{\bar{Y}^+_{1,0} - \bar{Y}^-_{1,0}\kappa_0}{1 - \kappa_0} \\
\mathbb{E}[Y_{1,1} \mid S_1 = 1, D_1 > D_0, k = k^*] &= \frac{\bar{Y}^-_{1,1} - \bar{Y}^+_{1,1}\kappa_1}{1 - \kappa_1}
\end{aligned}$$

where:

$$\kappa_0 = \frac{f_1(k^{*-})}{f_1(k^{*+})}\frac{(1 - \bar{D}^-)}{(1 - \bar{D}^+)} \qquad \kappa_1 = \frac{f_1(k^{*+})}{f_1(k^{*-})}\frac{\bar{D}^+}{\bar{D}^-}$$

which completes the proof.