College Merit-Aid and Academic Renewal Thresholds: Evidence from Interactions of Discontinuous Rules

Ben Blemings*

This Version: Friday 6th November, 2020, Most Recent Version: Link

PRELIMINARY DRAFT:

PLEASE DO NOT CITE, SHARE, OR CIRCULATE

Abstract

To increase college completion, wages, and human capital, financial aid is offered to college students. Prior treatment estimates of myriad effects of merit-aid could be biased by selection into sample or known minimum eligibility cutoffs. This paper formally models, then tests viability and validity of an extension to measuring effects of merit-aid that improves upon assessing merit-aid's impacts. Under a grade progression assumption, students finish 1 grade per year on average, the Kindergarten cohort birthdate cutoff for beginning school can be used as discontinuous variation in college scholarship receipt. The effect of aid on credits and GPA are not as different as previously estimated. Different effects by gender are found for 3 year renewal for GPA, but not credits. Robustness checks are inconsistent with relative age bias or differences in sample. The rules for this design to work are widespread, making this design relevant for identification challenges in education policy.[1]

Keywords: Merit-Aid, Promise, Difference-in-Discontinuity

JEL Classification: H42, H75, I22

*Benjamin Blemings, Ph.D. Candidate, Department of Economics, West Virginia University, Morgantown WV 26506. E-mail: bbleming@mix.wvu.edu

[1]I thank my adviser, Adam Nowak, and my dissertation committee, Dan Grossman, Scott Cunningham, and Bryan McCannon. I thank Bill Even and John Winters for many helpful comments and suggestions. I thank seminar participants at Eastern Economic Association 2020, West Virginia University Brown Bag 2020, Southern Utah University, Kansas University, and the Computational Justice Lab at Claremont Gradutate University. I thank Liz Reynolds for graciously providing detailed administrative data.

1 Introduction

College graduates make $1 million more in lifetime earnings on average than non-college graduates (Carnevale et al., 2015), however college costs remain a primary obstacle to college (Cowan, 2016; Solis, 2017). To combat financial constraints, college financial aid has grown substantially.[2] West Virginia spends $40 million yearly (West Virginia Higher Education Policy Commission, 2014) on merit-based aid and Georgia had, from 1993-2011, 1,383,718 students receive merit aid.[3] Figure 1 shows merit grants exist in 27 states.[4] Abbott et al. (2019) suggests expanding ability-tested grants is superior to labor supply tax cuts or expanding student loans, but merit-aid is criticized for being ineffective (Ortagus et al., 2018) and a transfer to the wealthy (Dynarski, 2004; Cornwell and Mustard, 2007).

To inform debate about merit-aid and compare it to other interventions, it is therefore crucial to get accurate treatment effects for the many outcomes merit-aid affects (Scott-Clayton, 2011; Scott-Clayton and Zafar, 2019; Goodman, 2008; Sjoquist and Winters, 2015a, 2014; Cornwell and Mustard, 2007; Cowan and White, 2015; Sjoquist and Winters, 2015b). A seemingly attractive research design to evaluate the effects of merit-aid is regression discontinuity (RD), because there are defined academic cutoffs which discontinuously change the probability of treatment. Using entry test (ACT, SAT) scores and/or high school GPA cutoffs follows the original RD study (Thistlethwaite and Campbell, 1960; Cook, 2008).

However, an important difference between current merit-aid research and Thistlethwaite and Campbell (1960) is many academic cutoffs for merit-aid scholarships are known. Known academic cutoffs present a challenge to the RD design, because individuals can manipulate

[2]College grant aid per full-time enrolled student has doubled from $4,360 in 1998 to $9,520 in 2018 (College Board, 2019).

[3]Geogria data accessed at: https://web.archive.org/web/20111006161755/https://www.gsfc.org/GSFCNEW/SandG_facts.CFM?guid=&returnurl=http://www.gacollege411.org/Financial_Aid_Planning/HOPE_Program/Georgia_s_HOPE_Scholarship_Program_Overview.aspx. This citation required the use of waybackmachine.com.

[4]This is the classification given by (Cowan and White, 2015). 9 states offer a generous large program. To be considered large, it means a large amount of students are covered and/or a large portion of expenses are covered.

scores into receiving treatment and there is a large economic incentive to do so.[5] Indeed, students strategically manipulate academic scores into being just high enough to receive treatment, by retaking entry exams such as the ACT, in Tennessee, Florida, and West Virginia (Bruce and Carruthers, 2014; Zhang et al., 2016; Scott-Clayton and Zafar, 2019).[6,7]

This paper estimates effects of merit-aid on academic performance, in a state with known cutoffs and a large program, leveraging variation in college scholarship receipt due to birthdates' distance to Kindergarten entry cutoff date. This date is set by statewide compulsory schooling laws in place 13 years before the scholarship began. This strategy is less vulnerable to running variable sorting, because students' birthdate is realized 16 years before the scholarship announcement and 19 years before it begins. The strategy also more precisely studies potential cohort differences by using the high-frequency, continuous nature of birthdate, opposed to the discrete yearly measure of entry cohort year.

Prior literature studying college merit-aid leverages cohort-based variation (e.g. Scott-Clayton (2011) and Sjoquist and Winters (2015a)). However, small differences in birthdate and compulsory Kindergarten entry date cutoffs, which were never intended to interact in any way with college merit-aid may also cause variation in college cohort and therefore scholarship eligibility.[8] Fusing compulsory entry laws as variation in education (Black et al., 2011; Clark and Royer, 2013) and a birthdate-based discontinuity in financial aid (Denning, 2019), my design uses early education "programs" as variation in college grant receipt.[9]

[5]Two academic criteria are entry test scores and high school GPA. Entry tests can be retaken and high school GPA can be inflated through easier courses being taken or idiosyncratic high school decisions.

[6]Another issue with using is state-wide administrative datasets in Scott-Clayton (2011); Zhang et al. (2016) do not record if entry test scores are from the first attempt, the best attempt, or last attempt.

[7]Evidence of running variable manipulation is not consistent with the identifying assumption of local smoothness (Hahn et al., 2001; McCrary, 2008).

[8]Scott-Clayton (2011) uses a out-of state students, who are not eligible as a control group, however the control group changes over time, because the largest university began offering larger out of state scholarships when the scholarship was introduced. Another concern in Scott-Clayton (2011) with using cohort-variation is that students may select into the sample due to the scholarship, so Lee (2009) bounds are used to bound the effect. Instead of bounding the effect, this strategy of using continuous temporal based variation allows for a difference in means test to characterize selection.

[9]A direction for future research in financial aid noted by Dynarski and Scott-Clayton (2013) is analyzing how financial aid interacts with other "programs".

This design addresses two pervasive challenges, selection into treatment and sample, in the voluminous, active quasi-experimental literature on merit-aid programs.

First, internal validity is investigated. The Kindergarten cohort birthdate cutoff creates adequately strong variation in college cohort and college scholarship receipt. Academically eligible students born slightly after the Kindergarten cohort birthdate cutoff are 50 percentage points more likely to enter in the first eligible college cohort and receive the scholarship than academically eligible students born slightly before the cutoff. Additionally, the treatment and control groups are reasonably similar, because there is no evidence birthdate is manipulated (McCrary, 2008) and pre-treatment covariates are not different. Reasonable similarity and relevant first stage are consistent with causal identification assumptions for the fuzzy RD method used.

The alternative strategy reveals new insights into merit-aid scholarships. There is no difference in the likelihood of reaching pace-based (credit hours) or distinction-based (GPA) performance goals. Points estimates indicate students who receive the scholarship are between 0.24 and 0.3 pp more likely to hit both the credits and GPA renewal thresholds in their first year of college. The most closely related prior work finds the merit-aid scholarship increases the likelihood of hitting pace-based performance was 0.24 and distinction-based was 0.08 (Scott-Clayton, 2011). Students must finish both requirements to renew the scholarship, so the large discrepancy warrants additional attention, especially given inferiority of performance-based goals compared to task-based goals for college students (Clark et al., 2016).[10]

Additionally, there are heterogeneous effects of the program by gender for likelihood of finishing the GPA requirement each year. While male students finish the GPA requirement for each of 3 years when receiving the scholarship, female students do not. The heterogeneous effects are important, especially considering the differences in major choice by gender (Bordón et al., 2020) and underrepresentation of women in STEM (Mourifie et al., 2020).

[10] An important difference is for this merit-aid program, students receive a goal and money.

A new concern with this alternative strategy is a confounding effect of relative age (Matta et al., 2016). However, before the scholarship, relative age changes at the cutoff and scholarship eligibility does not. Thus, prior cutoffs are used to estimate and "control" for confounding effects of relative age, using difference-in discontinuities (Grembi et al., 2016; Smith, 2016; Canaan, 2020), providing support results are not caused by conflating relative age effects with the scholarship.[11] The estimator is not vulnerable to concerns that students born on either side of the cutoff have mothers with different characteristics (Buckles and Hungerman, 2013) for the same reason.

Beyond policy relevance of merit-aid, this research fits the broad direction of applied microeconomics. The alternative source of variation is consistent with estimating treatment effects using natural experiments (Angrist and Pischke, 2010).[12] In contrast to methodological approaches to addressing heterogeneity in treatment effects (Hsu and Shen, 2019), monotonicity (Bertanha and Imbens, 2019), or extrapolating (Angrist and Rokkanen, 2015; Dong and Lewbel, 2015) away from cutoffs in regression discontinuity designs, this paper estimates a "more average" effect using a different natural experiment. By using a birthdate running variable, the estimated effects are not local to an academic cutoff; estimates are local near a date, making generalizations away from the cutoff more likely to be appropriate.

Exact birthdate is continuous (while ACT is more discrete) which is consistent with identifying assumptions typically used (Hahn et al., 2001). The continuous running variable is less vulnerable to error due to specification bias caused by discrete variables not being arbitrarily close to the cutoff (Lee and Card, 2008) or choice of standard errors (Kolesár and Rothe, 2018). The estimates may be more heterogeneous, because potential students have less exact control of birthdate compared to entry test (ACT/SAT) scores. Higher running variable variance (as is arguably the case with a birthdate running variable compared to an

[11] The intuition of this strategy is outlined in Figure 8.

[12] The new source of variation provides a new way to form a counterfactual to solve the fundamental problem of causal inference (Holland, 1986).

entry test score) implies a greater mixture of students affected by the birthdate cutoff (Lee, 2008; Bloom, 2012) meaning more groups represented by estimates.

The following section discusses program details, prior research, and the relevant compulsory Kindergarten entry laws. Then, the programs' rules and resulting variation are formally modeled, then validated as being viable and internally valid. Then, method and results are presented, followed by a brief discussion of implications for college interventions for aiding student success and research on goal-setting.

2 Literature and Background

First, details about Promise are outlined. Then prior research is discussed. Finally, for understanding the design, historical compulsory entry laws in West Virginia are discussed.

2.1 Promise Program Details

The Promise scholarship began in 2002, after the original merit-aid scholarship began in Georgia in 1991. Forbes.com ranks West Virginia as second to last in education in the nation (McCann, 2020), which the scholarship was meant to address. The scholarship aimed to influence students to attend college in state and thereby hopefully remaining in state after graduating.[13] The program was funded by lottery sales. The state announced the program in 1999 and after accumulating funding, began awarding scholarships in 2002. At first, the scholarship covered the entirety of tuition and fees at public schools. However in 2010, the scholarship was capped at an annual payout of $4,750.[14]

[13] Merit-aid lead students to remain in state in West Virginia after graduating, but not other states Sjoquist and Winters (2014).

[14] The growth in the cost of college nationwide and at West Virginia University is shown in Figure A.1.

2.1.1 Requirements To Renew

Additionally, the program has had renewal requirements since its inception. First, students must finish at least 30 credits in each year to renew their scholarship.[15] Second, students must exceed a minimum GPA. In the first year they must obtain at least a 2.75 GPA. For every year afterwards they must obtain at least a 3.0 GPA. The scholarship is renewable for 8 semesters, 4 academic years.

2.2 Research on Merit-Aid

2.2.1 Renewal Requirements

An important design aspect of aid programs are scholarship renewal academic requirements. Renewal requirements set college academic performance thresholds students must exeed in each year of college to renew their scholarship in the next year. Programs with renewal requirements have been shown to more expediently graduate students (Cornwell et al., 2005; Jia, 2019). Scott-Clayton (2011) finds no differences in major choice due to scholarship receipt in an RD setup using academic eligibility measures in a similar sample to this paper.

2.2.2 Outcomes

Merit-aid affects many individual outcomes, including but not limited to: academic performance (Scott-Clayton, 2011), college choice (Goodman, 2008), major choice (Sjoquist and Winters, 2015a), student loan debt (Scott-Clayton and Zafar, 2019), post-college location choice (Sjoquist and Winters, 2014; Fitzpatrick and Jones, 2016), car purchases (Cornwell and Mustard, 2007), regional growth in enrollment (Sjoquist and Winters, 2015b), and alcohol consumption (Cowan and White, 2015).[16] Focusing on first year academic performance is warranted, because many students do not renew their scholarship for the entire period of eligibility (Carruthers and Özek, 2016). In terms of academic performance for a sample of

[15] They have the Fall, Spring, and Summer semester to do so.

[16] There are also studies of regional effects such as college attainment (Sjoquist and Winters, 2015b).

all West Virginia higher education students, the West Virginia Promise Scholarship has been shown to increase freshman credits by 1.8 (Scott-Clayton, 2011). The likelihood, freshman year, of exceeding the number of credits needed to renew the scholarship is 0.24 and the likelihood of exceeding the GPA required to renew is 0.08 (Scott-Clayton, 2011).

2.2.3 Research Designs

Research designs estimating causal effects of merit-aid typically use two approaches. One approach, with individual data and academic eligibility measures, is an RD design leveraging small differences in academic scores. Many merit-aid programs suffer from manipulation due to the academic cutoffs being known.[17] Another issue with merit-aid RD estimates leveraging academic cutoffs is the possibility some schools provide the best score and others the first attempt and which attempt is used matters; Bruce and Carruthers (2014) show no manipulation in the first score, but manipulation in the best score.[18] The other design is cohort-level variation, which can naturally be used for differences-in-differences methodology as long as pre and post policy outcomes are available. This approach assumes students in the post-policy cohorts are similar to pre-policy cohorts. The cohort variation and academic eligibility have been combined by examining academically eligible students before and after using 2SLS, where the instrument for scholarship receipt is a dummy variable for entering college in a cohort after the policy began.[19] The drawback of cohort-level approaches is differential selection, because students have an incentive to sort into the sample.[20]

[17] For large merit-aid programs these cutoffs are known (Scott-Clayton, 2011; Scott-Clayton and Zafar, 2019; Bruce and Carruthers, 2014; Zhang et al., 2016). For merit-aid in non-large states and/or targeted to needy students, cutoffs are typically unknown,(e.g. Page et al. (2019); Bettinger et al. (2019).

[18] Students also strategically retake entry tests for reasons other than desiring to receive merit-based aid (Vigdor and Clotfelter, 2003). Re-taking also leads to higher scores as students learn by doing (Frisancho et al., 2016).

[19] In survey datasets such as American Community Survey (ACS) and the College Alcohol Survey (CAS).

[20] Scott-Clayton (2011) attempts to bound selection effects “in the spirit of” Lee (2009) bounds.

2.3 Compulsory Kindergarten Entry Laws

The compulsory entry laws relevant to this paper are birth date cutoffs for when children must begin Kindergarten. These laws vary across states and times. Since Promise begins in 2002 and most students enter college after completing grades K-12, the relevant entry laws were in effect during the 1980's. In West Virginia, before 1982, students who were 5 years old by November 1 were supposed to enter Kindergarten in that year (Whaley, 1985, pg. 20). In 1982, the rule changed to students who were 5 by September 1 of the current school year were supposed to enter Kindergarten.

The Kindergarten date cutoff determining year of college cohort entry depends on how strictly the cutoff is enforced (i.e. whether students born later are allowed to enter early).[21] (Whaley, 1985, pg. 20) states, "local districts may permit earlier entry if a child shows readiness based on a test of basic skills." Due to the potential for students to enter early it is highly improbable that the cutoff changes the probability of entering in the first eligible cohort from 0 to 1, thus dictating a fuzzy RD research design. However, (Whaley, 1985, pg. 20) also states, "Complaints occur about the fact one district may allow early entry, while a neighboring one may not," suggesting that the laws are enforced to a certain extent. West Virginia compulsory age of Kindergarten entry does not apply to students that enter Kindergarten in another state, then move into West Virginia before which works against there being a strong first stage.[22] Additionally, in 1988 the law was changed again which means that students in year after are sorted at a different date cutoff, November 1.

In the years prior to the scholarship becoming available, the cutoffs affected relative age, but not scholarship receipt. These prior cutoffs will be used to ensure the estimated effects are not being driven by relative age effects using difference-in-discontinuity (Smith, 2016;

[21] If the cutoff is not enforced strictly enough, then the cutoff will not have explanatory power for cohort of entry or scholarship receipt.

[22] (Lleras-Muney, 2002, pg. 407) notes a similar point in footnote 36, going further to point out that inter-state mobility was probably not related to compulsory entry laws, which is likely also the case for this paper. Further, Lleras-Muney (2002) notes that mobility is low in the US in the period of study (Card and Krueger, 1992). Mobility into West Virginia is likely also low. Finally, it is highly unlikely that someone moved into West Virginia early in grade school to receive the scholarship.

Grembi et al., 2016). Under certain assumptions, difference-in-discontinuity estimates are not vulnerable to effects of the scholarship being conflated with relative age.

2.3.1 Related Prior Research Designs

Some literature uses closely related designs. First, a birthdate running variable has already been used to study financial aid (Denning, 2019). The difference is that Denning (2019) uses a different birthdate cutoff, one exploiting tax rules, which awards a smaller amount of financial aid than the scholarship studied here. One study that uses a birthdate cutoff is explicitly interested in relative age within cohort effects in college (Matta et al., 2016). Finally, Canaan (2020) uses a kindergarten entry date cutoff to study the effect of early school tracking, deciding a course of study, in France on long term wages. To deal with the potential relative age effects that are identified by Matta et al. (2016), Canaan (2020) uses a difference in discontinuity to difference out relative age effects. This paper is different than Canaan (2020), because it is interested in leveraging a new source of variation in college cohort, not primary school cohort and studying a college "intervention".[23]

3 Data

The data is student-level of three cohorts at a large, public university in West Virginia.[24] The sample consists of only in-state students who are first-time freshman. Exact birthdate, gender, race, high school GPA, entry test scores, and academic outcomes up to the third year of college are observed. Table A.1 presents summary statistics of the sample.

3.1 Scholarship Eligibility

The data has information on 3 requirements required to be eligible for the scholarship.

[23] Additionally, Canaan (2020) does not observe who is actually tracked, making the estimates intent to treat, while this paper observes treatment status.

[24] After multiple rounds of data requests over multiple years, the West Virginia Higher Education Policy Commission (WVHEPC) ultimately denied the data request for statewide data.

1. Earn at least a 3.0 GPA in high school
2. ACT Composite score > 20.5 OR SAT Combined Score > 1,000
3. One must be a college freshman in the 2002 cohort or after

As the program is merit-based, 1 and 2 are requirements related to academic proficiency. The third requirement is one must enter college after the program begins. To simplify, requirements 1 and 2 are combined into a single variable, academic eligibility, defined as

$$\text{Acad. Eligible} = 1[(\text{HS GPA} \geq 3) \cap (\text{ACT Comp} > 20.5 \cup \text{SAT Comb} \geq 1{,}000)].$$[25]

3.2 Temporal, Cohort-Based Eligibility

The third requirement will be referred to as temporal eligibility. Prior studies of the same program (Scott-Clayton, 2011; Scott-Clayton and Zafar, 2019) define temporal eligibility of individual i as

$$Temporally\ Eligible_i = 1(Cohort \geq Cohort_{Initial=2002}), \tag{1}$$

in which $Cohort$ is the observed cohort of entry and $Cohort_{Initial}$ is the first cohort that a student can receive funding, 2002 in the case of the Promise scholarship. The crucial insight of this paper is that small differences in birthdate, combined with Kindergarten entry date cutoffs, leads to variation in the likelihood of entering in the initial cohort.

4 Kindergarten Cohort and College Aid Receipt

Fulfilling the temporal eligibility requirement is related to ones birthdate.[26] Combining birthdate with a compulsory Kindergarten entry date cutoff provides a potential source

[25] HS GPA stands for high school GPA, ACT Comp stands for ACT composite score, and SAT Comb stands for SAT combined score.

[26] Students born many years before the scholarship begins, for example 30 years or more, are unlikely to begin college in the first year or after the scholarship begins. The reason for this is typically students enter college around the age of 18, after completing the pre-requisite grades of Kindergarten through 12$^{\text{th}}$ grade.

of variation with which to estimate the effects of the Promise scholarship program. The alternative source of variation hypothesized can be succinctly written

$$P(Cohort_{Init=2002}) = \begin{cases} p + c & if\ BD \geq \overline{BD}_{y=1} \\ p & if\ BD < \overline{BD}_{y=1} \end{cases} \tag{2}$$

in which BD is the observed birthdate.[27] $\overline{BD}_{y=1}$ is the date 14 years before the scholarship began in which children aged 5 born before $\overline{BD}$ were required to enter Kindergarten and children born after were more likely to enter in the next cohort. Entering Kindergarten 13 years before the college scholarship begins is fortuitous, because if one finishes 1 grade per year and then enters college, the individual is temporally eligible to receive the scholarship.[28]

For the design to be valid, there needs to be a high level of statistical confidence that $p < p + c$ (i.e. you are discontinuously more likely to enter college later if born after the Kindergarten date.)[29,30] Cancelling the p's, the design requires c>0, because then the Kindergarten date cutoff sorts students into different college classes in a predictable way, students finishing 1 grade per year.[31] To demonstrate how $c > 0$ could realistically arise, Figure 2 shows a timeline comparing children with slightly different birth dates. All terms in Equation 2 are observed, except for c, so c can be non-parametrically estimated.

The constant c can be thought of as the predictive power of the cutoff for Kindergarten cohort combined with the finish one grade per year hypothesis. In the special case where p = 0 and p+c = 1, the running variable would perfectly sort students into treatment and sharp regression discontinuity would be appropriate. However there are reasons to believe that $(p + c) - p < 1$, because students born before the cutoff date might enter in a later cohort and those born after the cutoff might enter in an earlier cohort. First, students repeating or

[27] I abstract away from other cutoffs in this equation. For now only consider birthdates reasonably local to the cutoff sorting students into starting college in 2002, like within 200 days on both sides of the cutoff.

[28] "Denning (2019), also using a birthdate-based discontuity to study effects of college financial aid, encodes the fortuity of barely having an eligible birthdate by being titled, "Born Under a Lucky Star..."

[29] As p is $\in$ (0,1), $0 \leq c < 1 - p$.

[30] Formally, $[\lim_{BD \to 0^-} P(Cohort_{2002}) = p] \neq [\lim_{BD \to 0^+} P(Cohort_{2002}) = p + c]$

[31] If $c = 0$, then the Kindergarten cutoff does not sort students into college cohorts.

skipping grades will push c towards 0. Next, locally varying imperfect enforcement means some students may have entered earlier or later than legislated. Finally, in-migration from other states with different cutoff dates during grade school could lead to students entering in a different cohort than predicted by the cutoff and finish one grade per year hypothesis.

Students who enter in the initial cohort are temporally eligible for the scholarship by definition, making them more likely to receive the scholarship

$$P(Scholarship = 1) = \begin{cases} q + k & if\ Cohort_{Init} = 1 \\ q & if\ Cohort_{Init} = 0. \end{cases} \tag{3}$$

k can be considered how important being in the appropriate cohort is for scholarship receipt, like strictness of the requirement to enter in an eligible cohort. Equation 3 is Equation 1. For kindergarten cohort cutoff to matter for scholarship receipt, the cohort requirement must be an important enough factor, requiring k>0 and q+k>q. With a sufficiently large c, Equation 3 can be restated as

$$P(Scholarship = 1) = \begin{cases} q + k' & if\ BD \geq \overline{BD}_{Y=1} \\ q & if\ BD < \overline{BD}_{Y=1} \end{cases} \tag{4}$$

Since birthdate may not perfectly sort students into cohorts, k is replaced with k'.

Equations 5, 6, and 7 outline the system of equations which leads to the variation in scholarship receipt exploited in this paper. Equation 5 is related to k' in Equation 4. Equations 6 and 7 are related to c in Equation 2.

$$Scholarship = f(College\ Cohort) \tag{5}$$

$$College\ Cohort = g(Kindergarten\ Cohort) \tag{6}$$

$$Kindergarten\ Cohort = h(Birthdate\ Relative\ to\ Cutoff) \tag{7}$$

Equations 5 and 7 are (discontinuous) rules about scholarship eligibility and Kindergarten cohort entry respectively. To systematically predict college cohort based on Kindergarten cohort requires students finish 1 grade per year, on average. Mathematically, linking Kindergarten cohort to college cohort requires a functional form assumption about g in Equation 6. The assumed functional form of g is

$$College\ Cohort = Kindergarten\ Cohort + 13.$$ [32]

That is students finish kindergarten through 12^{th} grade and then enter college. This is a reasonable assumption for the purposes of estimating the effects of merit-based aid, because it is unlikely that students who receive merit-based aid are being held back.

By substitution,

$$College\ Scholarship = f(g(h(Birthdate\ Relative\ to\ Kindergarten\ Cutoff))). \quad (8)$$

Equation 8 shows the potential source of variation, requiring c, $k' > 0$, for estimating treatment effects of merit-aid. The next section infers c and k' from the data. Figure 3 plots how scholarship receipt depends on birthdate and academic eligibility, suggesting $c > 0$.

5 Internal Validity of Research Design

The objective of this section is establishing internal validity. First, the first stage is tested, specifically examining whether small differences in birthdate lead to different cohort year of entering college (c in Equation 2) and differential likelihood of scholarship receipt (k' in Equation 4.) Second, implications of the identification assumption are examined in two ways. A McCrary density test is used to investigate distributional smoothness of number

[32] There are 13 years between entering Kindergarten and entering college. A more abstract way to think about it is # of years until a particular policy begins.

of observations at the date cutoff (McCrary, 2008). Second, in the vein of a randomized controlled trial (RCT), pre-treatment covariate balance at the birth date cutoff is examined.

5.1 First Stage

For the approach to be valid, there must be a large enough jump in the probability of treatment at the entry date cutoff. A predictable, systematic jump in probability of treatment at the cutoff requires $c > 0$ and $k' > 0$, Figure 4 tests these hypotheses in the data. Figure 4a examines whether the Kindergarten entry date cutoff has a discontinuous effect on the year students enter college at the cutoff in the year prior to the date cutoff also determining scholarship receipt. At the cutoff there is an approximately 50 percentage point (pp) increase in the likelihood of entering college in the year that is consistent with finishing one grade per year (i.e. 13 years later after finishing grades K-12). Figure 4b shows that this cutoff does not have an effect on scholarship receipt, because the scholarship does not begin until the next college cohort enters.

Figure 4c shows the year the cutoff determines scholarship receipt has a nearly identically sized discontinuity to the year before. The $\lim_{BD \to 0^-} \approx 0.2$ and $\lim_{BD \to 0^+} \approx 0.8$, implying $c_{y=1} \approx 0.6$. Since $0.6 \not\approx 0$, the first condition is satisfied. Additionally, $c_{y=-1} \approx c_{y=1}$ which reassuringly suggests the scholarship did not alter the likelihood that students are sorted into college cohorts by the Kindergarten cutoff. As shown by the non constant shape of the points, the probability of entering in the next cohort increases as distance from the cutoff increases. This is to be expected, because it implies that students born farther from the cutoff are less likely to enter school early.[33]

Figure 4d tests whether cohort of entry leads to large differences in the likelihood of treatment by the merit-aid scholarship. The $\lim_{BD \to 0^-} \approx 0.1$ and $\lim_{BD \to 0^+} \approx 0.5$, implying $k'_{y=1} \approx 0.4$. While $0.4 \neq 0$, a larger discontinuity would be more ideal, because it means

[33] Alternatively, on the left side of the cutoff, as distance from the cutoff increases the likelihood of entering in the next cohort also decreases.

the proposed variation more strongly predicts receiving the scholarship. One reason k' is relatively low is because the academic eligibility requirements are not incorporated.[34]

Figure 4e and Figure 4f repeats the visualizations in Figures 4c and 4d after subsampling to only academically eligible students. In Figure 4e, the discontinuity in cohort entry year is similarly sized to Figure 4c. In contrast, the discontinuity in Figure 4f, k', is about 25 percentage points larger at the cutoff compared to Figure 4d and gets close to 1 as the distance from the cutoff increases. Overall, Figure 4 shows Kindergarten entry date cutoffs are adequately relevant for the first stage of a fuzzy regression discontinuity design.

5.2 McCrary Density for ACT Compared to Birthdate

Earlier work in West Virginia has shown that using academic eligibility cutoff scores fail the McCrary density test (Scott-Clayton, 2011), there is a larger mass of observations to right of the cutoff which suggests local smoothness does not hold.[35] The most similar papers, Scott-Clayton (2011) and Scott-Clayton and Zafar (2019) have data for all of West Virginia, of which only subsample is contained in this paper. Figures 5a-5d focus on showing that the evidence of entry test score manipulation is also present in this subsample.

Figure 5a shows that there is a no discontinuity in the distribution of ACT scores at the ACT cutoff in the years before the scholarship is available. Figure 5b shows that after the scholarship becomes available, there is a discontinuity in the distribution of ACT scores at the required score to receive the ACT, suggesting that the local smoothness assumption is violated. Figures 5c and 5d show the corresponding density tests for running variable manipulation (McCrary, 2008), using the local polynomial density estimator of Cattaneo et al. (2019b), are consistent with the histograms. There is evidence the academic eligibility running variables are manipulated in this subsample, calling into question the local smoothness identification assumption (Hahn et al., 2001; McCrary, 2008) if academic scores were to

[34] Figure A.4 repeats the analyses in Figures 4c and 4d, using optimal binning procedures (Calonico et al., 2015), and the results are not substantively different.

[35] This sorting on the academic running variable is also observed in Tennessee and Florida (Bruce and Carruthers, 2014; Zhang et al., 2016).

be used as a running variable. This suggests studies using academic eligibility as a running variable use designs are not identified.

Figure 5e tests whether there is a difference in the birthdate distribution at the entry date cutoff before it is relevant to determining scholarship eligibility. One might expect a discontinuity in the distribution, despite it not mattering for scholarship eligibility due to the effect of relative age within a cohort. For example, students who are older compared to other students in their cohort may be more likely to go to college, which would potentially generate a discontinuity. It appears that there is a difference in the McCrary Density Test presented in Figure 5e, but the difference is not statistically significant at 95 percent confidence.

Figure 5f presents the McCrary density test for manipulation at the cutoff which determines cohort and scholarship eligibility. It fails to reject that there is a discontinuity in the distribution of birthdates at 95 percent confidence. Comparing Figure 5f to Figure 5e suggests there is not a major concern that the birthdate entry cutoff matters more in the first year of the scholarship. Comparing Figure 5d to Figure 5f suggests that using a birthdate running variable is less vulnerable to concerns of running variable manipulation.[36]

5.3 Pre-Treatment Covariate Balance

It is possible to examine how similar the treated group is to the non-treated group by examining whether the two groups have similar averages for pre-treatment covariates at the cutoff used for treatment assignment. Practically this is performed by using pre-treatment covariates as the outcome and looking/testing for discontinuities at the cutoff. Figure 8a shows 1 pre-treatment covariate that is, by definition, discontinuous at the cutoff.

That covariate is within-cohort age rank, because students immediately to the right of the cutoff are the oldest in their cohort. However, the cutoff exists in prior years where the cutoff

[36] The graphical results of two additional McCrary density tests are reported in Figures A.5a and A.5b. The difference in these tests is two additional restrictions, an equal cdf on both sides and the existence of higher order derivatives, which increase the power of the test if the assumptions are satisfied. Increased power is desirable, because 1 school sample results in potentially too few observations at the date cutoff. The results are still unable to reject the null of a continuous distribution on either side of the cutoff with these additional restrictions. The optimal bandwidth is wider than available data in Figure A.5b.

does not also determine scholarship receipt, so this relative age effect can be differenced out as described in Section 7.4.[37]

5.3.1 Plots

Figure 6 examines the continuity of academic eligibility variables at the cutoff and also displays the prior year cutoff so the relative age effect can be compared to relative age and scholarship effect at the cutoff. In Figure 6a, there is a slight discontinuity at the cutoff entry date which is likely indicative of a relative age effect. The downward slope as distance from the cutoff on the right also suggests this is the case. Figure 6b shows a similarly sized discontinuity at the cutoff to Figure 6a. One difference is that the slight downward slope is not there, suggesting that the effect of moving away from the cutoff increasing likelihood of entering in the initial cohort is working against the relative age effect.

Figure 6c shows the difference in ACT composite score at the date cutoff in the year prior to the scholarship beginning. Somewhat unexpectedly, individuals to the left have slightly higher cutoffs which is not consistent with the relative age effect. However, this difference is small is unlikely statistically significant.

In Figure 6d, contrasting Figure 6c, individuals immediately to the right have a slightly higher ACT on average than those immediately to the left. This slight movement is perhaps not surprising given the incentive to earn an ACT score which is eligible if one is in the eligible cohort. There is a slight upward slope, just like Figure 6b, which could reflect the increased likelihood of entering in an eligible class farther to the right of the cutoff. However, this slope may not be much different than individuals to the left of the cutoff in Figure 6c. In Figure 6d, the averages on either side of the cutoff are close enough to be considered consistent with the identifying assumption, local smoothness.

Figures 6e and 6f show SAT scores, and since fewer students take the SAT compared to the ACT in West Virginia, the conditional means are more dispersed than those in Figures

[37] Qureshi and Gangopadhyaya (2019)- medicaid human capital and Cohodes et al. (2016)- eligibility for health insurance as a kid.

6c and 6d. Figure 6e shows the combined SAT scores in the year before the scholarship begins. There is a discontinuity at the cutoff, students to the right have higher combined SAT scores. Figure 6f has a similar sized discontinuity, so its plausible that local smoothness is adequate.

6 Methods

Fuzzy regression discontinuity is used to estimate the effect of scholarship receipt on academic performance.[38] It exploits psuedo-random assignment of the scholarship, which is a result of state-wide Kindergarten compulsory entry age assigning students to different cohorts based on small differences in birth-date. The two-stage approach follows

$$S_i = \lambda + \psi(After_c) + \gamma(BDdist_i * Before_c) + \xi(BDdist_i * After_c) + \epsilon_i, \tag{9}$$

in which subscript i stands for individual and c stands for cohort. S stands for scholarship receipt. The fitted values for scholarship receipt from Equation 9 are then inserted into Equation 10,

$$y_i = \alpha + \beta(\hat{S}_i) + \zeta(BDdist_i * Before_c) + \pi(BDdist * After_c) + e_i.^{39} \tag{10}$$

The predicted values of scholarship for individual i are represented by $\hat{S}_i$. BDist stands for number of days away from being more likely to be born into the first Kindergarten cohort that was likely to receive funding. The estimated effect of the scholarship is β, the observed difference between the regressions on either side of the cutoff, at the cutoff.[40]

[38]Difference in Discontinuities is used as a robustness check.

[39]Conventional RD estimates using MSE-optimal bandwidths will not be small enough, in general, to eliminate bias from undersmoothing (Calonico et al., 2017), leading to poorly performing confidence intervals. To address this issue, a bias-correction is applied according to Calonico et al. (2017). However this bias-correction estimation leads to increased variability that also must be taken into account, so the RD estimates also adjust the confidence intervals to account for this appropriately (Calonico et al., 2017).

[40]The interaction terms enable the fitting of separate regressions on either side of the cutoff.

6.1 Outcomes: Academic Performance Based Renewal

The dependent variables, y, are binary indicators for whether a student finished the minimum required credits ($y_{Credits>Renewal}$) or GPA ($y_{GPA>Renewal}$), respectively. While continuous outcomes are interesting, binary outcomes for exceeding renewal requirements have been used before (Scott-Clayton, 2011), performance based funding is criticized (Ortagus et al., 2018), and performance based goals have been shown to be less effective than task based goals (Clark et al., 2016). Prior work finds large discrepancies between $\tau_{Credits}$ (=0.24pp) and τ_{GPA} (=0.08pp), this paper applies an improved design to investigate whether this discrepancy is sustained or if the new parameter estimates are closer to 0.24 or 0.08.

6.1.1 Mechanism: Goals

Clark et al. (2016) investigates the relative efficacy of performance-based and task-based goals. Models about goals (Koch and Nafziger, 2011) hinge on both present-bias (Strotz, 1955) and reference-dependence (Kahneman and Tversky, 1979). A present-biased student might set out to exert a preferred effort level, but when the time comes to attend class or review for a test lack the self-control. A loss-averse student with a salient reference point might work hard to achieve the reference point. Clark et al. (2016) finds performance-based goals have small, positive, but insignificant effects, while task-based goals are effective. However, the academic renewal requirements which are academic performance-based goals have been shown to be effective (Scott-Clayton, 2011).[41]

6.1.2 Mechanism: Money

One of the reasons the conclusions about performance-based goals may differ between Scott-Clayton (2011) and Clark et al. (2016) is merit-aid comes with performance goals and a financial subsidy, while Clark et al. (2016) studies performance based goals with no subsidy.

[41] Additionally, program requirements for renewing aid can alter the incentives for pursuing more difficult majors (Sjoquist and Winters, 2015a).

If students and parents share the financial windfall from aid, then students will have more disposable income. Increases in disposable income lead to increased consumption of normal goods. Second, by increasing income, merit aid could affect the allocation of time. Students may choose to work fewer hours in formal employment, study more, or engage in other extracurricular activities.

6.2 Bandwidth

Observations far away from the cutoff bias the estimated effect at the cutoff by affecting the fit on either side of the cutoff. However, dropping them leads to less observations and increases the variance. This is the well-established bias-variance tradeoff.[42] To leverage statistical methods for calculating optimal bandwidths, local linear regression is used. The approach used is a mean squared error (MSE)-optimal bandwidth as shown in Calonico et al. (2014). The MSE-optimal bandwidths are desirable for good point estimation, but coverage error rates (CER) are desirable for optimal inference (Cattaneo et al., 2019a, pg. 8).[43] In light of this fact, some specifications make use of CER-optimal bandwidths.

6.3 Kernel

Another required choice when using local linear regression is which kernel, or weighting scheme to use. A uniform kernel is unweighted, giving every observation equal weight, regardless of its distance to local mean being estimated. [44] A different approach would be to use a triangular kernel. A triangular kernel puts more weight on observations close to the cutoff. A desirable feature of the triangular kernel is that when it is used with MSE-optimal

[42] One could use a global estimator and ad-hoc drop observations, but this approach is unlikely to reach an optimal bias-variance tradeoff.

[43] Despite the fact that CER-optimal bandwidths are typically smaller than MSE-optimal bandwidths.

[44] The use of a uniform kernel is supported in (Lee and Lemieux, 2010, pg. 319), noting that a uniform kernel and estimating a model with different bandwidths leads to easy to interpret results. A desirable property of the uniform kernel is that it minimizes the asmyptotic variance under technical conditions (Cattaneo et al., 2019a, pg. 44). Indeed the simulations in Imbens and Kalyanaraman (2012) shows uniform kernels minimize error in several simulations compared to other methods.

bandwidths, it leads to point estimates with optimal properties (Cattaneo et al., 2019a, pg. 44). It is usually the case that kernel choice does not matter much (Lee and Lemieux, 2010) and (Calonico et al., 2017, pg. 376) notes that either are popular choices. Compared to most other RD designs, this data has a relatively limited number of observations, so it is possible that a uniform kernel will be required to make optimal inferences.

6.4 Identification

Next, assumptions for β to represent the causal effect of receiving the scholarship on academic performance are considered. First, regression discontinuity assumptions are discussed. One of the core assumptions is, by definition, violated, necessitating a difference-in-discontinuity approach to ensure scholarship effects are not conflated with relative age effects.

6.4.1 Usual Two-Stage Assumptions

Equations 9 and 10 outline a two-stage least squares approach where the treatment assignment rule is known and therefore must satisfy the usual instrumental variable assumptions.[45] For the design to be valid, there must be a discontinuous change in the likelihood of scholarship receipt at the date cutoff. If the birthdate cutoff does not generate a large difference between the probability on the left-hand side from the probability on the right-hand side, then ψ (from Equation 9) is too small to yield precise estimates of β (from Equation 10).

Two other usual assumptions of instrumental variables are monotonicity and the Stable Treatment Unit Value Assumption (SUTVA.) Monotonicity simply means that being to the right of the date cutoff does not reduce the likelihood of entering in the later cohort and subsequently receiving the scholarship for any individual student. This seems like a reasonable assumption.

Next, SUTVA assumes that the potential outcomes of students to the left of the cutoff are not affected by the treatment of students to the right of the cutoff. This assumption is

[45] Dong (2018) requires less assumptions.

quite reasonable for year 1 outcomes, but when students on either side of the cutoff are in school at the same time, one cohort being treated may cause students in the other cohort to also enroll in more credits to be on pace with peers. This is part of the reason to focus on the first year outcomes. Another reason to focus on 1st year outcomes is SUTVA assumes equal dosages of treatment, which is violated for long-term outcomes, because students lose their scholarship if they fall below the scholarship, as noted by Carruthers and Özek (2016).[46]

6.4.2 Local Smoothness

Although Thistlethwaite and Campbell (1960) motivates the original RD design as a locally randomized experiment, that is actually a stronger requirement than is required for identification of causal effects (Hahn et al., 2001; Cattaneo et al., 2016).[47] The identifying assumption is local independence, which is typically weakened to local smoothness (Hahn et al., 2001; Dong, 2018).[48,49] Researchers support local smoothness using the McCrary Density test (McCrary, 2008) and ensuring that pre-treatment covariates are balanced for treatment and control groups.

Figure 8a shows how this within-cohort age rank changes discontinuously at the cutoff. Dealing with this difference in within-cohort age rank requires a slight adjustment of the regression discontinuity design. Assuming away this difference in within-cohort age rank, under the above conditions, β could be considered the local average treatment effect of receiving Promise. Naturally, the treatment effect is for compliers and local to those with births around the cutoff. So, because the birthdate cutoff does not induce out-of-state students to take the treatment, the effect is not valid for them- they are never-takers. More

[46] Figures A.2 and A.3 show that many lose their scholarship, but there is still a discontinuity when considering the treatment being 4 years of scholarship.

[47] Local randomization is discussed in Lee (2008), discussed as the stronger assumption required for randomization inference in RD designs in Cattaneo et al. (2016) and used in Cattaneo et al. (2017) to re-evaluate the effects of Head Start program on child mortality.

[48] Figure A.7 shows the identification assumption. To be a causal effect, potential outcomes must be smooth at the cutoff.

[49] These outcomes are not observed, so this assumption is not directly testable. This is the essence of the "fundamental problem of causal inference," (Holland, 1986, pg. 947).

assumptions are required for the treatment effect to be relevant for units farther away from the cutoff (Bertanha and Imbens, 2019).

7 Results

First, reduced form plots show the relationship between the date cutoff and academic outcomes and the local linear regression estimates are presented. Next, robustness checks investigate whether differences in the results from prior work are due to confounding relative age effects or differences in sample. Afterwards, the analysis is extended to outcomes over 3 years and heterogeneity by gender is examined.

7.1 Reduced Form Plots, Year 1 Outcomes

Figure 7 presents plots of the reduced form relationship between birthdate cutoffs (the year before the scholarship begins and the year it begins) and the likelihood of meeting first year renewal requirements in GPA and credits. Reduced form plots are shown, because upcoming estimates are local cases of 2 stage Wald estimates from instrumental variable designs.[50] No visually detectable reduced form effect would call regression estimates into question.

Figure 7a shows the effect of exceeding the birthdate cutoff, and thus being the oldest in ones' class, on the likelihood of completing at least 30 credits. This birthdate cutoff does not determine scholarship receipt, because it is the year before the scholarship is available, so the difference at the cutoff represents the effect of relative age on finishing at least 30 credits. There is no visually detectable jump in likelihood of completing renewal credits due to being the oldest in ones' class. Figure 7b looks at the same cutoff in the next year, when it also determines scholarship receipt. Those barely exceeding the cutoff are discontinuously more likely to complete 30 credits, highly suggestive of a reduced form effect of birthdate

[50] $\tau_{LATE} = \frac{\beta_{Reduced\ Form}}{\beta_{First\ Stage}}$.

on the outcome when birthdate also predicts scholarship receipt. The outcomes follow the same curved pattern as the likelihood of receiving treatment shown in Figures 4.

Figure 7c shows the effect of exceeding the birthdate cutoff, before the scholarship, on likelihood of exceeding the first year renewal GPA (2.75). There is no discontinuous increase at the cutoff. However, when birthdate determines scholarship receipt in Figure 7d, there is an increase in the likelihood of finishing the GPA requirement at the cutoff.Figure A.8 repeats Figure 7 using only academically ineligible kids, because academically ineligible kids do not gain access to a scholarship by being born after the cutoff date. Figure A.8 shows no evidence of any change in outcomes at any discontinuity. No change in outcomes for the academically ineligible is consistent with access to a scholarship being the reason for the jump in outcomes in Figures 7b and 7d.

7.2 Regression Discontinuity Estimates, Year 1 Outcomes

Table 1 presents estimates of scholarship receipt on first year academic outcomes. There are varying choices of kernel, sample, covariates, coefficient adjustments, and bandwidth selection criterion to account for several justifiable ways of estimating the effects. Table 1, Panel A focues on the likelihood of completing 30 credits in ones' first year. In column 1, no bias-correction or robustness corrections to the estimates are done and the scholarship increases the likelihood of completing the scholarship by 36 percentage points, which is significant above 95 percent confidence. In column 2, the bias-correction and robustness adjustments are done (Calonico et al., 2017), and the results are essentially unchanged.

In columns 1 and 2 (in Panel A and B) academically ineligible students are included, making the control group less ideal. In column 3, academically ineligible students are dropped and the effect size decreases to 27 percentage points, probably due to a more academically gifted control group being utilized. Additionally, the statistical significance drops to 90%. In column 4, academic controls for HS GPA and entry test (ACT/SAT) scores are included to compare individuals of similar scores.[51] The effect size drops to 0.24, which is the exact point estimate Scott-Clayton (2011) reports, but the coefficient is no longer statistically significant at conventional levels (p<0.128). Column 5 uses CER-optimal bandwidths, which can be more desirable for inference, but less desirable for point estimates. The estimated effect does not change, but the standard error grows, reducing the p-value.

Columns 6 and 7 use a uniform kernel, because it minimizes asymptotic variance under certain conditions (Cattaneo et al., 2019a, pg. 44). The difference in scholarship receipt at the cutoff increases noticeably with the distance from the cutoff for a few months before flattening, so triangular kernels place more weight on potentially unaffected observations. Column 6 uses the MSE-optimal bandwidth criterion with a uniform kernel, the effect size remains similar, but the standard error drops and the effect is statistically significant above 90% (p<0.085). In column 7, a uniform kernel is used with the coverage error rate optimal bandwidths, the effect size grows to a 31 percentage point increase. The standard error also grows, causing this estimate to remain statistically significant at 94.6 percent confidence. The point estimates are similar to Scott-Clayton (2011), sometimes slightly larger depending on bandwidth, sample, controls, and correction.

Table 1 Panel B replicates Panel A, using the likelihood of meeting the renewal GPA of 2.75 as the dependent variable. Column 1 shows the scholarship increases the likelihood of hitting the GPA threshold for renewal by 31 percentage points. The effect is statistically significant above 95%. Columns 2-4 add parameter corrections, subsample to academically eligible students, and include academic controls; all changes leave the effect size and statistical significance almost unchanged. Column 5 uses the CER-optimal bandwidth and finds a 33% larger effect. This point estimate should be interpreted cautiously as the point estimate does not have optimal point estimate properties (Cattaneo et al., 2019a, pg. 44).

[51] Non-academically eligible students are dropped before academic controls are included, because covariate inclusion in RD assumes that the controls are additively separable from the running variable (Calonico et al., 2019). Dropping non-academically eligible students increases the plausibility of this assumption.

Columns 6 and 7 of Table 1, Panel B use the uniform kernel. In column 6, the coefficient is 33% smaller, but the standard error is reduced. Still, the coefficient is only statistically significant at 9% confidence. In column 7, a uniform kernel is used with the CER-optimal bandwidths and the estimate becomes significant at 95 percent confidence. and the effect grows back to near the previously estimated effects. These estimates are similar to the increased likelihood of completing the renewal credits that are presented in Panel A of Table 1, in contrast to Scott-Clayton (2011).

7.3 Addressing Whether Sample Affects Results

One possibility is that differences in the sample is responsible for the discrepancy between these estimates and Scott-Clayton (2011). To investigate, Table 2 uses a strategy closer to that of (Scott-Clayton, 2011), by using a non-local approach that makes global functional form assumptions. In Panel A and B, column 1, using the pre (2001) and post (2002) cohort and a linear model, the estimates suggest a 0.09 pp increase in likelihood of hitting the GPA renewal requirement in the first year. This global approach is as close to Scott-Clayton (2011) as possible and the coefficient estimates are nearly identical.

Moving to the right, observations far away from the cutoff are dropped and the effect size increases. Since observations away from the cutoff increase specification bias in the treatment effects, these estimates suffer less bias from observations away from the cutoff. Another difference worth emphasizing is that a quadratic model typically finds larger results (Panel C compared to A and Panel D compared to B). This suggests that imposing a linear assumption may lead to under-estimated effects. My approach does not need to choose between functional forms by using local linear regression, allowing the functional form to fit the data. Using a similar approach to Scott-Clayton (2011), the results are the same suggesting the difference is not sample-driven.

7.4 Addressing Relative Age Differences at Cutoff

To address the confounding effect of differential relative age at the entry date cutoff is difference-in-discontinuity is used (Grembi et al., 2016; Smith, 2016; Canaan, 2020).[52] The effect of relative age at the cutoff is present in the years before the scholarship becomes available. To remove the confounding effect of relative age, one simply subtracts off the effect at the birth cutoff in prior periods

$$\mathrm{T}^{\mathrm{DD}} \equiv (Y^{\text{-}} \text{ - } Y^{+}) \text{ - } (\tilde{Y}^{\text{-}} \text{ - } \tilde{Y}^{+}),$$

where $\tilde{Y}$ are the outcomes at the cutoff in period before the cutoff determines scholarship receipt. To implement this estimator, the data is formatted so that 6 months on either side of each cutoff are related to that particular cutoff. Figure 8 represents how this looks visually.

7.4.1 Identification Assumptions

A new assumption that is introduced by the use of cross-sectional data is the effect of relative age, in the case of no treatment, is constant over time. This is similar to the parallel trands assumption for differences-in-differences, with the important difference that this assumption is local. Under typical IV assumptions (due to the fuzzy cutoff), local smoothness, and local "parallel trends", T^{DD} is a causal effect of receiving the scholarship.

7.4.2 Assumption to Interpret T^{DD} as Local Average Treatment Effect

One more assumption is required for T^{DD} to represent the local causal effect of receiving the scholarship. The final assumption is that the effect of receiving the scholarship does not depend on relative age. Stated differently, there must be no interaction between relative age effects and the effects of the scholarship. There must be additive separability between relative age and receiving the scholarship.

[52] This estimator differs from other RD estimators that exploits longitudinal data, because the variation it uses is cross-sectional, not within-unit. Other estimators include fixed effect RD estimators (Pettersson-Lidbom, 2012), first-difference RD estimator (Lemieux and Milligan, 2008), and dynamic RD designs (Cellini et al., 2010).

7.4.3 Results

Table 1 presents estimates of scholarship receipt on first year academic outcomes. There are varying choices of kernel, sample, covariates, coefficient adjustments, and bandwidth selection criterion to account for several justifiable ways of estimating the effects. Table 1, Panel A focuses on the likelihood of completing 30 credits in ones' first year. In column 1, no bias-correction or robustness corrections to the estimates are done and the scholarship increases the likelihood of completing the scholarship by 36 percentage points, which is significant above 95 percent confidence. In column 2, the bias-correction and robustness adjustments are done (Calonico et al., 2017), and the results are essentially unchanged.

In columns 1 and 2 (in Panel A and B) academically ineligible students are included, making the control group less ideal. In column 3, academically ineligible students are dropped and the effect size decreases to 27 percentage points, probably due to a more academically gifted control group being utilized. Additionally, the statistical significance drops to 90%. In column 4, academic controls for HS GPA and entry test (ACT/SAT) scores are included to compare individuals of similar scores.[53] The effect size drops to 0.24, which is the exact point estimate Scott-Clayton (2011) reports, but the coefficient is no longer statistically significant at conventional levels ($p<0.128$). Column 5 uses CER-optimal bandwidths, which can be more desirable for inference, but less desirable for point estimates. The estimated effect does not change, but the standard error grows, reducing the p-value.

Columns 6 and 7 use a uniform kernel, because it minimizes asymptotic variance under certain conditions (Cattaneo et al., 2019a, pg. 44). The difference in scholarship receipt at the cutoff increases noticeably with the distance from the cutoff for a few months before flattening, so triangular kernels place more weight on potentially unaffected observations. Column 6 uses the MSE-optimal bandwidth criterion with a uniform kernel, the effect size remains similar, but the standard error drops and the effect is statistically significant above

[53] Non-academically eligible students are dropped before academic controls are included, because covariate inclusion in RD assumes that the controls are additively separable from the running variable (Calonico et al., 2019). Dropping non-academically eligible students increases the plausibility of this assumption.

90% (p<0.085). In column 7, a uniform kernel is used with the coverage error rate optimal bandwidths, the effect size grows to a 31 percentage point increase. The standard error also grows, causing this estimate to remain statistically significant at 94.6 percent confidence. The point estimates are similar to Scott-Clayton (2011), sometimes slightly larger depending on bandwidth, sample, controls, and correction.

Table 1 Panel B replicates Panel A, using the likelihood of meeting the renewal GPA of 2.75 as the dependent variable. Column 1 shows the scholarship increases the likelihood of hitting the GPA threshold for renewal by 31 percentage points. The effect is statistically significant above 95%. Columns 2-4 add parameter corrections, subsample to academically eligible students, and include academic controls; all changes leave the effect size and statistical significance almost unchanged. Column 5 uses the CER-optimal bandwidth and finds a 33% larger effect. This point estimate should be interpreted cautiously as the point estimate does not have optimal point estimate properties (Cattaneo et al., 2019a, pg. 44).

Columns 6 and 7 of Table 1, Panel B use the uniform kernel. In column 6, the coefficient is 33% smaller, but the standard error is reduced. Still, the coefficient is only statistically significant at 9% confidence. In column 7, a uniform kernel is used with the CER-optimal bandwidths and the estimate becomes significant at 95 percent confidence. and the effect grows back to near the previously estimated effects. These estimates are similar to the increased likelihood of completing the renewal credits that are presented in Panel A of Table 1, in contrast to Scott-Clayton (2011).

7.5 Reduced Form Plots, Year 3 Outcomes

Figure A.9 shows reduced form plots for outcomes after 3 years. For exceeding 30 credits each year, Figure A.9a shows a negative effect of being the oldest in ones class. Figure A.9b shows a large positive effect of the relative age and the scholarship. The relationship between scholarship and GPA is less visually discernible for reasons discussed below.

7.6 Regression Discontinuity Estimates, Year 3 Outcomes

Table A.3 shows regression discontinuity estimates. Estimates for credits are between 0.3 and 0.4 and are almost always statistically significant at 95%, regardless of bandwidth, kernel, sample, or controls. Point estimates for GPA, in Panel B, are between 0.03 and 0.21, but not statistically significant.

7.7 Difference in Discontinuity Estimates, Year 3 Outcomes

Figure A.10 plots the difference in discontinuity estimates across bandwidths. The confidence intervals are wide at low bandwidths as expected by estimating 2 differences with a modest sample. However, significance is found at longer bandwidths and the bandwidth does not affect the point estimates much. The credits point estimates in Panel A are slightly larger than regression discontinuity estimates, consistent with Figure A.9a showing a negative effect. For some bandwidths, the scholarship also increases GPA renewal, with similar magnitude and statistical significance as credit renewal.

7.8 Difference in Discontinuity Estimates by Gender

Figure 10 presents difference in discontinuity estimates at varying bandwidths, separately by gender. The scholarship is about equally effective for males and females for completing credits requirements. For GPA requirements, female point estimates are very close to 0 at all bandwidths, but male point estimates are consistently near 0.5 percentage points. This heterogeneity is responsible for the failure to estimate statistically significant effects on GPA for the entire sample.

7.8.1 Outcomes By Year (1, 2, 3) By Gender

Figures 11 and 12 shows estimates by year by gender. With this limited sample size, the effects are still noisy.[54] Although not outside the 95% confidence intervals, a general pattern emerges that male point estimates are higher than the female point estimates. The female point estimates straddle 0 for GPA across different years at a bandwidth of 160 days, but for credits they are all positive. This is why the females are equally likely to finish credits each of the first 3 years like males, but not GPA. Figures 11 and 12 support that there might be bigger effects for males at every year and for each renewal threshold and that women may not meet the GPA requirement each year for the first 3 years.

7.9 Coarsened Birthdate

[WEEKS]

[MONTHS]

8 Conclusion

The interaction of two education programs' rules, compulsory Kindergarten entry birthdate cutoffs and inaugural college cohorts being eligible for financial aid, along with an assumption that students finish 1 grade per year on average, is used to estimate causal effects of merit-aid. The Kindergarten birthdate cutoff sorts students into different college cohorts, affecting likelihood of receiving the merit-aid scholarship. Additionally, treatment and control groups are not appreciably dis-similar based on pre-treatment covariates and a McCrary density test (McCrary, 2008). These results imply the alternative design is viable and internally valid.

The effect of the merit-aid scholarship is to increase the likelihood of exceeding renewal thresholds for GPA and credits by 24 to 30 percentage points. Only under very specific assumptions about bandwidth and kernel, those most closely resembling prior work, the point

[54] At 160 days, there's about 550 total observations.

estimate on exceeding the GPA threshold differs from credits point estimate. Additionally, there are heterogeneous effects by gender, receiving the scholarship helps men finish the GPA threshold each year for 3 years, but not for women.

The first contribution of this paper is proposing, formalizing, and validating an alternative design for assessing the causal treatment effects of college merit-aid. This alternative strategy is critical, because prior approaches find evidence that is inconsistent with the identifying assumptions. The running variable, birthdate, is less vulnerable to these criticisms of manipulated variation and selection into/out of sample causing differences in treatment and control groups. Some additional advantages of the birthdate cutoff is it is plausibly more appropriate to extrapolate away from a birthdate cutoff than an academic entry score cutoff and effects are feasibly representative of a greater mix of types of people, therefore more heterogeneous (Lee, 2008; Bloom, 2012).

The second contribution is investigating the large, three-fold discrepancy between pace-based and distinction-based goals that has been found in prior research in a similar sample (Scott-Clayton, 2011). I find no dis-similarity between the estimates for finishing pace-based and distinction-based goals and that the estimates are large, positive, and statistically significant. The lack of a discrepancy and additional results suggest that the discrepancy may be due to functional form and weighting scheme, not a substantive behavior-driven mechanism. The results are more intuitive from the perspective that students must finish both requirements to renew the scholarship in the next year.

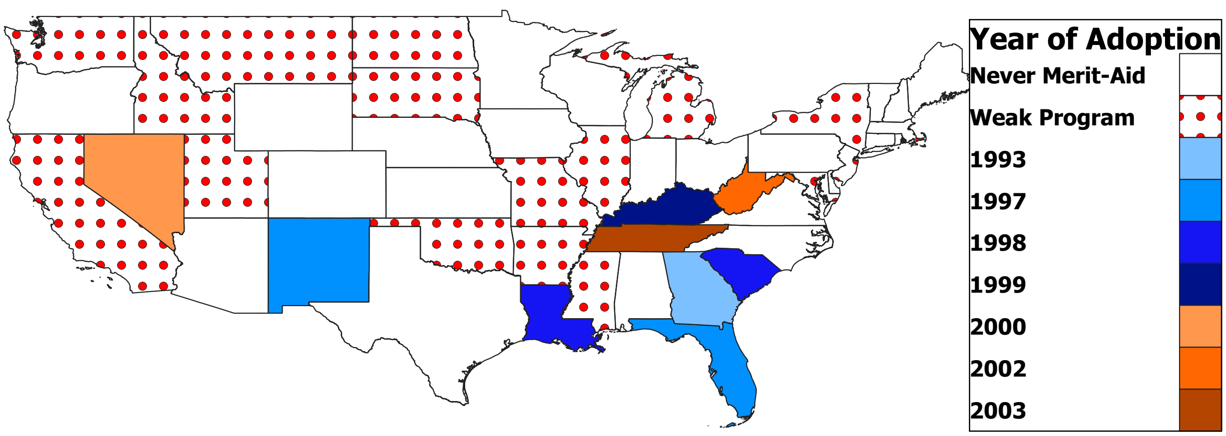
The third contribution is uncovering heterogeneous effects of the scholarship by gender. Receiving the scholarship helps men finish the GPA requirement for each of the first 3 years, but it does not help women finish the GPA requirement. There is no difference for the credits requirement for 3 years by gender. Future research might consider the reason for this gap.

There are some other limitations that would benefit from additional investigation. From an internal validity perspective, it would be better to have more pre-treatment covariates to investigate balance between treatment and control groups. Next, if more pre-treatment data were available, more precise effects could be estimated and the local parallel trends assumption required for difference in discontinuity estimates could be investigated.

A caveat to the results is that the results are from 1 university. This is more of a advantage for future research, because compulsory entry laws are widespread and the assumptions required for this advantageous alternative source of variation are mild and can be investigated with the data. Future research may be able to adapt this design to solve other identification challenges in education policy. The value of c and k are critical, c is not guaranteed to be high enough. In my sample, there is relatively little in-migration which is a headwind to a high c, but the program begins 13 years later which is a tailwind.

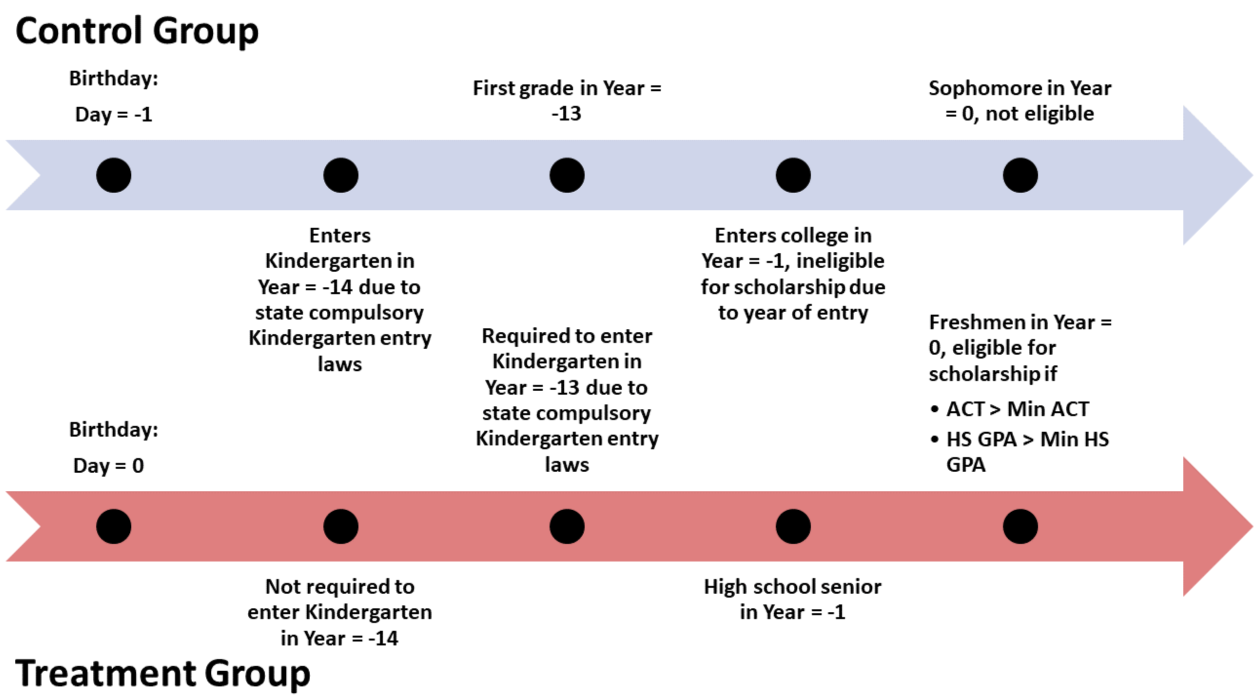
Finally, it is unclear from these results which mechanism, money or motivation, is the cause of the large treatment effects. However, this paper takes a small step towards finding a more defendable research design. Using this alternative design may be informative for disentangling mechanisms and heterogeneity in future work.

Figure 1: Merit-Aid Adoption by State and Time



Note: Table from Cowan and White (2015). Weak program means the program offers aid for a few students and/or covers small amount of costs, Weak program dates of start not shown.

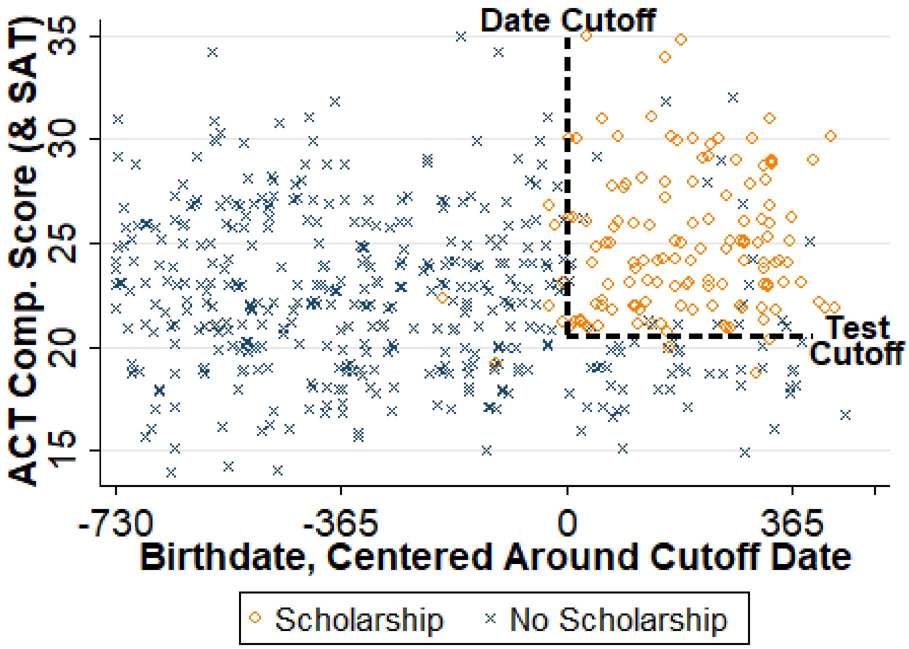
Figure 2: Illustration of College Financial Aid Differentially Assigned by Small Changes in Birth Date



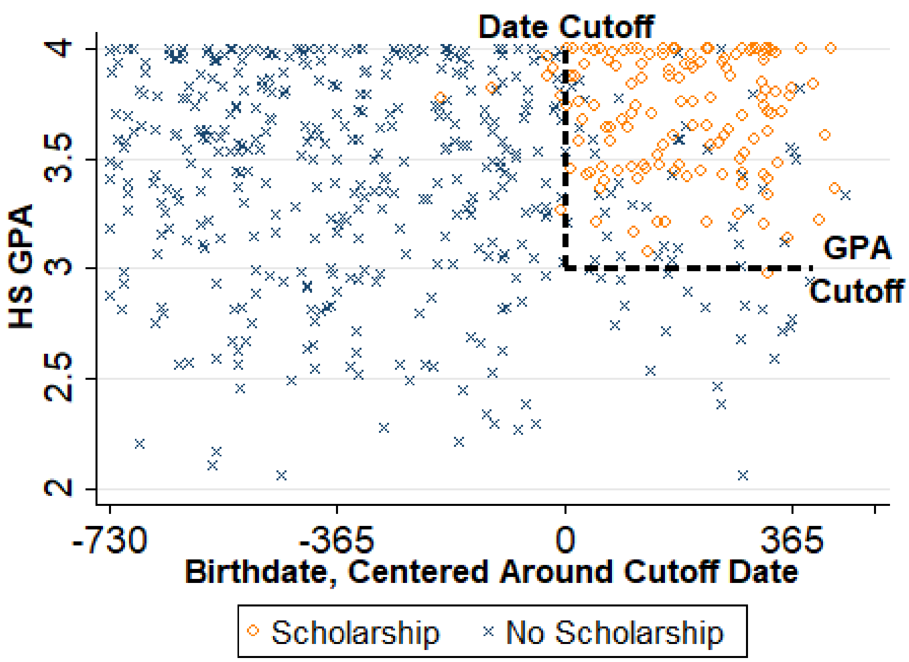
Note: Year 0 stands for initial year of this college scholarship. Day = -1 is a student born right before the birthdate cutoff to enter Kindergarten in the first cohort that is eligible for the college scholarship if they finish 1 grade per year. Day = 0 is a student born right after the cutoff which allows them to enter in the the Kindergarten cohort that will enter college in the first college eligible cohort, finishing 1 grade per year.

Figure 3: Eligibility Measures and Scholarship Receipt

Panel A: Birthdate, Entry Test Score, and Scholarship

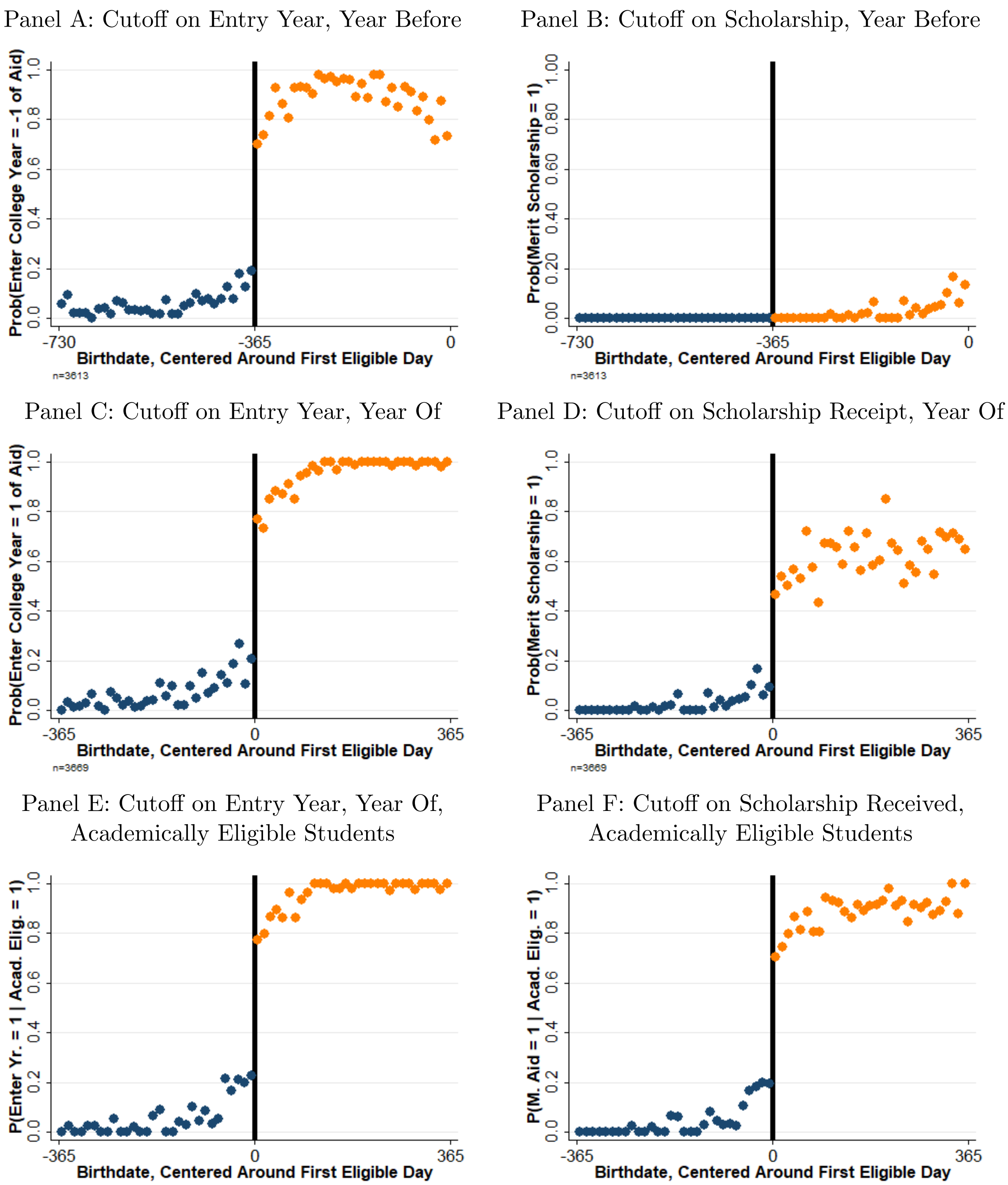


Panel B: Birthdate, HS GPA, and Scholarship



Note: 10% random sample. N = 559. Birthdates centered at 0, the first eligible date for the scholarship. The cutoff ACT score is 20.5. SAT score converted to ACT using the conversion chart at Edwards (2020). The cutoff HS GPA is 3.0.

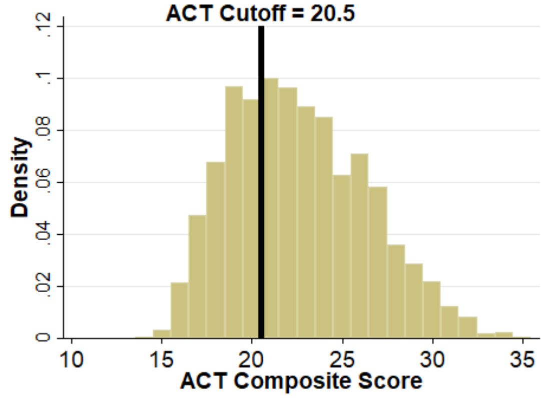
Figure 4: First Stage Discontinuities at Kindergarten Entry Date Cutoffs



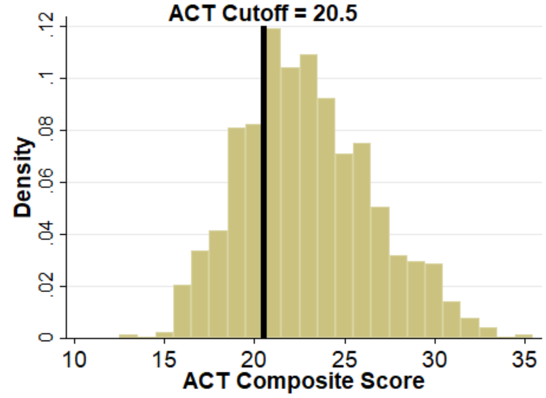
Note: Bin-length = 11.375 days. Plots are the conditional mean within bins. The red vertical line indicates cutoff birthdates. The birth-month is centered at the cutoff birthdate, meaning 0 is the first birthdate of increased likelihood of treatment. In Panels A, C, and E, the y-variable is 1/0 for being a first time freshman in the first year of the scholarship. In Panels B, D, and F, the y-variable is mean of receiving at least 1 year of financial merit-based aid.

Figure 5: Comparing Academic and Birthdate Running Variables' Continuity Around Cutoffs

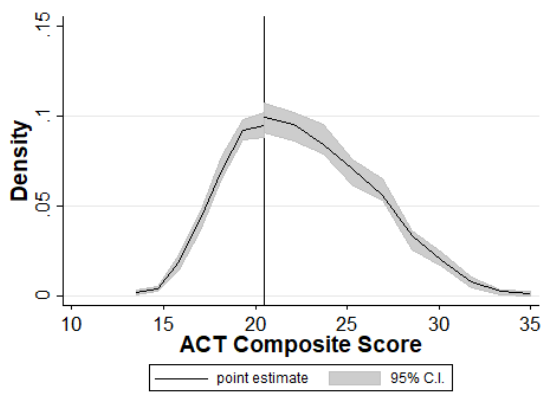
Panel A: ACT Scores, Two Cohorts Pre-Scholarship



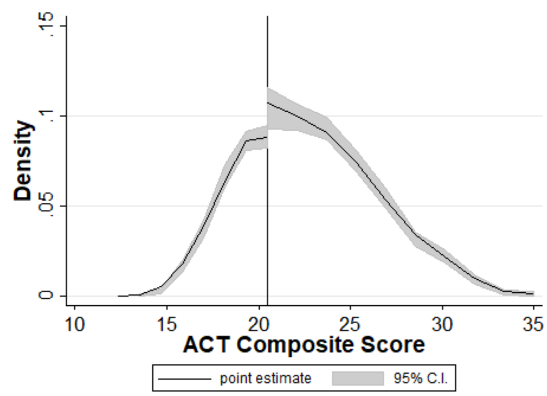
Panel B: ACT Scores, First Cohort of Scholarship



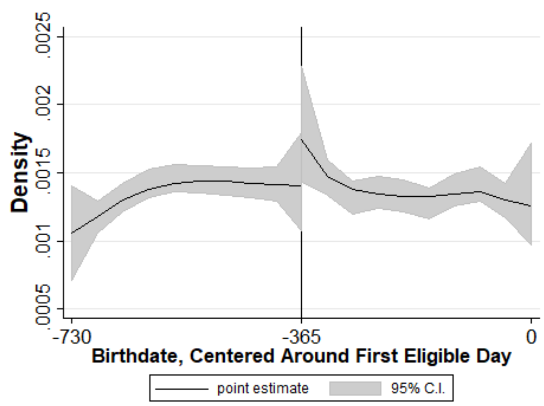
Panel C: McCrary Density ACT Scores, Two Cohorts Pre-Scholarship



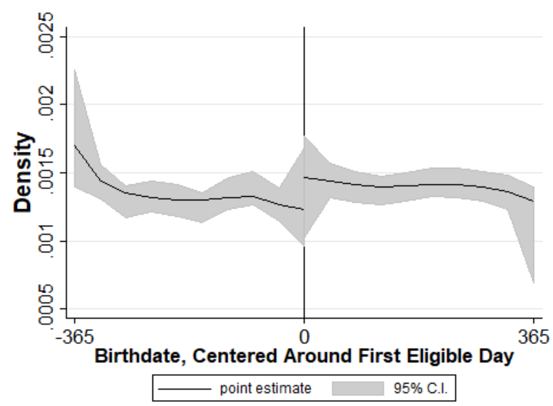
Panel D: McCrary Density ACT Scores, First Cohort of Scholarship



Panel E: McCrary Density Test, Birthdate Running Variable, Two Cohorts Pre-Scholarship

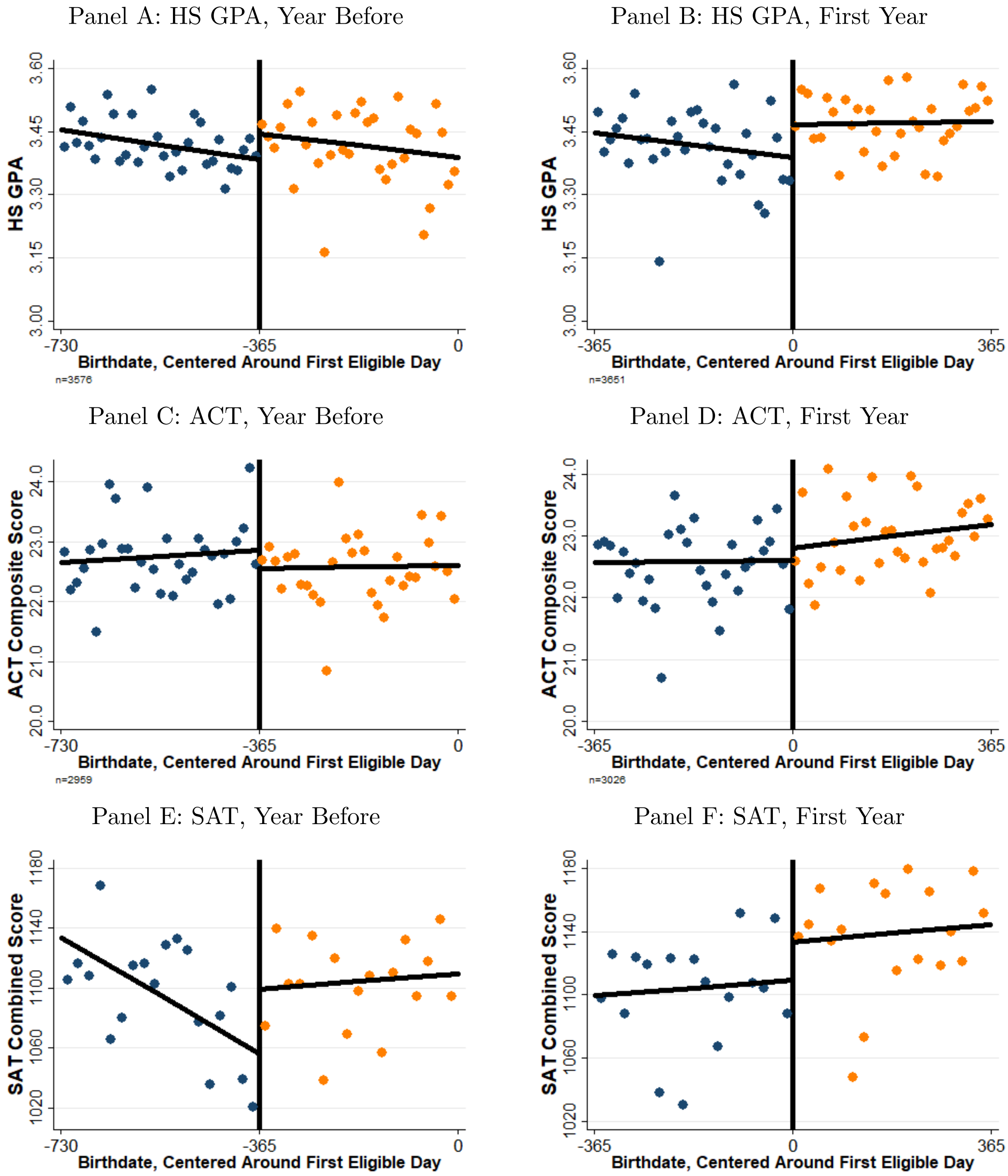


Panel F: McCrary Density Test, Birthdate Running Variable, First Scholarship Cohort



Note: Panels C-F estimated using rddensity in Stata. Optimal bandwidths are calculated according to Cattaneo et al. (2019b). Birthdate centered means 0 equals the relevant date for discontinuously changing scholarship receipt. The cutoff in Panel E is -365, because the cutoff is the same date - 1 year in the year before.

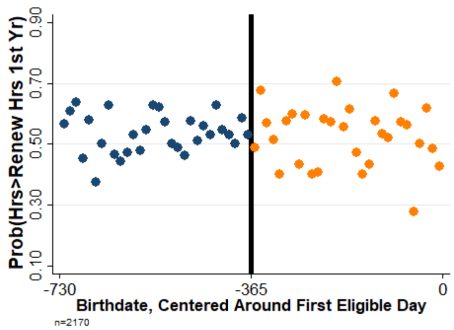
Figure 6: Pre-Treatment Covariate Continuity at Date Cutoff



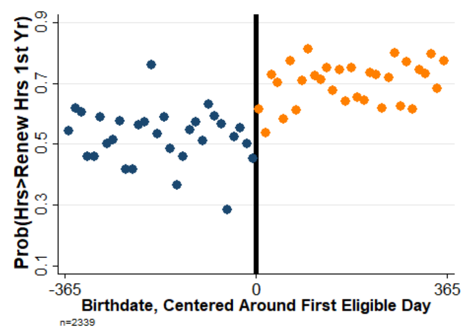
Note: Red line represents cutoff birthdate. Panels A, C, and E use the year prior to the scholarship being available cutoff, 365 days before 0. Panels B, D, and F use the year where the cutoff determines eligibility. Points are means conditional on point being within a bandwidth (bw) of h of a given birthdate. Plot A and B bw = 11.375. Plots C and D bw = 11.74. Plots E and F bw = 20.11. Bandwidths chosen as default histogram option. Specifically: number of bins chosen as: $min(\sqrt{N}, \frac{10ln(N)}{ln(10)})$, where N is number of observations.

Figure 7: Reduced Form Effect of Birthdate on Likelihood of Exceeding Renewal Thresholds in Freshman Year, Including Before Scholarship Began, Academically Eligible

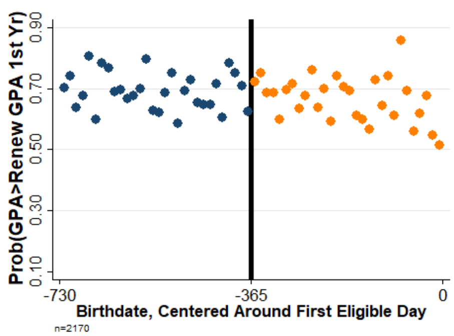
Panel A: Exceed 30 Credits, Pre-Scholarship, Academic Eligible Students



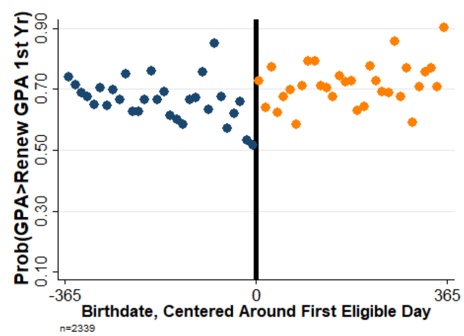
Panel B: Exceed 30 Credits, Scholarship Year, Academic Eligible Students



Panel C: Exceed 2.75 Freshman GPA, Pre-Scholarship, Academic Eligible Students



Panel D: Exceed 2.75 Freshman GPA, Scholarship Year, Academic Eligible Students



Note: Red line represents cutoff birthdate. Panels A & C use the year prior to the scholarship being available cutoff, 365 days before 0. Panels B & D use the year where the cutoff determines scholarship eligibility. Points are means conditional on point being within a bandwidth (bw) of h of a given birthdate. Plot A bw = 14, 15.125. Plot B bw = 14. Plot C bw = 11.75. Plot D bw = 12.14. Plot D bw = 14. Plots A-D Bandwidths chosen as default histogram option. Specifically: number of bins chosen as: $min(\sqrt{N}, \frac{10ln(N)}{ln(10)})$, where N is number of observations.

Table 1: Local Linear Estimates of Scholarship on Meeting First Year Credits and GPA Renewal Thresholds

Panel A: Dep. Var.: Renew Credits							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scholarship	0.36**	0.35**	0.27*	0.24	0.24	0.23*	0.31*
	(0.15)	(0.18)	(0.16)	(0.16)	(0.19)	(0.13)	(0.16)
Corrections	-	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R
Sample	All	All	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.
Bandwidth	141	141	135	134	91	133	90
Eff. Obs. Left	680	680	395	395	248	393	248
Eff. Obs. Right	740	740	471	471	319	462	319
Obs-Left	1970	1970	1081	1081	1081	1081	1081
Obs-Right	1979	1979	1356	1356	1356	1356	1356
Acad. Controls				X	X	X	X
BW Criterion	MSE	MSE	MSE	MSE	CER	MSE	CER
Kernel	Tri	Tri	Tri	Tri	Tri	Uni	Uni
Panel B: Dep. Var.: Renew GPA							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scholarship	0.31**	0.34**	0.34**	0.30**	0.40**	0.20	0.32**
	(0.15)	(0.18)	(0.16)	(0.15)	(0.19)	(0.13)	(0.16)
Corrections	-	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R
Sample	All	All	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.
Bandwidth	151	151	140	142	96	138	93
Eff. Obs. Left	720	720	404	408	268	398	261
Eff. Obs. Right	777	777	491	500	336	482	327
Obs-Left	1970	1970	1081	1081	1081	1081	1081
Obs-Right	1979	1979	1356	1356	1356	1356	1356
Acad. Controls				X	X	X	X
BW Criterion	MSE	MSE	MSE	MSE	CER	MSE	CER
Kernel	Tri	Tri	Tri	Tri	Tri	Uni	Uni

Note: * p<0.1, ** p<0.05, *** p<0.01. Estimates displayed are the second stage Wald estimator from a fuzzy regression discontinuity design using birthdate in relation to Kindergarten entry date as the cutoff/instrument. The estimate is the effect of receiving the Promise scholarship. In Panel A, the dependent variable is a dummy variable for complete 30 credits (renewal amount) in first year of college. In Panel B, the dependent variable is a dummy variable for receive at least a 2.75 GPA (renewal GPA) in first year of college. In all each column and panel, the first stage cutoff has a strong effect on scholarship receipt, consistent with 4d. In column 1, conventional estimates are reported. In columns 2 and beyond the bias-correction, robust estimates are reported (Calonico et al., 2017). In columns 1 and 2, the full sample is used. Columns 3 and beyond use only academically eligible students. From column 4 onwards, high school GPA and entry test (ACT/SAT) score are included. In columns 1-4 and 6, the bandwidth selection procedure optimizes mean square error (MSE). In columns 5 and 7, the bandwidth selector optimizes coverage error rate. In columns 1-5 a triangular kernel is used. In columns 6 and 7 a uniform kernel is used.

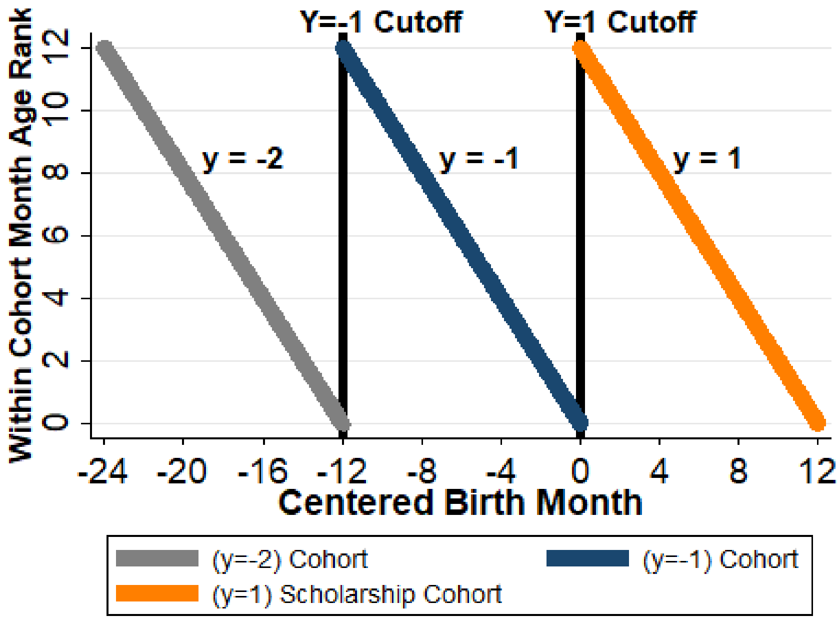
Table 2: Global Estimates of Promise Scholarship on Exceeding Renewal GPA

	(1) All Months	(2) Drop 3 Months	(3) Drop 6 Months
Panel A: Linear			
Received Any Promise	0.08*	0.10*	0.16**
	(0.04)	(0.05)	(0.06)
Kleibergen-Paap F	190.6	140.3	224.7
Observations	3003	2268	1515
Panel B: Linear, Academic Controls			
Received Any Promise	0.07	0.08	0.13
	(0.06)	(0.07)	(0.11)
Kleibergen-Paap F	455.5	287.2	124.4
Observations	2988	2256	1505
Panel C: Quadratic			
Received Any Promise	0.17	0.25	0.30
	(0.13)	(0.18)	(0.26)
Kleibergen-Paap F	104.5	53.23	27.27
Observations	3003	2268	1515
Panel D: Quadratic, Academic Controls			
Received Any Promise	0.14	0.20	0.31
	(0.12)	(0.17)	(0.24)
Kleibergen-Paap F	111.3	54.38	30.27
Observations	2988	2256	1505

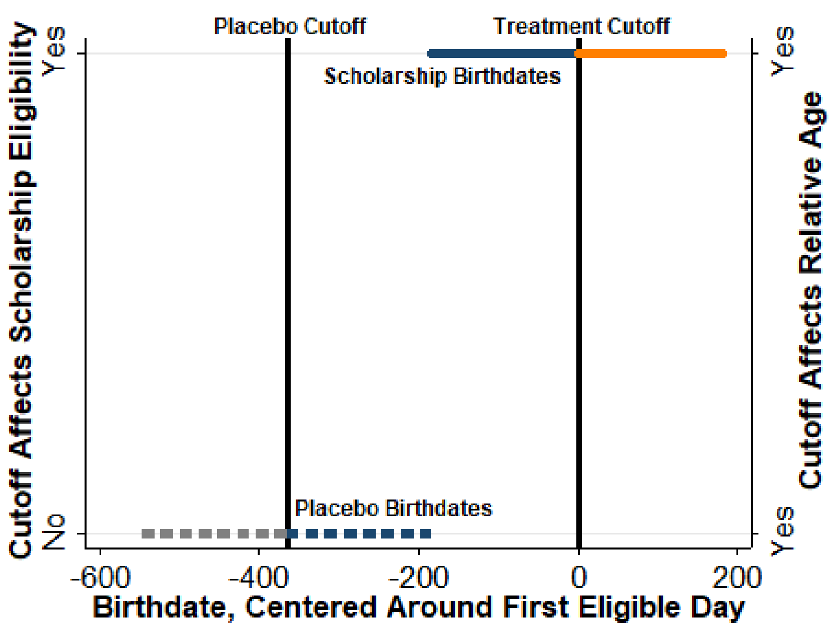
Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parentheses. Dependent variable is a dummy variable for whether a student exceeded a 2.75 GPA in their freshman year. All panels display 2SLS regressions, fit separately on either side of the birth-date cutoff. Only academically eligible students included. Column 1 uses all birth months. Column 2 drops 3 months, on both sides, farthest from cutoff. Column 3 drops 6 months, on both sides, farthest from the cutoff. Months are dropped, because students far away bias the estimate of interest. Panel A assumes linearity. Panel B assumes linearity and includes HS GPA and entry-test (ACT or SAT) score. SAT is converted to ACT score using College Board conversion chart. Panel C assumes a quadratic functional form. Panel D assumes a quadratic functional form and includes same academic controls as Panel B.

Figure 8: Cohort Age Rank and Difference in Discontinuity Visualization

Panel A: Discontinuous Age Rank At Cutoff



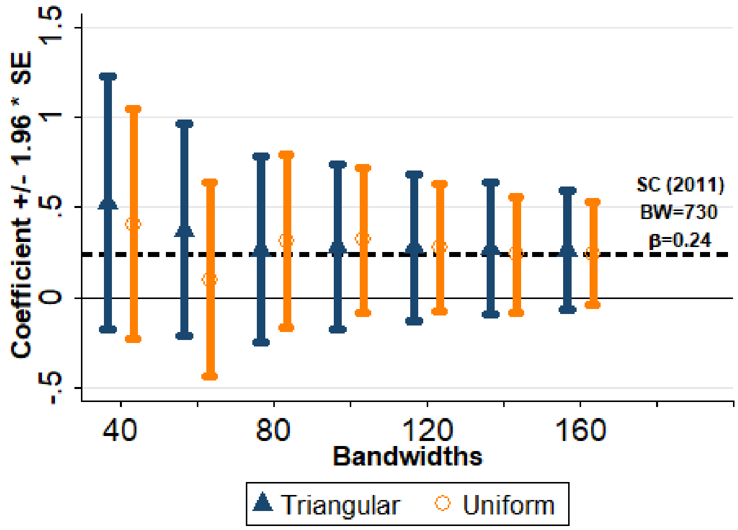
Panel B: Difference in Discontinuities Design



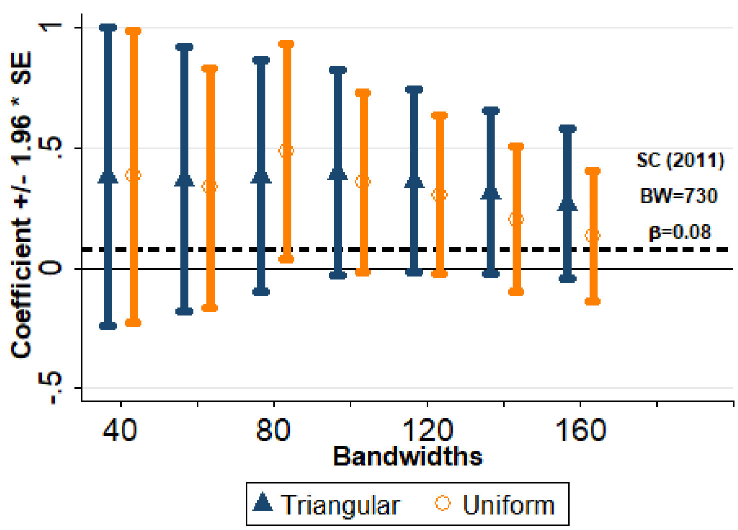
Note: Panel A shows age rank (in months), relative to cohort, on the y-axis. Birth month centered at 0 is the x-axis. Day 0 is the first day where scholarship eligibility discontinuously increases. In every cohort, age rank is a function declining at the same linear rate. Panel B sketches out the canonical difference-in-discontinuity estimate. For the placebo birthdates, relative age affects outcomes at the placebo cutoff. For the scholarship birthdates, relative age and the scholarship affect outcomes at the treatment cutoff. There are 182 days from each cutoff to both sides.

Figure 9: Coefficient Estimates from Difference in Discontinuity on Freshman Outcomes

Panel A: Probability of Completing Renewal Credits



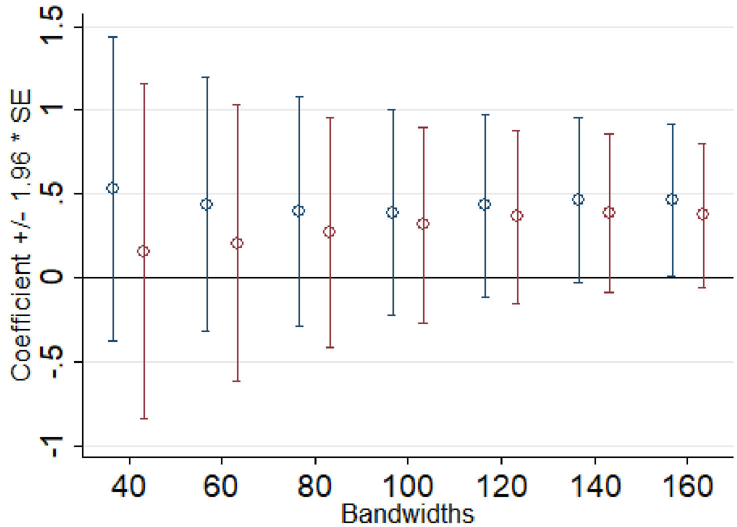
Panel B: Probability of Completing Renewal GPA



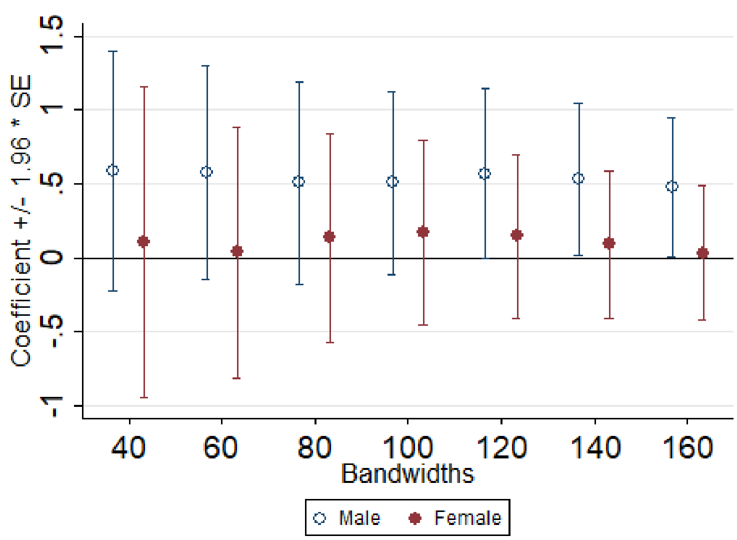
Note: Coefficients represent "conventional", not bias-corrected or robust for MSE-optimal, treatment estimates from difference-in-discontinuity estimates (Grembi et al., 2016). Diff-in-discontinuity is the RD estimate from the kindergarten entry date cutoff before it determined the scholarship receipt, but did determine relative age subtracted from the RD estimate in where it did determine scholarship receipt. This approach differences out the relative age effect under fairly mild assumptions. The different bandwidths indicate the number of days used from the respective cutoffs. Triangular and uniforms kernels used. Panel B legend valid for both. All regressions subsample to academically eligible students meaning [(ACT 20.5 OR SAT $\geq$ 1000) AND HS GPA 3.0] and include controls for HS GPA and ACT composite score (w/ SAT converted to ACT). Panel A dependent variable is exceed renewal credits in freshman year. Panel B dependent variable is renewal GPA in freshman year. The horizontal dashed lines represent the Scott-Clayton (2011) point estimates which correspond approximately to 730 day bandwidths, since Scott-Clayton (2011) uses 2 cohorts before and after.

Figure 10: Coefficient Estimates from Difference in Discontinuity on 3 Year Outcomes

Panel A: Probability of Completing Renewal Credits



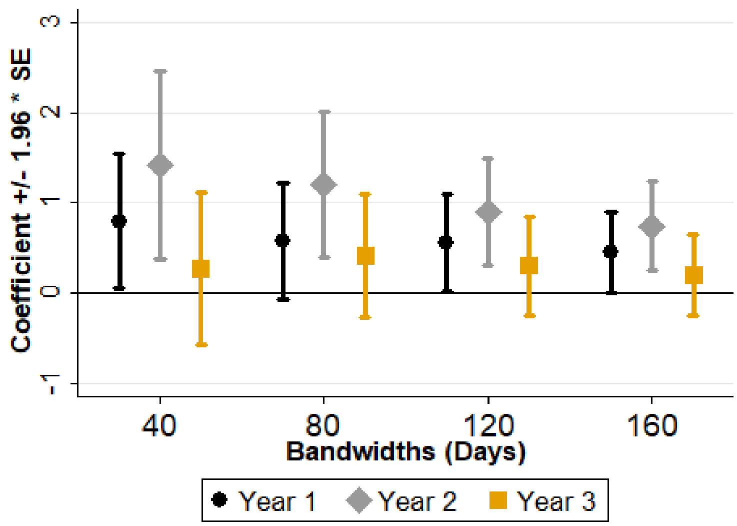
Panel B: Probability of Completing Renewal GPA



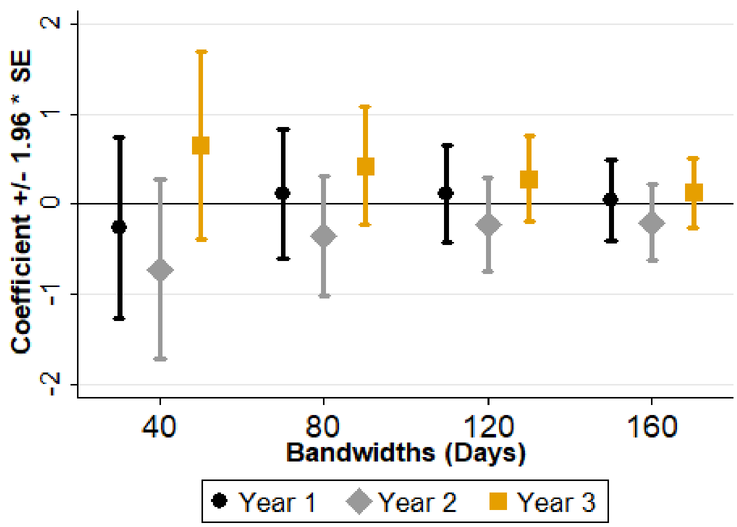
Note: Coefficients represent "conventional", not bias-corrected or robust for MSE-optimal, treatment estimates from difference-in-discontinuity estimates (Grembi et al., 2016). Diff-in-discontinuity is the RD estimate from the kindergarten entry date cutoff before it determined the scholarship receipt, but did determine relative age subtracted from the RD estimate in where it did determine scholarship receipt. This approach differences out the relative age effect under fairly mild assumptions. The different bandwidths indicate the number of days used from the respective cutoffs. Triangular and uniforms kernels used. Panel B legend valid for both. All regressions subsample to academically eligible students meaning [(ACT 20.5 OR SAT $\geq$ 1000) AND HS GPA 3.0] and include controls for HS GPA and ACT composite score (w/ SAT converted to ACT). Panel A dependent variable is exceed renewal credits in freshman year. Panel B dependent variable is renewal GPA in freshman year. The horizontal dashed lines represent the Scott-Clayton (2011) point estimates which correspond approximately to 730 day bandwidths, since Scott-Clayton (2011) uses 2 cohorts before and after.

Figure 11: GPA Estimates by Year by Gender

Panel A: Probability of Completing Renewal GPA, Male



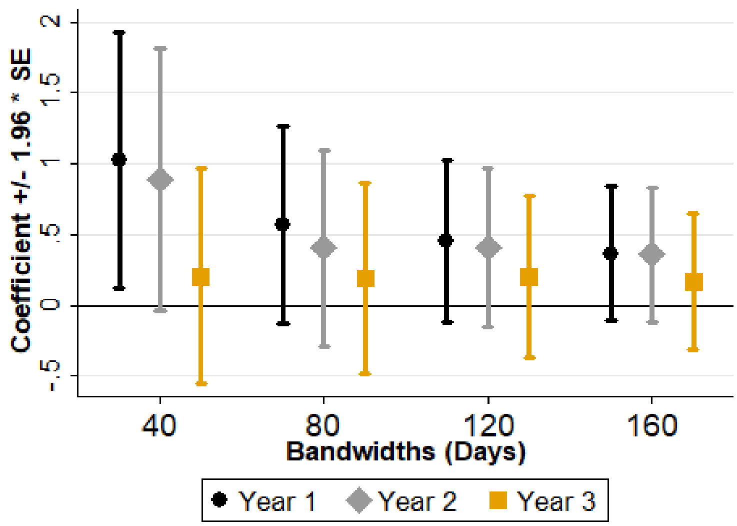
Panel B: Probability of Completing Renewal GPA, Female



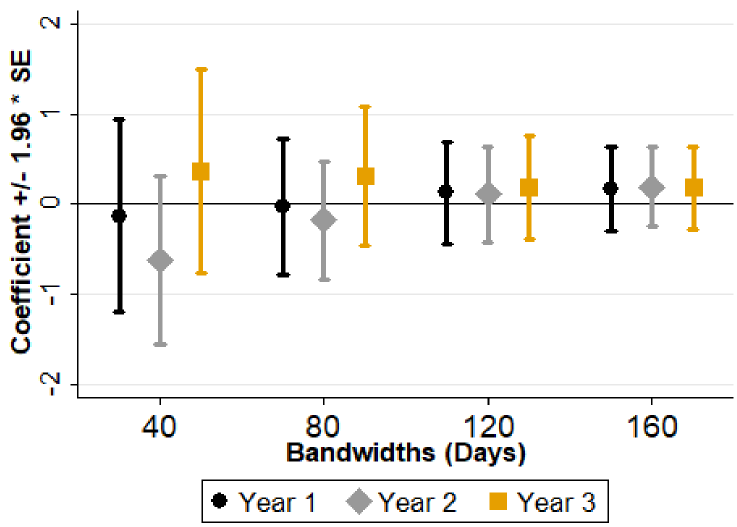
Note: Dependent variable is: finish renewal GPA in that year only. Panel A is male. Panel B is female. N $\leq$ 524 in Panel A. N $\leq$ 541 in Panel B. Coefficients represent "conventional", not bias-corrected or robust for MSE-optimal, treatment estimates from difference-in-discontinuity estimates (Grembi et al., 2016). All regressions subsample to academically eligible students meaning [(ACT 20.5 OR SAT $\geq$ 1000) AND HS GPA 3.0] and include controls for HS GPA and ACT composite score (w/ SAT converted to ACT).

Figure 12: Credits Estimates by Year by Gender

Panel A: Probability of Completing Renewal Credits, Male



Panel B: Probability of Completing Renewal Credits, Female



Note: Dependent variable is: finish renewal credits in that year only. Panel A is male. Panel B is female. N $\leq$ 524 in Panel A. N $\leq$ 541 in Panel B. Coefficients represent "conventional", not bias-corrected or robust for MSE-optimal, treatment estimates from difference-in-discontinuity estimates (Grembi et al., 2016). All regressions subsample to academically eligible students meaning [(ACT 20.5 OR SAT $\geq$ 1000) AND HS GPA 3.0] and include controls for HS GPA and ACT composite score (w/ SAT converted to ACT).

References

Abbott, B., Gallipoli, G., Meghir, C., and Violante, G. L. (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*, 127(6):2569–2624.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

Angrist, J. D. and Rokkanen, M. (2015). Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association*, 110(512):1331–1344.

Bertanha, M. and Imbens, G. W. (2019). External validity in fuzzy regression discontinuity designs. *Journal of Business & Economic Statistics*, pages 1–39.

Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., and Stevens, M. (2019). The long-run impacts of financial aid: Evidence from California's Cal Grant. *American Economic Journal: Economic Policy*, 11(1):64–94.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2011). Too young to leave the nest? The effects of school starting age. *The Review of Economics and Statistics*, 93(2):455–467.

Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, 5(1):43–82.

Bordón, P., Canals, C., and Mizala, A. (2020). The gender gap in college major choice in Chile. *Economics of Education Review*, 77:102011.

Bruce, D. J. and Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, 81:30–44.

Buckles, K. S. and Hungerman, D. M. (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics*, 95(3):711–724.

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

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

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Canaan, S. (2020). The long-run effects of reducing early school tracking. *Journal of Public Economics*, 187:104206.

Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100(1):1–40.

Carnevale, A. P., Cheah, B., and Hanson, A. R. (2015). The economic value of college majors. Technical report, Center on Education and the Workforce.

Carruthers, C. K. and Özek, U. (2016). Losing HOPE: Financial aid and the line between college and work. *Economics of Education Review*, 53:1–15.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019a). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Jansson, M., and Ma, X. (2019b). Simple local polynomial density estimators. *Journal of the American Statistical Association*, pages 1–7.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016). Inference in regression discontinuity designs under local randomization. *The Stata Journal*, 16(2):331–367.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2017). Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality. *Journal of Policy Analysis and Management*, 36(3):643–681.

Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1):215–261.

Clark, D., Gill, D., Prowse, V., and Rush, M. (2016). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, pages 1–45.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, 103(6):2087–2120.

Cohodes, S. R., Grossman, D. S., Kleiner, S. A., and Lovenheim, M. F. (2016). The effect of child health insurance access on schooling: Evidence from public insurance expansions. *Journal of Human Resources*, 51(3):727–759.

College Board (2019). Average aid per student over time. https://research.collegeboard.org/trends/student-aid/figures-tables/average-aid-student-over-time.

Cook, T. D. (2008). "Waiting for life to arrive": A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, 142(2):636–654.

Cornwell, C. and Mustard, D. B. (2007). Merit-based college scholarships and car sales. *Education Finance and Policy*, 2(2):133–151.

Cornwell, C. M., Lee, K. H., and Mustard, D. B. (2005). Student responses to merit scholarship retention rules. *Journal of Human Resources*, 40(4):895–917.

Cowan, B. W. (2016). Testing for educational credit constraints using heterogeneity in individual time preferences. *Journal of Labor Economics*, 34(2):363–402.

Cowan, B. W. and White, D. R. (2015). The effects of merit-based financial aid on drinking in college. *Journal of Health Economics*, 44:137–149.

Denning, J. T. (2019). Born under a lucky star? Financial aid, college completion, labor supply, and credit constraints. *Journal of Human Resources*, 54(3):760–784.

Dong, Y. (2018). Alternative assumptions to identify LATE in fuzzy regression discontinuity designs. *Oxford Bulletin of Economics and Statistics*, 80(5):1020–1027.

Dong, Y. and Lewbel, A. (2015). Identifying the effect of changing the policy threshold in regression discontinuity models. *Review of Economics and Statistics*, 97(5):1081–1092.

Dynarski, S. (2004). The new merit aid. In *College choices: The economics of where to go, when to go, and how to pay for it*, pages 63–100. University of Chicago Press.

Dynarski, S. and Scott-Clayton, J. (2013). Financial aid policy: Lessons from research. Technical report, National Bureau of Economic Research.

Edwards, H. (2020). Official act to sat (new 1600 and old 2400) conversion charts. https://blog.prepscholar.com/act-to-sat-conversion.

Fitzpatrick, M. D. and Jones, D. (2016). Post-baccalaureate migration and merit-based scholarships. *Economics of Education Review*, 54:155–172.

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

Goodman, J. (2008). Who merits financial aid? *Journal of Public Economics*, 92(10-11):2121–2131.

Grembi, V., Nannicini, T., and Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Hsu, Y.-C. and Shen, S. (2019). Testing treatment effect heterogeneity in regression discontinuity designs. *Journal of Econometrics*, 208(2):468–486.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959.

Jia, N. (2019). Heterogeneous effects of merit scholarships: Do program features matter? *Applied Economics*, 51(27):2963–2979.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Koch, A. K. and Nafziger, J. (2011). Self-regulation through goal setting. *Scandinavian Journal of Economics*, 113(1):212–227.

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, 108(8):2277–2304.

Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675–697.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674.

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

Lemieux, T. and Milligan, K. (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics*, 142(2):807–828.

Lleras-Muney, A. (2002). Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939. *The Journal of Law and Economics*, 45(2):401–435.

Matta, R., Ribas, R. P., Sampaio, B., and Sampaio, G. R. (2016). The effect of age at school entry on college admission and earnings: A regression-discontinuity approach. *IZA Journal of Labor Economics*, 5(1):9.

McCann, A. (2020). Most & Least Educated States in America. https://wallethub.com/edu/e/most-educated-states/31075/.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Mourifie, I., Henry, M., and Méango, R. (2020). Sharp bounds and testability of a Roy model of STEM major choices. *Journal of Political Economy*, 128(8):3220–3283.

Ortagus, J., Kelchen, R., Rosigner, K., and Voorhees, N. (2018). Performance-based funding in American higher education: A systematic synthesis of the intended and unintended consequences. *Educational Evaluation and Policy Analysis*, 53(2):299–329.

Page, L. C., Kehoe, S. S., Castleman, B. L., and Sahadewo, G. A. (2019). More than dollars for scholars the impact of the Dell Scholars Program on college access, persistence, and degree attainment. *Journal of Human Resources*, 54(3):683–725.

Pettersson-Lidbom, P. (2012). Does the size of the legislature affect the size of government? evidence from two natural experiments. *Journal of Public Economics*, 96(3-4):269–278.

Qureshi, J. and Gangopadhyaya, A. (2019). Childhood medicaid eligibility and human capital. Technical report.

Scott-Clayton, J. (2011). On money and motivation a quasi-experimental analysis of financial incentives for college achievement. *Journal of Human Resources*, 46(3):614–646.

Scott-Clayton, J. and Zafar, B. (2019). Financial aid, debt management, and socioeconomic outcomes: Post-college effects of merit-based aid. *Journal of Public Economics*, 170:68–82.

Sjoquist, D. L. and Winters, J. V. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, 80:39–50.

Sjoquist, D. L. and Winters, J. V. (2015a). State merit aid programs and college major: A focus on STEM. *Journal of Labor Economics*, 33(4):973–1006.

Sjoquist, D. L. and Winters, J. V. (2015b). State merit-based financial aid programs and college attainment. *Journal of Regional Science*, 55(3):364–390.

Smith, A. C. (2016). Spring forward at your own risk: Daylight saving time and fatal vehicle crashes. *American Economic Journal: Applied Economics*, 8(2):65–91.

Solis, A. (2017). Credit access and college enrollment. *Journal of Political Economy*, 125(2):562–622.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309.

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

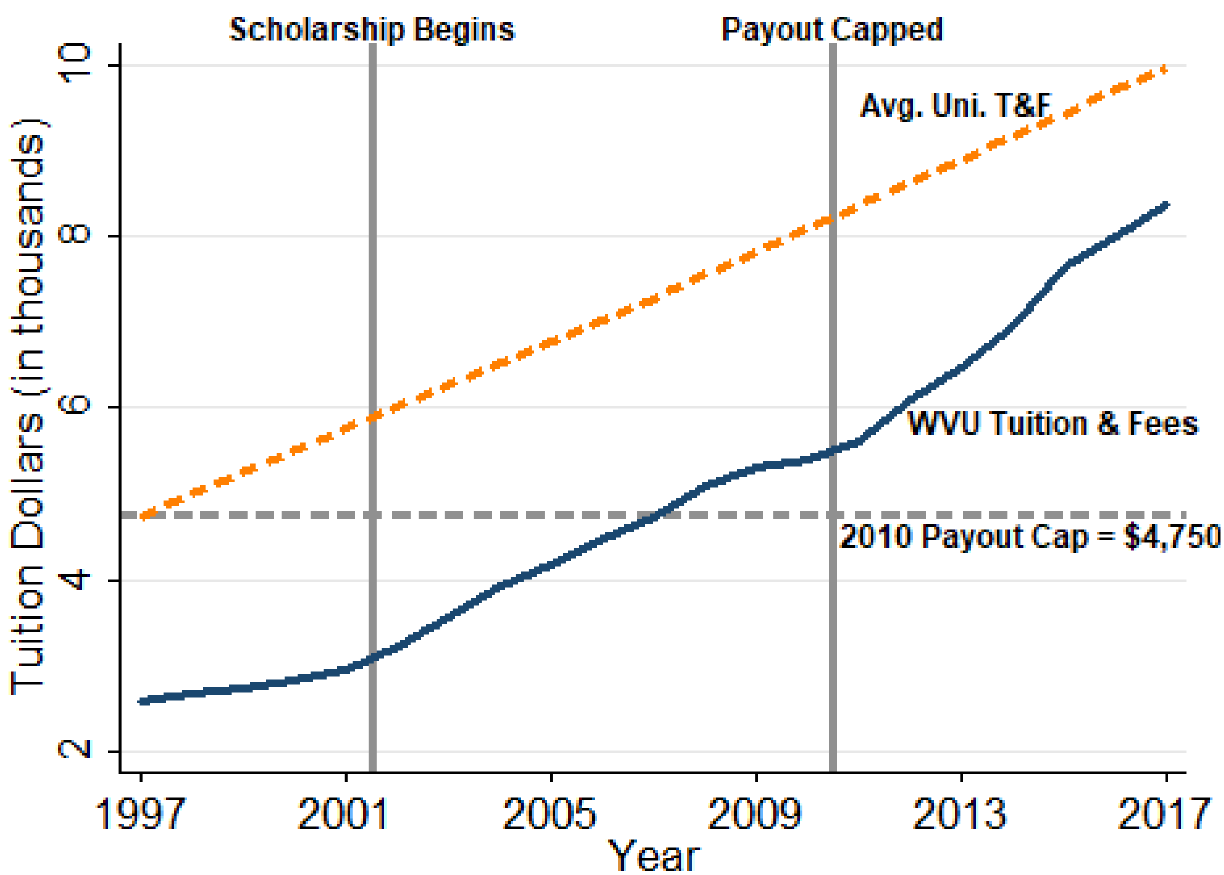
West Virginia Higher Education Policy Commission (2014). Final report of promise scholarship ad-hoc advisory comittee. http://www.wvhepc.edu/wp-content/uploads/2014/01/PROMISEfinalreport.pdf.

Whaley, M. (1985). The status of kindergarten: A survey of the states.

Zhang, L., Hu, S., Sun, L., and Pu, S. (2016). The effect of Florida's Bright Futures Program on college choice: A regression discontinuity approach. *The Journal of Higher Education*, 87(1):115–146.

A Online Appendix

Figure A.1: Trends in Cost of College and West Virginia Promise Scholarship Program



Note:

Table A.1: Descriptive Statistics By Academic and Temporal Eligiblity

Panel A: Acad Elig., Born Day ≥ 0						
	Mean	Median	Std Dev	Min	Max	Count
At Least 1 Year Scholarship	0.89	1	0.31	0	1	1356
Freshman Credits	29.16	31	7.46	0	47	1356
Freshman Renewal Credits	0.70	1	0.46	0	1	1356
Freshman GPA	3.05	3	0.70	0	4	1335
Freshman Renew GPA	0.72	1	0.45	0	1	1356
High School GPA	3.67	4	0.28	3	4	1356
Act Composite Scores	24.71	24	2.97	21	35	1114
Sat Combined Scores	1195.83	1190	125.60	1000	1560	242

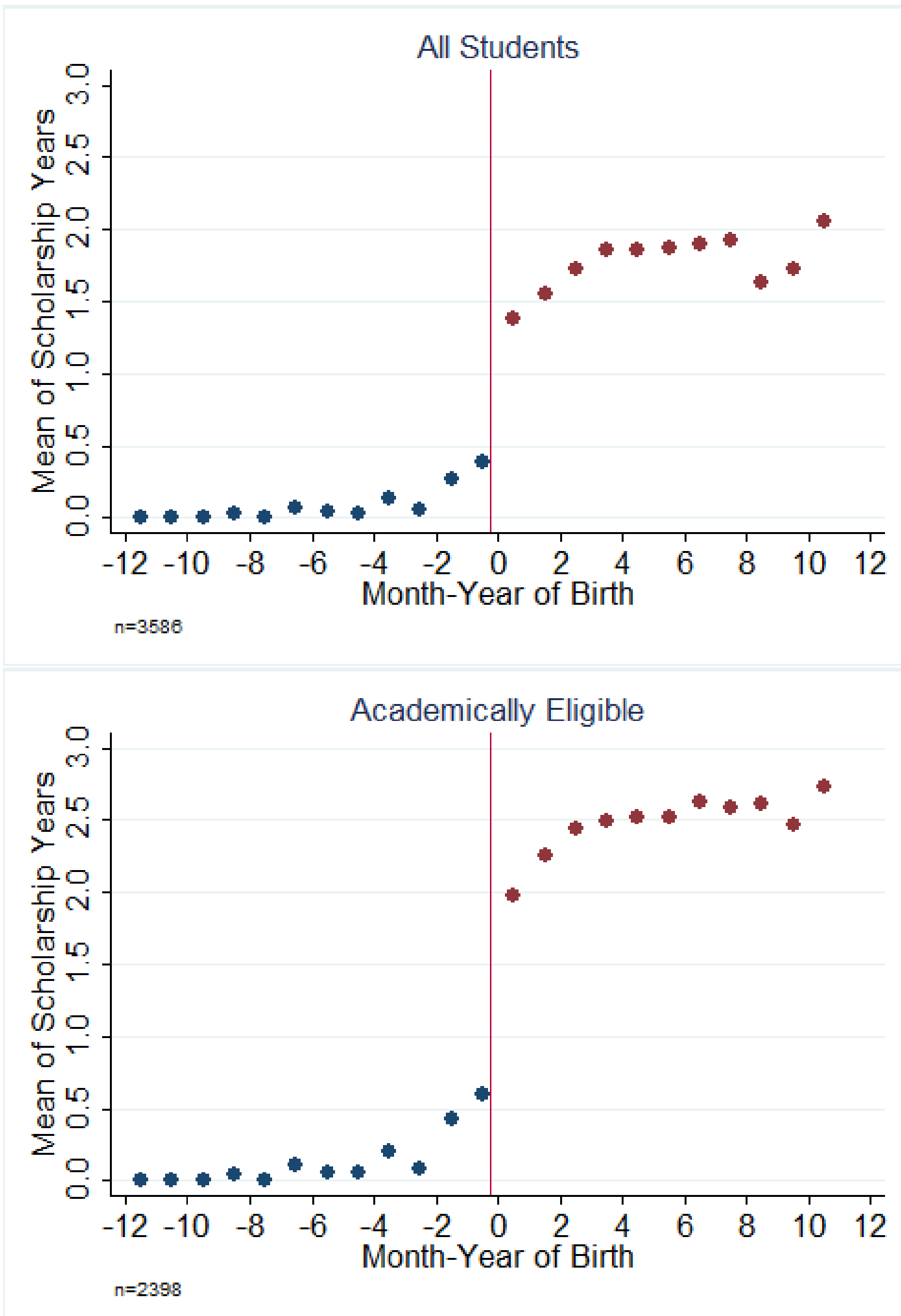
Panel B: Acad Elig., Born Day < 0						
	Mean	Median	Std Dev	Min	Max	Count
At Least 1 Year Scholarship	0.04	0	0.20	0	1	1081
Freshman Credits	27.63	30	8.31	0	53	1081
Freshman Renewal Credits	0.52	1	0.50	0	1	1081
Freshman GPA	2.97	3	0.74	0	4	1055
Freshman Renew GPA	0.66	1	0.47	0	1	1081
High School GPA	3.67	4	0.29	3	4	1081
Act Composite Scores	24.72	24	2.83	21	34	874
Sat Combined Scores	1176.81	1170	119.17	1000	1550	207

Panel C: Not Acad Elig., Born Day≥ 0						
	Mean	Median	Std Dev	Min	Max	Count
At Least 1 Year Scholarship	0.05	0	0.21	0	1	617
Freshman Credits	21.86	25	9.88	0	49	617
Freshman Renewal Credits	0.24	0	0.43	0	1	617
Freshman GPA	2.21	2	0.79	0	4	585
Freshman Renew GPA	0.30	0	0.46	0	1	617
High School GPA	3.04	3	0.41	2	4	617
Act Composite Scores	19.56	19	2.23	13	30	519
Sat Combined Scores	1003.06	980	130.40	780	1420	98

Panel D: Not Acad Elig., Born Day <0						
	Mean	Median	Std Dev	Min	Max	Count
At Least 1 Year Scholarship	0.00	0	0.05	0	1	784
Freshman Credits	21.52	25	10.01	0	41	784
Freshman Renewal Credits	0.20	0	0.40	0	1	784
Freshman GPA	2.26	2	0.79	0	4	730
Freshman Renew GPA	0.33	0	0.47	0	1	784
High School GPA	3.01	3	0.46	1	4	784
Act Composite Scores	19.40	19	2.48	13	32	663
Sat Combined Scores	974.96	970	111.48	750	1260	121

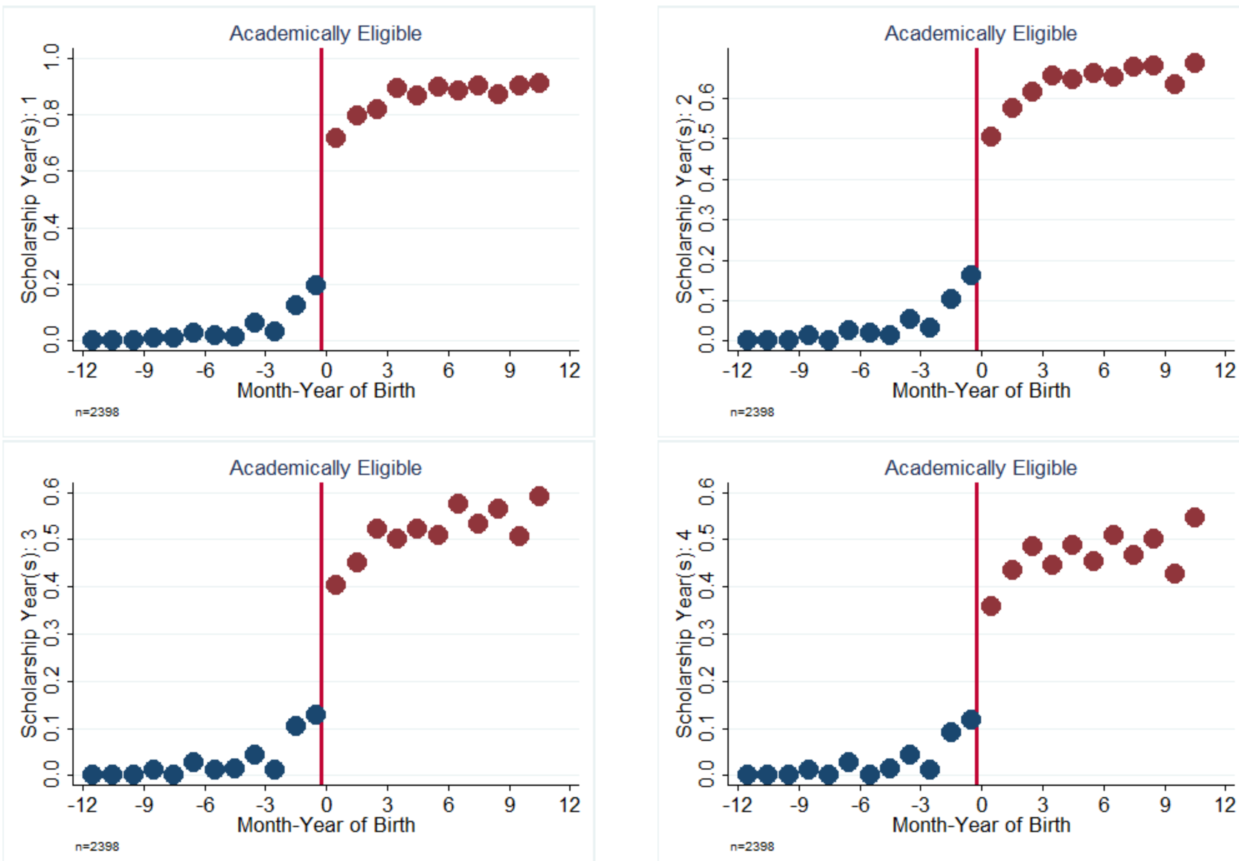
Note: Student-level observations. In-state students at WVU. First-time freshman only. Only cohorts 2001 and 2002. Freshman renewal GPA = 2.75. Freshman renewal credits = 30. Academic eligibility obstained by (HS GPA $>$ 3) AND (ACT Composite $>$ 20 OR SAT Composite $>$ 1000). Temporal eligibility if born on or after cutoff date. Panel A includes academically eligible and temporally eligible students. Panel B includes academically eligible, temporally ineligible. Panel C includes non-academically eligible, temporally eligible. Panel D includes academically ineligible and temporally ineligible students.

Figure A.2: Effect of Birth Month on Years of Scholarship



Note: Academic eligiblity determined by high school GPA and ACT score.

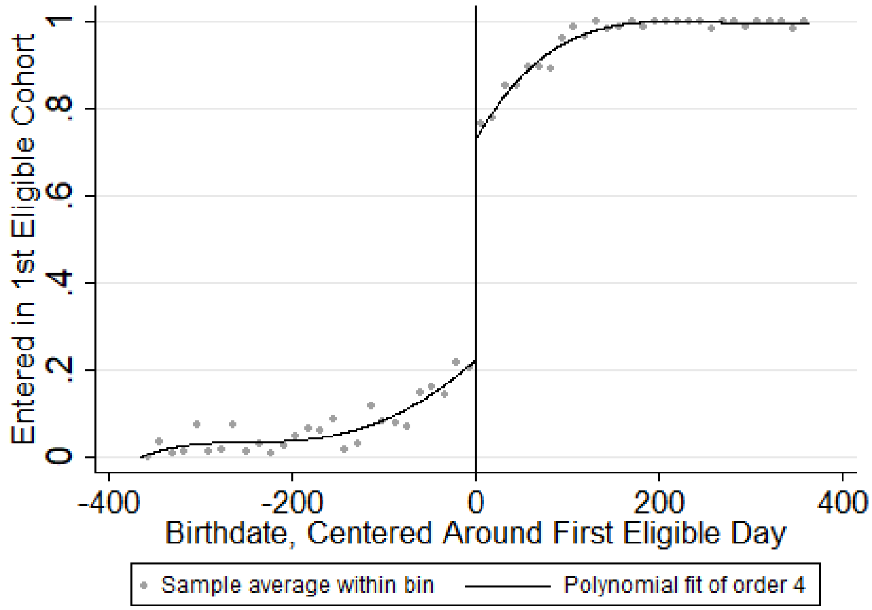
Figure A.3: Relationship between Birth Date and Years of Scholarship Received, Academically Eligible



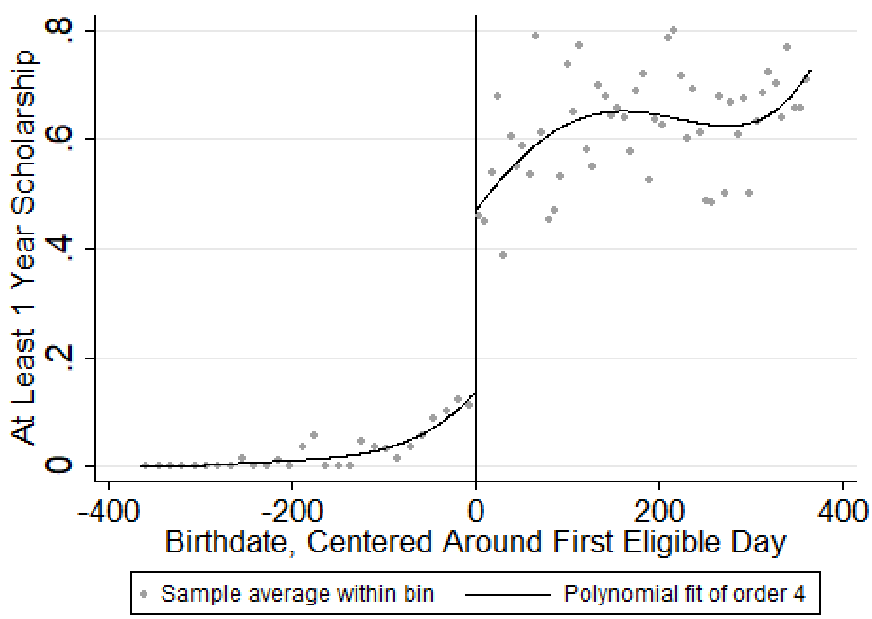
Note: Each successive graph uses a more restrictive criteria for the likelihood of scholarship. The y-axis measures likelihood of receiving scholarship for 1-4 years, 4 being the maximum number of years allowed by the policy.

Figure A.4: First Stage Discontinuities in Year Entered and Scholarship Receipt at Cutoff

Panel A: Effect of Birthdate, Relative to Cutoff, on Year of College Entry



Panel B: Effect of Birthdate, Relative to Cutoff, on Received Any Promise



Note: N = 3,669. Bin-length evenly spaced mimicking variance number of bins using spacings estimators. Plots are the conditional mean within bins. The black vertical line indicates the cutoff birth-month. The birth-month is centered at the cutoff month, meaning 0 is the first month of increased likelihood of treatment. In Panel A, the y-variable is 1/0 for being a first time freshman in the first year of the scholarship. In Panel B, the y-variable is mean of receiving at least 1 year of financial merit-based aid.

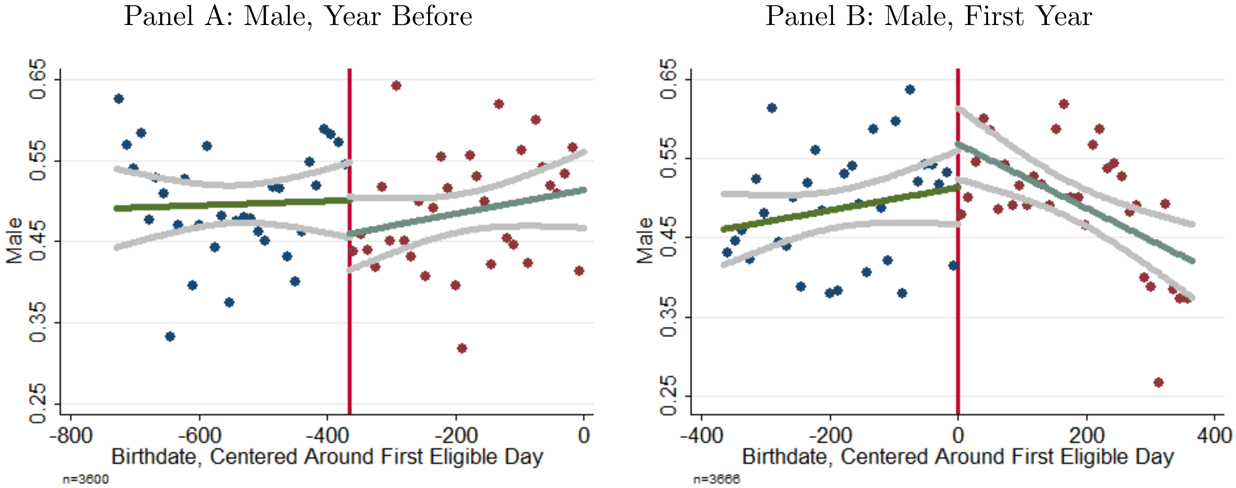
Figure A.5: Restricted McCrary Density Tests

Panel A: Effect of Birthdate, Relative to Cutoff, on Year of College Entry

Panel B: Effect of Birthdate, Relative to Cutoff, on Received Any Promise

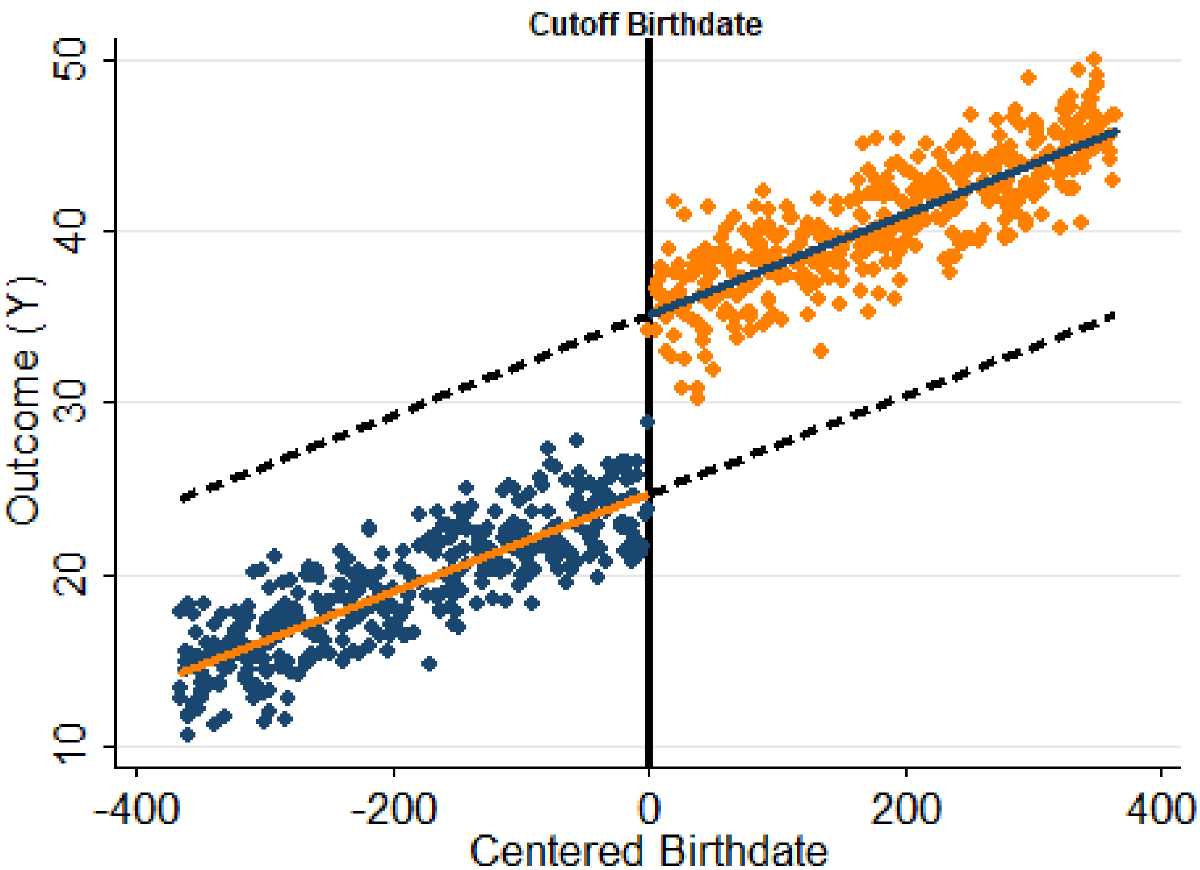
Note: Panels A & B estimated using rddensity in Stata. Optimal bandwidths are calculated according to Cattaneo et al. (2019b). Birthdate centered means 0 equals the relevant date for discontinuously changing scholarship receipt. The cutoff in Panel A is -365, because the cutoff is the same date - 1 year in the year before. The difference between this Figure and Figure 5, Panels E, F is this figure places restricts the CDF on both sides to an equal CDF and assumes existence of higher order derivatives. These restrictions increase the power of the test. Note that in Panel B, the estimated window size is longer than the frame of the data and should be interpreted cautiously.

Figure A.6: Tests of Gender Continuity at Date Cutoff



Note: Red line represents cutoff birthdate. Panels A, C, and E use the year prior to the scholarship being available cutoff, 365 days before 0. Panels B, D, and F use the year where the cutoff determines eligibility. Points are means conditional on point being within a bandwidth (bw) of h of a given birthdate. Plot A and B bw = 11.375. Plots C and D bw = 11.74. Plots E and F bw = 20.11. Bandwidths chosen as default histogram option. Specifically: number of bins chosen as: $min(\sqrt{N}, \frac{10ln(N)}{ln(10)})$, where N is number of observations.

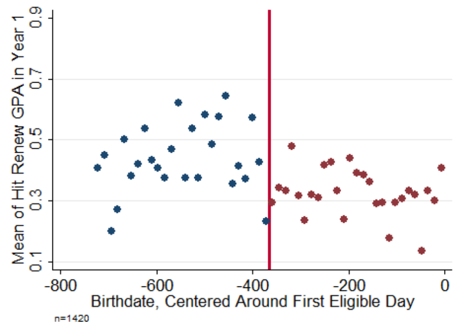
Figure A.7: Regression Discontinuity Identification Assumption: Local Smoothness of Potential Outcomes



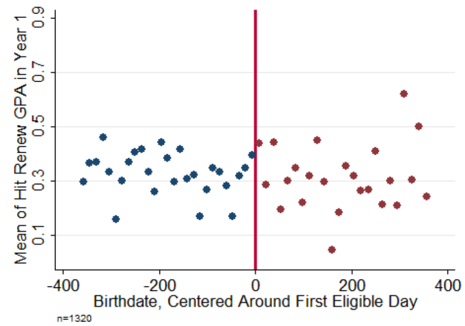
Note: Data is author-simulated. Picture depicts the assumption regression discontinuity, local smoothness, makes to estimate causal effects. Under local smoothness of potential outcomes, units just above and below make a valid counterfactual for one another.

Figure A.8: Reduced Form Effect of Birthdate on Likelihood of Exceeding Renewal Thresholds in Freshman Year, Including Before Scholarship Began, Academic Ineligible

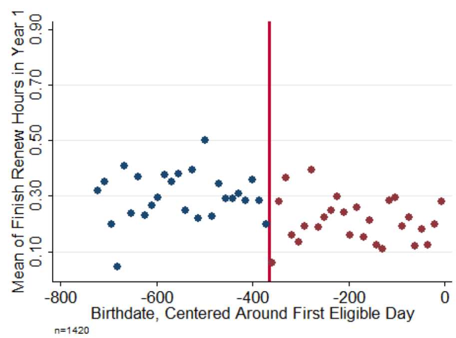
Panel A: Exceed 2.75 Freshman GPA, Pre-Scholarship, Academic Ineligible Students



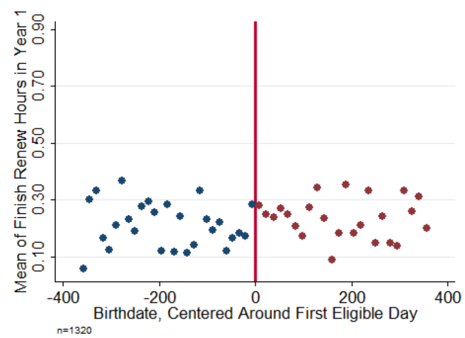
Panel B: Exceed 2.75 Freshman GPA, Scholarship Year, Academic Ineligible Students



Panel C: Exceed 30 Credits, Pre-Scholarship, Academic Ineligible Students



Panel D: Exceed 30 Credits, Scholarship Year, Academic Ineligible Students



Note: Red line represents cutoff birthdate. Panels A & C use the year prior to the scholarship being available cutoff, 365 days before 0. Panels B & D use the year where the cutoff determines scholarship eligibility. Points are means conditional on point being within a bandwidth (bw) of h of a given birthdate. Plot A bw = 14, 15.125. Plot B bw = 14. Plot C bw = 11.75. Plot D bw = 12.14. Plot D bw = 14. Plots A-D Bandwidths chosen as default histogram option. Specifically: number of bins chosen as: $min(\sqrt{N}, \frac{10ln(N)}{ln(10)})$, where N is number of observations.

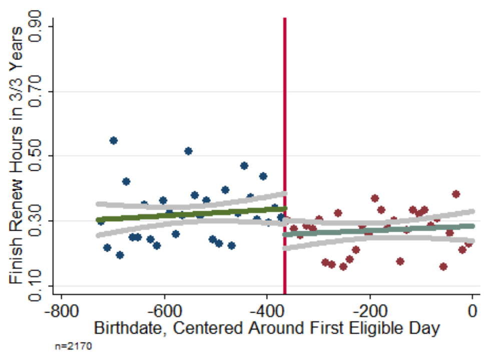
Table A.2: Local Linear Estimates of Scholarship on First Year Credits and GPA

Panel A: Dep. Var.: Continuous Credits			
	(1)	(2)	(3)
Scholarship	4.87*	5.61*	0.54
	(2.91)	(3.39)	(2.49)
Corrections	-	B-C, R	B-C, R
Sample	All	All	Ac. Elg.
Bandwidth	134	134	133
Eff. Obs. Left	658	658	393
Eff. Obs. Right	706	706	462
Obs-Left	1970	1970	1081
Obs-Right	1979	1979	1356
Panel B: Dep. Var.: Continuous GPA			
	(1)	(2)	(3)
Scholarship	0.34	0.39*	0.51**
	(0.21)	(0.23)	(0.22)
Corrections	-	B-C, R	B-C, R
Sample	All	All	Ac. Elg.
Bandwidth	193	193	146
Eff. Obs. Left	868	868	407
Eff. Obs. Right	978	978	500
Obs-Left	1865	1865	1055
Obs-Right	1925	1925	1335

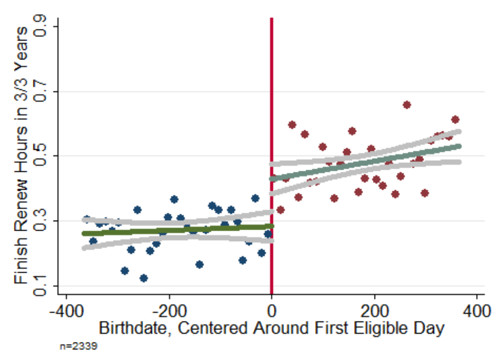
Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Estimates displayed are the second stage effect of scholarship from fuzzy regression discontinuity design using birthdate in relation to Kindergarten entry date as the cutoff/instrument. The estimate is the effect of receiving the Promise scholarship. In all panels, the first stage cutoff has a strong effect on scholarship receipt. In all results, bandwidths are calculated using [CITE FOR MSE-OPTIMAL BW'S], in which a common bandwidth is assumed on each side. In columns 1 and 2, the full sample is used. In column 3, only students who exceed the academic cutoffs are included in the estimation sample. In columns 2 and 3, the bias-correction and standard error adjustment in Calonico et al. (2017) are used, because of the additional bias from using MSE-optimal bandwidths. Panel A the dependent variable is a dummy for finishes 30 credits (the renewal requirement) freshman year. Panel B the dependent variable is the number of continuous credits completed freshman year. Panel C the dependent variable is a dummy for exceeds a 2.75 GPA (the renewal requirement) freshman year. In Panel D the dependent variable is continuous GPA.

Figure A.9: Reduced Form Effect of Birthdate on Likelihood of Exceeding Renewal Credits and/or GPA For 3 Years, Including Before Scholarship Began

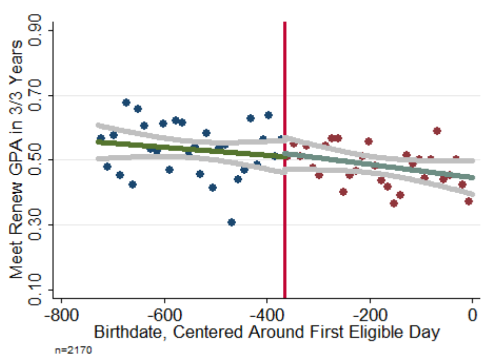
Panel A: Exceed Renew Credits For 3 Years, Pre-Scholarship, Academically Eligible



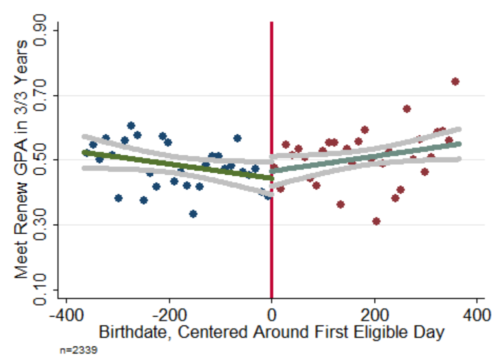
Panel B: Exceed Renew Credits For 3 Years, Scholarship Year, Academically Eligible



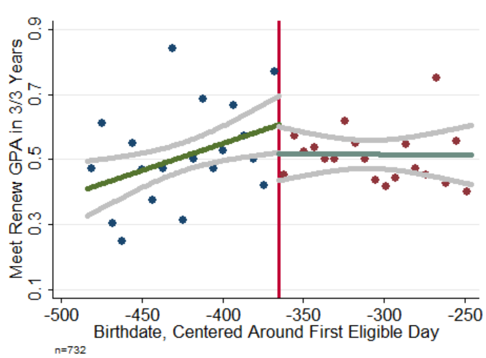
Panel C: Exceed Renew GPA For 3 Years, Pre-Scholarship, Academically Eligible



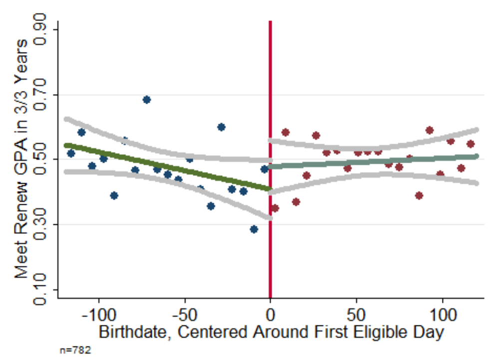
Panel D: Exceed Renew GPA For 3 Years, Scholarship Year, Academically Eligible



Panel E: Exceed Renew GPA For 3 Years, Scholarship Year, Academically Eligible



Panel F: Exceed Renew GPA For 3 Years, Scholarship Year, Academically Eligible



Note: Red line represents cutoff birthdate. Panels A, C, and E use the year prior to the scholarship being available cutoff, 365 days before 0. Panels B, D, and F use the year where the cutoff determines eligibility. Points are means conditional on point being within a bandwidth (bw) of h of a given birthdate. Plot A and B bw = 11.375. Plot C bw = 15.125. Plot D bw = 14. Plot E bw = 11.74. Plot F bw = 12.14. Bandwidths chosen as default histogram option. Specifically: number of bins chosen as: $min(\sqrt{N}, \frac{10ln(N)}{ln(10)})$, where N is number of observations.

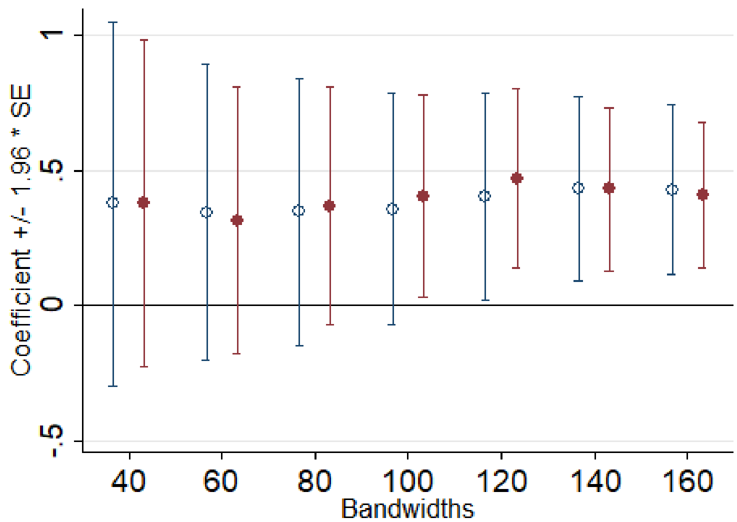
Table A.3: Local Linear Estimates of Scholarship on Meeting Third Year Credits and GPA Renewal Thresholds

Panel A: Dep. Var.: Renew Credits							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scholarship	0.39***	0.40***	0.34**	0.30*	0.25	0.36***	0.33**
	(0.12)	(0.14)	(0.16)	(0.16)	(0.18)	(0.13)	(0.16)
Corrections	-	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R
Sample	All	All	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.
Bandwidth	147	147	129	121	82	131	89
Eff. Obs. Left	697	697	379	358	225	390	246
Eff. Obs. Right	763	763	445	423	294	452	313
Obs-Left	1970	1970	1081	1081	1081	1081	1081
Obs-Right	1979	1979	1356	1356	1356	1356	1356
Acad. Controls				X	X	X	X
BW Criterion	MSE	MSE	MSE	MSE	CER	MSE	CER
Kernel	Tri	Tri	Tri	Tri	Tri	Uni	Uni
Panel B: Dep. Var.: Renew GPA							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scholarship	0.19	0.21	0.14	0.11	0.12	0.03	0.16
	(0.15)	(0.18)	(0.17)	(0.16)	(0.19)	(0.14)	(0.17)
Corrections	-	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R	B-C, R
Sample	All	All	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.	Ac. Elg.
Bandwidth	141	141	135	134	91	127	86
Eff. Obs. Left	680	680	395	395	253	379	238
Eff. Obs. Right	740	740	471	471	320	443	307
Obs-Left	1970	1970	1081	1081	1081	1081	1081
Obs-Right	1979	1979	1356	1356	1356	1356	1356
Acad. Controls				X	X	X	X
BW Criterion	MSE	MSE	MSE	MSE	CER	MSE	CER
Kernel	Tri	Tri	Tri	Tri	Tri	Uni	Uni

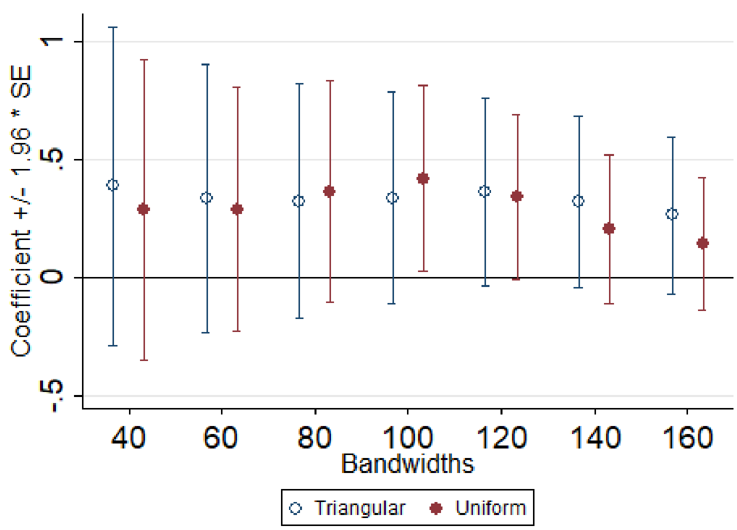
Note: * p<0.1, ** p<0.05, *** p<0.01. Estimates displayed are the second stage Wald estimator from a fuzzy regression discontinuity design using birthdate in relation to Kindergarten entry date as the cutoff/instrument. The estimate is the effect of receiving the Promise scholarship. In all panels, the first stage cutoff has a strong effect on scholarship receipt. In all results, bandwidths are calculated using [CITE FOR MSE-OPTIMAL BW'S], in which a common bandwidth is assumed on each side. In columns 1 and 2, the full sample is used. In column 3, only students who exceed the academic cutoffs are included in the estimation sample. In columns 2 and 3, the bias-correction and standard error adjustment in Calonico et al. (2017) are used, because of the additional bias from using MSE-optimal bandwidths. Panel A the dependent variable is a dummy for finishes 30 credits (the renewal requirement) freshman year. Panel B the dependent variable is the number of continuous credits completed freshman year. Panel C the dependent variable is a dummy for exceeds a 2.75 GPA (the renewal requirement) freshman year. In Panel D the dependent variable is continuous GPA.

Figure A.10: Coefficient Estimates from Difference in Discontinuity on 3 Year Outcomes

Panel A: Probability of Completing Renewal Credits



Panel B: Probability of Completing Renewal GPA



Note: Coefficients represent "conventional", not bias-corrected or robust for MSE-optimal, treatment estimates from difference-in-discontinuity estimates (Grembi et al., 2016). Diff-in-discontinuity is the RD estimate from the kindergarten entry date cutoff before it determined the scholarship receipt, but did determine relative age subtracted from the RD estimate in where it did determine scholarship receipt. This approach differences out the relative age effect under fairly mild assumptions. The different bandwidths indicate the number of days used from the respective cutoffs. Triangular and uniforms kernels used. Panel B legend valid for both. All regressions subsample to academically eligible students meaning [(ACT 20.5 OR SAT $\geq$ 1000) AND HS GPA 3.0] and include controls for HS GPA and ACT composite score (w/ SAT converted to ACT). Panel A dependent variable is exceed renewal credits in freshman year. Panel B dependent variable is renewal GPA in freshman year. The horizontal dashed lines represent the Scott-Clayton (2011) point estimates which correspond approximately to 730 day bandwidths, since Scott-Clayton (2011) uses 2 cohorts before and after.