MERIT-AID RENEWAL BEYOND A CUTOFF: EVIDENCE FROM PROMISING BIRTHDATES

Benjamin Blemings*

This Version: Thursday 18$^{\text{th}}$ August, 2022, Updated Version: Link

Abstract

Financial aid renewal requirements are ubiquitous and give students a concrete goal which is a promising strategy for students transitioning beyond high school. Yet, our understanding of merit-based aid's renewal requirements could be deepened with alternative identification strategies that address selection into programs with publicly known academic eligibility cutoffs. This paper estimates how receiving West Virginia's merit-based Promise scholarship affects students' likelihood of meeting academic renewal requirements using an unmanipulated birthdate cutoff that produces an internally valid fuzzy regression discontinuity. In contrast to known entry test requirements, density tests do not find significant enrollment differences at the birth cutoff. Scholarship-receiving students are 24 percentage points (pp) more likely to exceed the credits requirement and 31 pp more likely to exceed the GPA requirement, on average. There are no significant effects of crossing this birth cutoff, when crossing it does not also result in receiving the scholarship which means the effects are driven by student's responses to the requirements and not relative age effects. These estimates are externally valid for complier students away from the cutoff, but not to GPA targets of non-compliers due to never-takers out-performing untreated compliers. These estimates advance knowledge on how responsive students are to renewal goals, particularly their relative responses to credits or GPA targets and how externally valid those estimates are, informing design of financial aid renewal incentives.[1]

Keywords: Merit-Aid, Renewal Requirements, West Virginia Promise Scholarship, External Validity, Regression Discontinuity, Difference in Discontinuity

JEL Classification: I22, H75

*Postdoctoral Researcher, Dyson School of Applied Economics and Management, Cornell University, E-mail: btb77@cornell.edu

[1]I thank Adam Nowak, Dan Grossman, Scott Cunningham, Bryan McCannon, Brad Humphreys, Bill Even, Daniel Hungerman, and Clinton Neill for many helpful comments and suggestions. I thank seminar participants at Eastern Economic Association 2020, Southern Economic Association 2020, Southern Utah University, Kansas University, Wayne State University, West Virginia University, and the Association for Education Finance Policy 2021. I thank Liz Reynolds for graciously providing detailed administrative data. The data can be applied for from the West Virginia Higher Education Policy Commission.

I Introduction

College student financial aid packages often feature academic renewal requirements such as minimum credit or GPA thresholds that must be met for aid to continue being offered in subsequent years. Merit-aid programs in Georgia, Tennessee, Florida, and West Virginia all feature these renewal requirements. Following their experience with statewide merit-aid renewal requirements, West Virginia University also added these requirements to institutional scholarships.[2] Renewal requirements are virtually ubiquitous, with 96% of 52 land grant universities currently offering at least 1 institutional scholarship with at least 1 renewal requirement. Goal setting is a critical aspect of financial aid packages, since students who set goals are less likely to drop out of college (Beattie, Laliberté, Michaud-Leclerc, & Oreopoulos, 2019).

Renewal requirements are widespread and existing evidence suggests that goals are promising for increasing academic performance (C. M. Cornwell, Lee, & Mustard, 2005; Scott-Clayton, 2011), but the literature does not provide an answer as to whether renewal requirements are effective, whether they differ across types of targets, and whether the effects are externally valid. This gap in the literature is surprising given the summary in Dynarski and Scott-Clayton (2013) which finds that requirements are likely effective and that program design should be an area for research. One important obstacle to estimating how students respond to academic performance targets for merit-aid renewal is selection bias that arises from students sorting into treatment due to known academic eligibility cutoffs.[3] Students strategically manipulate academic scores into being just high enough to receive the merit aid scholarships by retaking entrance exams, such as the ACT, in Tennessee, Florida, and West Virginia (Bruce & Carruthers, 2014; Scott-Clayton & Zafar, 2019; Zhang, Hu, Sun, & Pu, 2016).

This paper quantifies how strongly students respond to academic targets, minimum credits completed and minimum GPA during freshman year, that must be met to renew a scholarship during subsequent years with previously unused variation in merit-based scholarship receipt. Variation derives from future college students who were born in 1983 on opposite sides of the cutoff date

[2] West Virginia University's policy change was discovered via conversations with financial aid administrators.

[3] Using entrance test scores (e.g., ACT and SAT) and/or high school GPA cutoffs is similar to test cutoffs used in the original RD design (T. D. Cook, 2008; Thistlethwaite & Campbell, 1960). However, one difference between contemporary merit-aid research and Thistlethwaite and Campbell (1960) is that academic cutoffs for many merit-aid scholarships are known.

for beginning kindergarten in 1988. The college scholarship was not announced until 1999, and was first available to the cohort that began college during 2002. By being born on opposite sides of the 1983 birthdate cutoff, college students from a major university in West Virginia entered in different kindergarten cohorts, which affected their college cohort and eligibility to receive the merit-based Promise Scholarship in 2002. Students were born 16 years before the scholarship was announced, and 19 years before it began making it impossible for future students, or their parents, to manipulate birthdates to increase the likelihood of becoming eligible for the scholarship.

This design, using cohorts of college entry as an instrument of scholarship receipt, is similar to that used in extant research (Scott-Clayton, 2011; Scott-Clayton & Zafar, 2019). However, this approach includes the additional insight that cohorts of college entry depend on kindergarten cohorts which depend on birthdate due to compulsory kindergarten entry laws. This insight is consequential to estimates of how strongly students respond to scholarship renewal targets on average and enables the investigation of the estimates' external validity.

First, the key parameters which determine whether the design provides sufficient variation are modeled and the design is shown to be internally valid. The fuzzy regression discontinuity design has a strong first stage, passes density tests around the birth cutoff, and returns reasonable covariate balance at the cutoff date. No statistical differences are evident in the densities on either side of the cutoff, suggesting an absence of birthdate manipulation, and that the identifying assumption is met (Cattaneo, Jansson, & Ma, 2019; Hahn, Todd, & Van der Klaauw, 2001; McCrary, 2008).[4]

Although the difference in densities is insignificant at the cutoff to a degree that supports identification, the density of students does remain elevated beyond the cutoff. This provides a nuanced angle to examine for which students the aid programs influence enrollment. It has been established that merit-aid increases in-state enrollment (C. Cornwell, Mustard, & Sridhar, 2006; Londono-Velez, Rodriguez, & Sánchez, 2020; Scott-Clayton, 2011; Sjoquist & Winters, 2013, 2014; Winters, 2020) and that age at school entry matters for college admission (Matta, Ribas, Sampaio, & Sampaio, 2016), yet this paper suggests that aid and age at school entry together do not to create larger effects on college enrollment.

[4]The inability to manipulate beyond the cutoff to receive treatment is consistent with the trends towards using such cutoffs in academic settings (Bettinger, Gurantz, Kawano, Sacerdote, & Stevens, 2019; Goodman, 2008; Goodman, Gurantz, & Smith, 2020) and is an alternative discontinuity in birthdate compared to tax deadlines used by Denning (2019).

This design is not new (Canaan, 2020; Takaku & Yokoyama, 2021), but this paper offers several firsts. This paper is first to apply the approach to address issues with evaluation of merit-aid scholarships and use exact birthdate. Use of exact birthdate more closely approximates a continuous running variable, which is advantageous because it is addresses issues with RD inference with a discrete running variable and enables econometric tests which assess whether the estimates are externally valid (Bertanha & Imbens, 2019; Dong & Lewbel, 2015; Kolesár & Rothe, 2018; Lee & Card, 2008). Extant research uses similar designs that affect kindergarteners (Takaku & Yokoyama, 2021) and sixth graders (Canaan, 2020), but this paper demonstrates that the design can be applied to policies that occur much later after a school entry birth cutoff. Specifically, it demonstrates that leveraging birth cutoffs due to compulsory entry laws work for policies that begin two times longer— 13$^{\text{th}}$ grade. To clarify the extension to a policy beginning longer after birth, a mathematical outline formalizes how the design provides quasi-random variation in education cohort-based treatments. The outline identifies the parameters which are relevant for enabling the design in other contexts and helps define the compliers.

Second, students respond strongly to the scholarship's academic renewal targets and this does not differ by type- credits completed (30) or GPA (2.75). Students who receive the scholarship are 24 percentage points (pp) more likely to reach minimum renewal requirements for both requirements. Relative age, unlike the scholarship, does not increase the likelihood that students reach these thresholds. No discontinuous increase is evident regarding the likelihood of exceeding these academic renewal thresholds at the birthdate cutoff during the year before when scholarship eligibility remains unchanged by exceeding the cutoff. This contrasts research that finds either of the renewal requirements have negative effects on hours attempted (C. M. Cornwell et al., 2005) or that students respond much more strongly to credits than hours requirements (Scott-Clayton, 2011). While the negative effects from C. M. Cornwell et al. (2005) are due to having unlimited semesters to use the financial aid from the scholarship, the discrepancy between targets in Scott-Clayton (2011) has remained unresolved. In this paper, the only time this discrepancy is observed is when students farther from the cutoff are included and the uniform kernel is used. This suggests that the discrepancy is an artifact of the weighting and comparison group used in the estimator by Scott-Clayton (2011) which is an important insight for designing renewal requirements of aid.

These estimates add to the debate surrounding which combination of policies are most effective at inducing greater post-secondary academic success (Dynarski & Scott-Clayton, 2013; Gneezy, Meier, & Rey-Biel, 2011; Lavecchia, Liu, & Oreopoulos, 2016). These results complement several extant studies that find positive effects of aid on academic achievement during students' early college years. For example, Leuven, Oosterbeek, and Van der Klaauw (2010) and Barrow, Richburg-Hayes, Rouse, and Brock (2014) find that financial incentives increase credits for early career post-secondary students. These estimates extend this finding to an important 4-year public university in West Virginia. Furthermore, J. Angrist, Autor, and Pallais (2021) finds evidence that the effect of merit aid on degree completion is mediated by credit hours during the first year of college and this study examines a target for credit completion. Merit-aid programs are a two-part intervention, providing goals and financial aid. Experimental evidence suggests that task-based rather than performance based goals are more effective among college students (Clark, Gill, Prowse, & Rush, 2020). Yet, the results of this study suggest that performance based goals can be effective when paired with financial incentives.

Third, the estimates in Scott-Clayton (2011) are valid for a subset of complier students and in this study the estimates are valid for complier students at the cutoff. While this paper finds that the average responses to first year requirements are the same, the credits and GPA targets differ in their robustness to extending the scholarship's effects to subsets of students that are not compliers at the cutoff. Specifically, the credits and GPA targets estimates are both valid for compliers away from the cutoff (Dong & Lewbel, 2015); however, only the credits are externally valid for always-takers and never-takers beyond the cutoff (Bertanha & Imbens, 2019).[5] For academic designs, even continuous, unknown ones such as those in Massachusetts for the Adams Scholarship, the estimated causal effects may not extend to those beyond the academic cutoff; this is the case for the effect of the Adams Scholarship on college choice (Dong & Lewbel, 2015; Goodman, 2008). The estimates from the birthdate cutoff generalize to compliers away from the cutoff which is a benefit to estimating effects of merit-aid by using this design. Furthermore, these results caution that GPA targets may be ineffective for those who do not complete 1 grade per year during kindergarten

[5] Under an additional assumption that Dong and Lewbel (2015) terms "local policy invariance," the birth cutoff could be moved without changing the effects. This is important given the variation in birth cutoffs across states.

through 12th grade and enter in the kindergarten cohort assigned to them by their birthdate relative to compulsory entry cutoff dates, but the effectiveness of credit targets are valid beyond this subset.

Section 2 outlines predictions mathematically regarding birthdates that affect scholarship receipt, the complier group, and verifies that college cohort entrance year depends on birthdate. Section 3 details the data, and Section 4 describes the fuzzy RD method, the difference in discontinuity method, and how the target estimand differs from others in merit-aid literature, and why that matters. Section 5 discussed the internal validity of the research design in this context. Section 6 reports results of whether receiving the scholarship affects the likelihood of meeting academic renewal thresholds. Section 7 investigates the external validity of the estimates, and Section 8 concludes.

II Kindergarten and College Cohort and College Scholarship

This section outlines discontinuous rules and related factors that lead to discontinuous college scholarship receipt at the kindergarten birthdate cutoff. It builds on an instrument, college cohort entry year, used in Scott-Clayton and Zafar (2019) by adding and testing the design insight that college cohort entrance year depends on birthdate due to kindergarten cohort entrance birthdate cutoffs that compulsory entry laws create. If students finish one grade per year, their college cohort is an additive function (with a constant) of their kindergarten cohort. Two other papers (Canaan, 2020; Takaku & Yokoyama, 2021) use this logic to investigate treatments that begin with a cohort, but the current paper is first to outline mathematically the complier group necessary for a comparison of other LATE estimands in merit-aid literature.

Equations 1, 2, and 3 outline the system of equations that leads to the variation in scholarship receipt used in this paper:

$$Promise\ Scholarship = f(College\ Cohort), \tag{1}$$

$$College\ Cohort = g(Kindergarten\ Cohort), \tag{2}$$

$$Kindergarten\ Cohort = h(Birthdate\ Relative\ to\ Cutoff). \tag{3}$$

Equations 1 and 3 are known, discontinuous rules regarding scholarship eligibility and kindergarten cohort entrance, respectively. Equation 2 is a grade progression function. By substitution:

$$Promise\ Scholarship = f(g(h(Birthdate\ Relative\ to\ Kindergarten\ Cutoff))). \tag{4}$$

Each equation is discussed below.

Equation 1 represents the rule that to receive the scholarship, students must enter college after the scholarship program began. Equation 1B expands Equation 1 into its piece-wise form:

$$P(Promise\ Scholarship = 1) = \begin{cases} q' + k' & if\ College\ Entry\ Cohort \geq 2002 \\ q' & if\ College\ Entry\ Cohort < 2002. \end{cases} \tag{1B}$$

'The constant *k'* captures how detrimental a college entry cohort is for college scholarship receipt. Existing research relies on $k' > 0$ and shows this is indeed true, meaning that entry cohort is important to scholarship receipt (Scott-Clayton, 2011; Scott-Clayton & Zafar, 2019).

Equation 3 represents the compulsory kindergarten entry birthdate cutoff rule, a rule set by law that was in place in 1988 in West Virginia. Students who have not turned 5 by the cutoff birthdate—September 1, 1988—are required to enter kindergarten during 1989, which means that students born after August 30, 1988 enter kindergarten in 1989. Equation 3B expands Equation 3 into its piece-wise form for the 1989 kindergarten cohort:

$$P(Kindergarten\ Cohort = 1989) = \begin{cases} p' + c' & if\ BD \geq 9/1/1983 \\ p' & if\ BD < 9/1/1983, \end{cases} \tag{3B}$$

where BD is the birthdate. This paper assesses students with birthdates close to the cutoff, so the following discussion concerns students whose birthdates are near the cutoff, $(9/1/1983 - \epsilon, 9/1/1983 + \epsilon)$, but this notation is suppressed throughout. A positive c' means that students are discontinuously more likely to enter kindergarten during 1989 if born after the cutoff date. If

$c' \leq 0$, the cutoff laws are not enforced.[6] Since kindergarten entry cohort is not observed, it is impossible to estimate c' (i.e. the degree of imperfect enforcement) using data in this paper.

Equation 2 represents student grade progression. For the kindergarten birthdate cutoff, to predict college scholarship receipt systematically, students must finish 1 grade per year (i.e. kindergarten through 12[th] grade) and then enter college. Thus, Equation 2 takes a specific functional form among compliers,

$$College\ Cohort = Kindergarten\ Cohort + 13.^{7} \tag{2B}$$

This is reasonable 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.

The proportion of students that follows Equation 2B can be verified with data. Under one-grade-per-year progression, Equation 3B implies:

$$P(College\ Cohort = 2002) = \begin{cases} p + c & if\ BD \geq 9/1/1983 \\ p & if\ BD < 9/1/1983. \end{cases} \tag{5}$$

If $c > 0$, then Equation 2B describes a positive share of the sample. c' captures only how strongly the kindergarten birthdate cutoff is enforced, but c also includes grade progression and immigration. Some ways that students might not follow Equation 2B is if students skip a grade, are held back, or redshirt at least one grade. Immigration also influences c because a student could be following the grade progression described by Equation 2B but was born in a different state with different kindergarten entrance rules.[8]

All elements of Equation 5, individual's college entry cohort and exact birthdate, are observed in the data, which allows non-parametric estimation of both p and c, presented in Figure I.[9] Figure IB shows that in comparison to students born just before the cutoff, being born after the cutoff

[6] As discussed in Section B, the cutoff is enforced imperfectly and costly to circumvent.

[7] There are 13 grades between kindergarten and college. For a program beginning with a different year, 13 would be a different number. For example, if the program began in 6[th] grade, like in Canaan (2020), then 13 would be 6.

[8] The average annual growth rate of the population of West Virginia from 1950 to 2000 was -0.12%. Students must also complete high school graduation requirements at a West Virginia high school, with at least half of the credits required for graduation obtained in-state (West Virginia Higher Education Policy Commission, 2012).

[9] More precisely, it allows estimation of $\lim_{\mathrm{BD}\to 9/1/1983^+} E[College\ Cohort\ Year = 2002 \mid Birthdate]\ (p + c)$ and $\lim_{\mathrm{BD}\to 9/1/1983^-} E[College\ Cohort\ Year = 2002 \mid Birthdate]\ (p)$.

during 1983 leads to a c_{2002} = 53.3 pp increase in the likelihood of entering college during 2002. The left side of the cutoff (p_{2002}) is not zero, which might be due to grade entrance exemptions, irregular grade progression, or immigration to West Virginia. It is impossible to disentangle the relative importance of these factors without additional data.

It is possible that the scholarship changed the likelihood that people would enter the college cohort predicted by their kindergarten cohort. This could happen if individuals delayed college by a year to become eligible for the scholarship. However, this would imply that the financial aid was more valuable than an additional year of post-graduation labor income (C. M. Cornwell et al., 2005, pg. 897). To investigate whether this delay occurs, Figure IA estimates c_{2001}, when crossing the birthdate cutoff does not affect scholarship receipt. If $c_{2002} \neq c_{2001}$, this implies that students delay college entry or in someway manipulate their cohort due to the scholarship. The estimate of c_{2001} is 52.6 pp, which $\approx c_{2002}$, suggesting that the scholarship does not meaningfully alter the relationship between kindergarten and college cohorts.

Most importantly, $c > 0$, which means that the proposed mechanism, college cohort differences, for why college scholarship receipt might differ discontinuously across the kindergarten entry birth cutoff has strong empirical support. Since beginning in the first eligible year matters for scholarship receipt (Scott-Clayton, 2011; Scott-Clayton & Zafar, 2019), Equation 5 implies

$$P(Promise\ Scholarship = 1) = \begin{cases} q + k & if\ BD \geq 9/1/1983 \\ q & if\ BD < 9/1/1983. \end{cases} \tag{6}$$

If beginning in the first year of the program is a sufficiently important requirement, then it is expected that $k > 0$. Without additional restrictions, $c \neq k$ because academic eligibility is also a criterion for eligibility. In Section IV, Equation 7 shows how k is estimated, in Section IV.1.2, k's role in identification is discussed more, and in Section V.1, k is estimated.

III Data

The research question is whether receiving a merit-aid scholarship affects the likelihood that students meet the academic renewal requirements for the scholarship. Thus, individual-level data on scholarship receipt and collegiate academic outcomes are the minimum data requirements. Extant research also uses data on cohorts of college entrance as an instrument of scholarship receipt (Scott-Clayton, 2011; Scott-Clayton & Zafar, 2019). The analysis uses these variables and adds additional data that have not yet been used to evaluate merit-aid programs. The new data are individuals' exact birthdates, combined with historical information about compulsory kindergarten entrance laws, which create lasting discontinuities in school cohorts by birthdate in West Virginia.

Data were collected at the student level among 2000 to 2002 cohorts at a large, public university in West Virginia. The sample consists of only in-state students who are first-time freshman which results 5,676 observations. Exact birthdate, gender, race, high school GPA, entry test scores, and first-year academic outcomes are observed. Table A.I reports summary statistics.

First-Year Academic Renewal Targets Outcomes The dependent variables are binary; they equal 1 if a student completed the minimum required credits or if he/she exceeded the GPA to renew the Promise scholarship during the first year, respectively. Students must finish at least 30 credits during each year—in Fall, Spring, and Summer—to renew the scholarship. They must obtain at least a 2.75 GPA during their freshman year, which increases to 3.0 during subsequent years. There are results for continuous outcomes, but it is appropriate to focus on this dichotimization because cutoff values are meaningful since they determine scholarship renewal and are relevant for policy and behavior. Figure A.II shows how these outcomes differ by cohorts and academic eligibility.

III.1 Scholarship Eligibility

The data include information on 3 requirements for scholarship eligibility:

1. Earn at least a 3.0 GPA in high school,
2. ACT Composite score $\geq$ 21 OR SAT Combined Score $\geq$ 1,000,
3. The student must be a college freshman in the 2002 cohort, or after.

Since the program is merit-based, 1 and 2 relate to academic proficiency. To simplify, requirements 1 and 2 are combined into a single variable—academic eligibility—defined as:

$$Acad.\ Eligible_i = 1[(HS\ GPA \geq 3) \cap (ACT\ Comp \geq 21 \cup SAT\ Comb \geq 1,000)],$$

where *HS GPA* represents high school GPA, *ACT Comp* is the ACT composite score, and *SAT Comb* is the SAT combined score.[10] Regarding 3, West Virginia is different from states that allowed merit-aid to be received retroactively for sophomores who entered earlier than the beginning year, such as Tennessee. This makes West Virginia an ideal context for this design.

Exact Birthdate and Cutoffs Exact birthdate is observed. According to standard practice, it is centered so that 9/1/1983 is day zero, since it is the first birthdate expected to be discontinuously more likely to receive the scholarship. The cutoff at day zero is called the post-scholarship or 1983 cutoff because being born on or after this date makes students more likely to receive the scholarship, as detailed in Section II. Students born on day -365 are more likely to be the oldest in their cohort, but are not more likely to receive the scholarship. Since it is a cutoff that does not determine scholarship receipt, it is called the pre-scholarship or 1982 cutoff.

III.1.1 Scholastic and Birthdate Cutoff Comparisons

Figure I shows that the birthdate cutoff is a viable candidate research design, but it might have less predictive power for scholarship receipt than academic outcomes do. Figure II shows scholarship receipt scatter plots for a 10% random sample by both academic and birthdate cutoffs. Figure IIA compares the ACT cutoff to the birthdate cutoff. There are about equal numbers of individuals on either side of the ACT cutoff prior to the birthdate cutoff because there is no incentive to be above the ACT cutoff before that score is used to determine scholarship eligibility. After the birthdate cutoff, fewer individuals lie just below the ACT cutoff, which accords with findings in Scott-Clayton (2011) and Scott-Clayton and Zafar (2019). A limited number of observations lies below the cutoff, and many above the cutoff, which is not ideal for RD because it suggests that students manipulate scores into treatment. To the left of the birth cutoff for individuals above the ACT score, most do not receive the scholarship, and that changes on the right side of the birth cutoff above the ACT score, meaning that the birthdate cutoff is likely to result in a strong first stage.

Figure IIB compares the HS GPA cutoff to the birthdate cutoff. Very few observations lie near the GPA cutoff before or after the birthdate cutoff, meaning that it is unlikely that the GPA cutoffs

[10] Academic eligibility thresholds change over time, but this is irrelevant here since only 1 eligible cohort is used.

are appropriate for use in an RD design in this context. Many students have near perfect or perfect high school GPAs on either side of the date cutoff, demonstrating one advantage of using a non-academic cutoff; it is possible to compare students of similar academic ability before and after the birth cutoff.

IV Method

Fuzzy regression discontinuity is used to estimate effects of the scholarship by comparing students born days apart who have discontinuously different rates of scholarship receipt. Scholarship take-up as a function of birthdate and the cutoff is estimated,

$$S_i = \lambda + \psi After83_i + f(Birthdate) + \epsilon_i \ \ \forall \mid Birthdate_i \mid < h, \tag{7}$$

in which *i* stands for individual. Scholarship receipt, S, is a dummy variable that equals 1 if a student received the scholarship. Being born after the post-scholarship birthdate cutoff is represented by $After83$, which is a binary variable that equals 1 if a student was born after August 30, 1983.

The relationship between the birthdate cutoff and academic targets is estimated as,

$$y_i = \alpha + \beta After83_i + f(Birthdate_i) + e_i, \ \ \forall \mid Birthdate_i \mid < h. \tag{8}$$

In both regressions, the effect of birthdate on scholarship receipt and outcomes, $f(Birthdate)$, is a local regression that is fit separately on either side of the cutoff. The regressions are fit to a subset of birthdates closest to the cutoff. The bandwidth is h which is the number of days away from the birthdate cutoff, that is used for estimation.[11]

IV.1 Target Estimand

Since the merit-aid scholarship was not assigned randomly, estimating an average treatment effect (ATE) is impossible without assuming (conditional) ignorability of treatment assignment, a strong and likely unrealistic assumption. Due to the implausibility of the assumptions that would enable

[11] Typically, h is chosen using MSE-optimal bandwidths, and noted if otherwise.

recovering the ATE with available data, other estimands are the focus of this paper. In Equation 8, β represents the estimator of the intention-to-treat (ITT) effect of the scholarship under local smoothness of potential outcomes across the birthdate cutoff (Hahn et al., 2001).

The local average treatment effect (LATE) is the target estimand, estimated as $\tau = \frac{\beta}{\psi}$, which represents the ITT effect, scaled by how strongly the birth cutoff influences scholarship receipt. LATE is local to students near the birth cutoff, and is valid only among compliers. In addition to receiving the scholarship by being on the right of the birth cutoff, compliers in this case are also those who enter the kindergarten cohort assigned to them by their birthdate, relative to the cutoff entry date, and finish 1 grade per year and then enter college.

IV.1.1 Other Merit-Aid Estimands

The target estimand differs from others in merit-aid literature. Many estimates come from regression discontinuity designs based on academic score cutoffs (Goodman, 2008; Scott-Clayton, 2011; Scott-Clayton & Zafar, 2019). Due to manipulation of best entrance test score when academic cutoffs are known, estimates in Bruce and Carruthers (2014) use only first test scores, which, by definition, are not manipulated.

Target estimands in extant merit-aid research might not be externally valid to non-compliers or those away from the cutoff. Ignoring the issue of manipulation into treatment due to known academic cutoffs, even unknown academic cutoff design estimates are internally valid only among complier students who are academically marginal (Dong & Lewbel, 2015; Goodman, 2008; Hahn et al., 2001). Using only the first academic entrance test to deal with manipulation in known academic cutoff designs, interpretation of estimates in Bruce and Carruthers (2014) are that they are internally valid only among compliers.[12]

IV.1.2 Identification

Interpreting τ as a LATE requires the same local smoothness assumption required to interpret β as the ITT effect, and three additional assumptions. First, there must be a first stage; students to the

[12] Compliers in Bruce and Carruthers (2014) are those who did not do sufficiently well on the first test and, for whatever reason, decided not to retake the test for an opportunity to increase the score beyond the academic threshold, which would have resulted in a large amount of financial aid.

right of the birth cutoff must be discontinuously more likely to receive the scholarship. Second, there needs to be restrictions on how exceeding the birth cutoff affects the likelihood of scholarship receipt. Third, the stable unit treatment value assumption (SUTVA) is required.

Before reporting diagnostics for the first stage and local smoothness, monotonicity and SUTVA are discussed briefly. Monotonicity means that exceeding the birth cutoff does not reduce the likelihood of receiving the scholarship, or vice versa.[13] SUTVA requires that those who receive the scholarship do not affect the potential outcomes of those who do not. It also assumes that everyone receives an equal amount of the scholarship, which is certainly met for examining the influence of the scholarship on first-year outcomes.

IV.2 Alternative Estimation Strategy: Difference in Discontinuity

An additional strategy that could be used to estimate effects of the scholarship is difference in discontinuity (Canaan, 2020; Grembi, Nannicini, & Troiano, 2016), defined by Grembi et al. (2016) as,

$$T^{DD} \equiv (Y^{+} - Y^{-}) - (\tilde{Y}^{+} - \tilde{Y}^{-}). \tag{9}$$

The second term, $(\tilde{Y}^{+} - \tilde{Y}^{-})$, is from a birth cutoff in which the probability of receiving the scholarship does not change discontinuously across the cutoff. In this application, the second term can be estimated from the birth cutoff on September 1, 1987 because students are sorted into the 1987 or 1988 kindergarten cohorts, which subsequently leads them to enter college in 2000 or 2001. The likelihood of scholarship receipt does not change discontinuously from the 2000 to 2001 cohort because it is unavailable until the cohort entering in 2002. This allows the effect of exceeding the birth cutoff, when it does not affect scholarship receipt, to be estimated and differenced out of the effect of the scholarship.

[13] Section II defines conditions for compliers. One, they enter the kindergarten cohort assigned by birthdate in relation to the kindergarten cutoff date. Two, they complete one grade per year in grade school, and then enter college.

IV.2.1 Advantages and Disadvantages

There are advantages and disadvantages to using the difference in discontinuity estimator in this context. First, local smoothness in fuzzy RD implies that the second term in Equation 9, $(\tilde{Y}^+ - \tilde{Y}^-)$, equals zero. When $(\tilde{Y}^+ - \tilde{Y}^-) = 0$, there is no net effect of relative age on realized outcomes, when the likelihood of receiving the scholarship is not discontinuous across the cutoff. This implies that relative age does not cause potential outcomes at either cutoff to be discontinuous.

Second, even if relative age affects the likelihood of reaching scholarship renewal thresholds (i.e. $(\tilde{Y}^+ - \tilde{Y}^-) \neq 0$), and is not due to receiving the scholarship, the difference in discontinuity estimator removes this bias and restores identification under additional assumptions. There are two disadvantages to difference in discontinuity in this context. First, it requires additional assumptions, and second, two differences must be estimated, which adds uncertainty and reduces precision without a sufficiently large number of observations.

Preferred Estimator Assessing whether τ (from Section IV.1) or T^{DD} is the most appropriate estimator depends on the assumptions required to interpret these estimators as causal. In addition to all of the assumptions above, interpreting T^{DD} as a heterogeneous LATE requires additional assumptions that are local (to the birth cutoff) parallel trends and additive separability of the effects of relative age and the scholarship (Grembi et al., 2016).[14]

V Internal Validity

This section reports diagnostics that assess the validity of assumptions necessary for τ to be interpreted as a LATE. These diagnostics examine how strongly the cutoff influences scholarship receipt (the first stage), and how likely it is that potential outcomes are smooth across the birth cutoff using density and covariate balance tests. Another test shows that realized outcomes are

[14] The local parallel trends assumption cannot be assessed with the data currently available because having 3 cohorts of data leads to observing only 2 cutoffs. Theoretically, cutoffs in which individuals on both sides either did not get the scholarship, as is the case for college cohorts of 2000 and 2001, that are used in this paper, or both got the scholarship, as is the case of college cohorts of 2002 to 2003, could be used. An issue with using post cohorts, such as 2002 to 2003, is that the birthdate cutoff changed in 1989 to November, which might add bias to estimates of relative age. Other similar designs uses pre and post cohorts (Canaan, 2020) to contribute to untreated discontinuity.

smooth across the birth cutoff when crossing the cutoff affects relative age, but not scholarship receipt.

V.1 First Stage

The first stage requires that students just to the left of the birth cutoff be discontinuously less likely to receive the scholarship than those just to the right. The strength of the first stage, the estimated difference at the cutoff, is ψ from Equation 7 and k' in Equation 6. Since Figure IB shows a $c_{2002} = 53.3$ pp likelihood that students to the right of the 1983 birth cutoff enter the 2002 college cohort, it is expected that $\psi \approx 53.3$ pp. The average likelihood that students receive a scholarship, conditional on birthdate, is shown in Figure III. As shown in Figure IIIA, there is no increase to the likelihood of scholarship receipt at the 1982 birth cutoff for entering college during 2001 because the scholarship was not offered until the 2002 college cohort. Figure IIIB shows a discontinuous increase in likelihood of scholarship receipt at the 1983 cutoff. Using local regression, fit separately on both sides of the cutoff, with a triangular kernel and MSE-optimal bandwidths, the estimated difference at the cutoff, ψ, is 35.5 pp and statistically significant.

The 35.5 pp difference in scholarship receipt is smaller than the 53.3 pp difference in cohort shown in Figure IB. Academic eligibility is also an important factor, so Figure IIIC subsamples to students who meet academic eligibility requirements. Conditional on being academically eligible, students born after the cutoff are 49.6 pp more likely to receive the scholarship, which is very close to the estimated likelihood, 53.3 pp, that being born after the cutoff leads to students entering college during 2002. This difference represents a sufficiently large first stage.[15,16]

V.2 Local Smoothness

The variation that quasi-randomly assigns the scholarship is due to students being born on slightly different days, 20 years before the scholarship begins. Such variation makes local smoothness

[15]The first stage of fuzzy designs in education varies widely. For example, Jacob and Lefgren (2004) has a 90 pp difference, Hoekstra (2009) 38.8 pp, and Goodman et al. (2020) 1–10 pp.

[16]Figures A.III and A.IV show attrition from the Promise scholarship, which accords with extant findings (e.g., Carruthers and Özek (2016)).

of potential outcomes across the cutoff, an identifying assumption for RD (Hahn et al., 2001), plausible ex ante. Three implications of local smoothness are examined below.

V.2.1 Density Tests

Students could sort into treatment by strategically retaking a grade, or sort into attending a particular university if being beyond the cutoff birthday makes them eligible for the scholarship. Either case of selection into treatment calls local smoothness into question. Density tests examine whether this selection is an issue for local smoothness by testing for bunching in the distribution of birthdates in the sample at the birthdate cutoff (Cattaneo et al., 2019; McCrary, 2008).

The results of density tests, using an implementation from Cattaneo et al. (2019), appear in Figure IV.[17] The null hypothesis is that the densities are equal on either side of the cutoff, so failing to reject supports local smoothness. First, the density test is executed at the pre-scholarship cutoff because relative age might lead to selection into or out of the sample, regardless of how exceeding the birth cutoff affects scholarship receipt. Shown in Figure IVA, there is an increase to the density at the 1982 cutoff, which might be due to relative age effects. The estimated density to the left of the cutoff is 0.00144, and to the right it is 0.00185. However, it is not statistically significant (p-value = 0.1618).

Results at the post-scholarship cutoff are shown in Figure IVB. The estimated density to the left of the cutoff is 0.00133 and the estimated density to the right is 0.0014. These values are not statistically different (p-value = 0.8131). The likely reason that no statistically significant manipulation beyond the birth cutoff was observed is that it is unknown whether the birthdate cutoff leads to increased likelihood of receiving the scholarship in 2002 for those born during the early 1980s. Unlike for known academic cutoffs, the density tests support the identification assumption of local smoothness of potential outcomes across the birthdate cutoff.[18]

V.2.2 Covariate Balance

The density is smooth at the post cutoff, but it is elevated longer in the post-cutoff (Figure IVB) than in the pre-cutoff (Figure IVA), which is expected because there are documented increases in

[17] The bandwidths are MSE-optimal in both, and the bias-correction method is used. The unrestricted model is used because placing restrictions on the higher-order derivatives is likely a poor assumption due to relative age.

[18] Shown in Figures A.VA and A.VB, ACT score manipulation is evident around the cutoff.

enrollment in state after the scholarship began (Scott-Clayton, 2011). In addition to how observations away from the cutoff introduce bias to the estimate, this might introduce bias because those students could be different somehow. To assess whether the other students are different, observed covariates are substituted for the outcome in Equation 8.

Table I reports results of covariate balance tests at the post-cutoff. No statistically significant differences in gender, ACT composite scores, SAT combined scores, or any ACT subscores at the cutoffs were evident, and most estimated differences are very small. The one statistically significant difference is for high school GPA; those just to the right of the cutoff have a 0.13 higher high school GPA. While estimated precisely, this is a very small difference in high school GPA, only one-tenth of a letter grade.

The increase to continuous HS GPA suggests that academic eligibility is manipulated by those beyond the birth cutoff to be eligible to receive the scholarship (Pallais, 2009). This relates to Panel C of Table I, which examines how exceeding the birth cutoff affects various academic eligibility measures. Students were 5 pp more likely to be academically eligible, but this is not statistically significant. In line with the slight increase to GPA, students were 7 pp more likely to obtain an academically eligible HS GPA (3.0), and this is statistically significant at the 90% confidence. This is similar to Pallais (2009), in which students increased their ACT scores in response to a merit-based college scholarship that required a minimum score. Empirical evidence is mixed regarding whether academic eligibility is endogenous. Covariate balance tests suggest that those just to either side of the post-cutoff are comparable.

Another way to gauge differences across the cutoff is comparing post- and pre-cutoff balance tests. Results appear in Table A.II, and are largely similar to the post-cutoff. Students were much less likely to be male at the pre-scholarship cutoff, but that changes to be about equally likely to be male at the post-scholarship cutoff. For a view of raw data on covariates, including academic eligibility, see the plots in Figures A.VI and A.VII.[19]

[19] At the pre-scholarship cutoff, there are gender differences at the cutoff with those just beyond the cutoff being more likely to be female. This difference in gender does not exist at the post-scholarship cutoff which suggests that the scholarship may play a role in reducing male enrollment gaps. As shown in Figure A.VIA, there may be a u-shaped relationship of percent male to the left of the cutoff. To more definitively conclude that the scholarship influences gendered enrollment gaps, it would be more appropriate to confirm that this u-shaped relationship exists in other pre-scholarship years. For now, this represents suggestive evidence.

V.2.3 Relative Age and Season of Birth Effects

Those just left of the birthdate cutoff are likely the youngest of their peers, and those to the right are likely the oldest, and relative age can affect education outcomes.[20] Such a discontinuity in relative age might cause outcomes to not be smooth. If relative age causes potential outcomes to be discontinuous across the threshold, the effect of the scholarship remains unidentified in the fuzzy RD design. To examine whether relative age affects outcomes in this context, Equation 8 is estimated using $After82$ instead of $After83$, which tests whether students beyond the pre-scholarship cutoff are discontinuously more likely to reach the scholarship's renewal targets. If a discontinuous change in hitting renewal thresholds for the scholarship exists, when the scholarship is not offered, local smoothness is likely violated. If effects on college outcomes are found at a birthdate cutoff when scholarship receipt does not change discontinuously at that cutoff, the difference to the discontinuity estimator is preferred because it adjusts for this specific violation under additional assumptions.

Figure V demonstrates the effect of exceeding the 1982 pre-scholarship cutoff on the likelihood of meeting academic renewal targets. Figure VA finds a statistically insignificant 2.8 pp effect on meeting the credits target, and Figure VB finds a statistically insignificant 2 pp effect on meeting the GPA target. Continuity in realized outcomes at the pre-scholarship cutoff suggests that relative age does not invalidate local smoothness at the post-scholarship cutoff.

VI Results

This section discusses results of how strongly students respond to scholarship renewal targets. First, unscaled estimates of exceeding the cutoff on meeting the targets are reported. Figure VI shows a discontinuous increase to the likelihood of finishing sufficient credits (Figure VIA) and GPA (Figure VIB) to renew the scholarship for those to the right of the birth cutoff.

Hypothesis tests support these conclusions. Hypothesis testing is conducted on β from Equation 8, using local polynomial regressions, with a triangular kernel and MSE-optimal bandwidths. Estimates for β shown above the scatter plots in Figure VI are larger than the same estimates from

[20] Deming and Dynarski (2008) discusses the ambiguous, long-term effects of relative age, and the delaying of grade school, and Peña (2020) discusses mechanisms from relative age to human capital. Matta et al. (2016) examines how starting school age affects college admissions and earnings.

the cutoff that does not alter scholarship receipt. There is a 14.4 pp difference at the cutoff for credits in Figure VIA (p-value = .051), and there is a 19.3 pp difference at the cutoff for GPA in Figure VIB (p-value = .023).

Time trends in academic outcomes might be causing this difference at the post-scholarship cutoff that is not at the pre-scholarship cutoff. If time trends cause this difference, academically ineligible students, who do not receive the scholarship at either cutoff, will have increases in meeting targets from the 1982 to 1983 cutoff. Shown in columns 3 and 6 of Table A.III, there are no statistically significant increases to completing renewal credits or GPA from the pre-scholarship cutoff to the post-scholarship cutoff for academically ineligible students.[21] No improvements from the 1982 to 1983 cutoff for academically ineligible students suggests that time trends are not causing the difference between the pre- and post-scholarship cutoffs for academically eligible students.

VI.1 Fuzzy Regression Discontinuity

Table II reports main estimates of how scholarship receipt affects the likelihood of reaching academic renewal thresholds. Panel A reports estimates of the effect of the scholarship on the likelihood of earning at least 30 credits during freshmen year. Columns 1 uses the full sample, suggesting that the scholarship increases the likelihood of completing sufficient credits to renew the scholarship by 36 pp. Column 2 applies the bias-correction and robust standard errors approach, detailed in Calonico, Cattaneo, and Titiunik (2014), and the estimate remains nearly unchanged and statistically significant at 95% confidence.

One reason to be concerned about the estimates in columns 1 and 2 is that academically ineligible students are included. A more appropriate comparison could be academically eligible students who were born a few days too soon to receive the scholarship in comparison to students who are academically eligible and born a few days later. Column 3 removes all academically ineligible students, and this reduces the point estimate to 27 pp, with the same standard error, reducing statistical significance slightly. This is close to the 24 pp point estimate from Scott-Clayton (2011) regarding the effect of the scholarship on meeting the credit target.

Panel B reports estimates of the effect of the scholarship on the likelihood of meeting the GPA target, 2.75 GPA, during freshmen year. Columns 1 through 3 estimate a 31–34 pp treatment effect,

[21] Figure A.VIII shows the scatter plots.

which is statistically significant at 95% confidence. These estimates are close to estimates of the effect of the scholarship on credits from Panel A, but they are not close to those reported for the effect of the scholarship on GPA targets in Scott-Clayton (2011), which is 8 pp. For the remainder of the columns, which include academic controls, changing the kernel, and using other confidence intervals, the estimated point effects of the scholarship on credits and GPA targets are at least 20 pp, and they are statistically significant above at least 90 percent confidence in 5/8 specifications.

Relative Age Since estimates are far from zero, regardless of the kernel, bandwidth type, sample, or estimation adjustment, exploring other explanations is warranted. One explanation is relative age, which was estimated to have only a 2–3 pp effect, in comparison to 20–35 pp estimates in Table II, an influence that was not statistically significant. If relative age was the driver of these estimates, it would not explain why the effect of the scholarship on credits is nearly the same as prior estimates (Scott-Clayton, 2011).

Parametric Results One difference between current and prior estimates (Scott-Clayton, 2011) is the sample, so there might be heterogeneous effects from students who attend universities besides this one. It is impossible to address this directly using the available data, but it is possible to mirror the sample and estimation procedure more closely in Scott-Clayton (2011) by including more data away from the birth cutoff and using parametric methods to estimate the difference at the cutoff. Thus far, the estimation procedure has used MSE-optimal bandwidths, which resulted in about 90–151 days on both sides of the cutoff being used during estimation. Table A.V reports results with students from full cohorts of 2001 and 2002, and global linear and quadratic fits. Using the full 365 days, the point estimate for credits is approximately 0.2, and the point estimate for GPA is 0.04 or 0.06, which is the same discrepancy between GPA and credits targets from Scott-Clayton (2011), but without students from other universities. When only half a year is used (i.e. 180 days on both sides of the cutoff), the GPA estimate increases to 0.12. It is impossible to rule out heterogeneous results for students from other universities, but this is highly suggestive that the difference is more likely to be the different method, rather than a different sample.

Donut RD One explanation is that those closest to the cutoff are different in some way, such as being more likely to enter late if they are just to the left. To address this possibility, the literature uses donut RD estimates, during which observations closest to the cutoff are removed (Dague, DeLeire, & Leininger, 2017; Lee & Lemieux, 2010) so that they do not drive results. Table A.IV follows this procedure, removing individuals within 1, 2, and 3 weeks of the birthdate cutoff. Removing observations results in greater imprecision, but point estimates remained large, above 0.2, for both GPA and credits.

Continuous Outcomes Continuous outcomes might be informative about why results differ from Scott-Clayton (2011). It is important to understand how much the scholarship is increasing continuous outcomes to get them to the targets. Table A.VI reports estimates of the scholarship on continuous outcomes. In Panel A, the scholarship increases continuous credits by 4.87 or 5.61 in columns 1 and 2, but once academically ineligible people are removed, the point estimate decreases to 0.54 and is no longer statistically significant. This might suggest that academic eligibility is endogenous and a bad control, though that is inconsistent with the small, statistically insignificant influence of the cutoff on academic eligibility. Estimates for continuous outcomes differ from Scott-Clayton (2011), in which the point estimate is a 1.8 credit increase due to the scholarship.

The opposite was observed in Panel B for credits; the estimates increased and are more statistically significant after removing academically ineligible people. Estimates for credits suggest that academically eligible individuals who receive the scholarship have a 0.34–0.51 higher GPA than those who are academically eligible but were born too soon, with the upper end of the range equal to half a letter grade. Estimates for continuous outcomes differ from Scott-Clayton (2011), in which the reported effects are 1.8 credits and no effects are reported for GPA.

VI.2 Difference in Discontinuity

Difference in discontinuity estimates represent an important robustness check because thus far, the negligible effects of relative age, which have significant effects on many outcomes, have been a function of optimal bandwidth.[22] Reliance on optimal bandwidths means it remains possible that at other bandwidths, the effect of relative age is larger and the effects of the scholarship are smaller, which would move the difference in discontinuity estimates of the scholarship effect closer to zero. Difference in discontinuity estimates for the effect of the scholarship, with 95% confidence intervals appear in Figure VII. Estimates vary by bandwidth (on the x-axis) and kernel (triangular or uniform).[23]

Shown in Figure VIIA, point estimates of the scholarship's effect on credits targets do not differ appreciably by kernel or bandwidth. The 95% confidence intervals sometimes include zero due to greater uncertainty associated with estimating differences at two cutoffs. Nevertheless, point estimates are that students are nearly 24 pp more likely to receive the scholarship, which is nearly the same estimate for credits in Scott-Clayton (2011) and indicated by the dashed horizontal black line. This magnitude suggests that the fuzzy RD estimates of the scholarship on reaching credits targets are not confounded by relative age, since these point estimates are consistent with a large effect of the scholarship and negligible effects of relative age.

Shown in Figure VIIB, point estimates of the scholarship's effect on GPA targets do not differ much across bandwidth or kernel. Similar to the credits targets, the confidence intervals around the point estimates are wide and often include zero. In comparison to the 8 pp effect in Scott-Clayton (2011), shown by the black horizontal dashed line, estimates are closer to 24 pp for nearly all combinations of kernels and bandwidths. With the uniform kernel and the largest bandwidth possible, which is the closest possible to Scott-Clayton (2011), the estimate on hitting the GPA target is closest to the 0.08 estimate reported in Scott-Clayton (2011). The difference in discontinuity estimates are noisier, but they support the conclusion that relative age is negligible and the targets are responded to more equally than extant research suggests.

[22] Some other outcomes that relative age affects include: ADHD diagnosis (Elder, 2010), test scores and behavior (Berlinski, Galiani, & Gertler, 2009; Hemelt & Rosen, 2016), parental cohabitation and mothers' labor supply (Landersø, Nielsen, & Simonsen, 2019), grade completion (McEwan & Shapiro, 2008), high school leadership and grit (Dhuey & Lipscomb, 2008; Peña & Duckworth, 2018), crime (P. J. Cook & Kang, 2016; Depew & Eren, 2016), college enrollment (Matta et al., 2016), and becoming a CEO (Du, Gao, & Levi, 2012). Nonetheless, extant research often finds that the advantges of relative age decline over time (Black, Devereux, & Salvanes, 2011; Elder & Lubotsky, 2009) and that delaying the beginning of school is not worth sacrificing a year of labor income (Deming & Dynarski, 2008).

[23] To implement this estimator, data were formatted so that 6 months on either side of each cutoff relate to that particular cutoff. The same bandwidth was imposed at both the 1982 and 1983 cutoffs.

Heterogeneous Effects By Gender Found by extant research, the effects might be heterogeneous across subgroups, such as gender. The difference in discontinuity estimates by gender appear in Figure A.IX. Figure A.IXA and Figure A.IXB show that point estimates are larger for males. Male estimates are slightly more steady as the bandwidths change, but female estimates switch signs. In all figures, the confidence intervals are wide, making it difficult to determine whether the estimates are truly larger for men, though there is some evidence that they are. For completing requirements during each of the first three years, shown in Figure A.X men and women are equally likely to complete renewal credits, but men respond more strongly to the GPA requirement.

Heterogeneous Effects by Academic Ability Another advantage of using variation from the birthdate cutoff, as opposed to academic cutoffs, is that heterogeneity by academic ability can be examined with subgroup analyses. The difference in discontinuity estimates, by academic ability, appear in Figure A.XI. Figures A.XIA and A.XIC show how estimates for the credits and GPA targets differ for students above and below the median high school GPA for academically eligible students (3.71). There are no obvious differences in meeting the credits target, but the scholarship appears to increase the likelihood that those below the median HS GPA reach the GPA target more than those above it. Figures A.XIB and A.XID show estimates for meeting the targets, split by median ACT score among those who are academically eligible (24).Figure A.XID suggests no heterogeneous effects by ACT score on reaching the GPA target. Figure A.XIB suggests that those above the median ACT score respond more strongly to the credits target. The estimates are imprecise, but they suggest that effects differ by academic ability and help those with higher academic ability regarding the credits target.

VII External Validity

One concern with the estimates discussed thus far is that they are valid only among compliers near the cutoff. This is a limitation found in all LATE target estimands from RD, and a concern in merit-aid literature that uses academic cutoffs and RD to produce causal estimates of merit-aid scholarships (Dong & Lewbel, 2015; Goodman, 2008).[24] Estimates from academic cutoffs are

[24] Academic, merit-aid cutoffs from Goodman (2008) is used by Dong and Lewbel (2015) to motivate the development of statistical methods to assess whether treatment effects generalize to compliers away from RD cutoffs.

concerning because they are unlikely to generalize to students with disparate academic measures. It is plausible that estimates from the birth cutoff generalize to students away from the cutoff more readily. This section assesses the external validity of the estimates, from the birthdate cutoff, to compliers away from the cutoff, and never-takers and always-takers at the cutoff using methods from Dong and Lewbel (2015) and Bertanha and Imbens (2019). This could also been see as understanding heterogeneous treatment effects across compliance types for the estimates from Scott-Clayton (2011).

VII.1 Compliers Away from Cutoff

One concern with RD estimates local to the 1 September birth cutoff is that extant research suggests that they do not generalize to other date cutoffs. Examples include birthdate cutoffs changing the quantity of education (Black, Devereux, & Salvanes, 2008; Black et al., 2011; Lleras-Muney, 2002), and that a mother's characteristics differ across season of birth (Buckles & Hungerman, 2013). So, estimated effects might be different were the birth cutoff in November or January.

Equal treatment effects of the scholarship away from the birth cutoff implies that small increases in birthdate away from the cutoff have the same marginal effect on outcomes (Dong & Lewbel, 2015), which requires estimating a treatment effect derivative (TED) that tests whether the slope of the fitted line (i.e. of outcomes on the running variable) to the right of the cutoff equals the slope of the fitted line to the left of the cutoff (Dong & Lewbel, 2015).[25] If the treatment effects and/or compliers change away from the cutoff, LATE estimates from RD are not stable away from the cutoff (Cerulli, Dong, Lewbel, & Poulsen, 2017; Dong & Lewbel, 2015).

Table III reports estimates of TED. Shown in the TED estimate in Panel A, an increase to the birthdate cutoff by 10 days leads to a 5 pp increase in the likelihood of exceeding the renewal target for credits, but it is small and statistically insignificant. The TED estimate for GPA targets in Panel B is similarly sized and also statistically insignificant, suggesting that treatment effects are valid among compliers away from the cutoff. Since this is a fuzzy design, the complier probability

[25]The TED approach (Dong & Lewbel, 2015) is used as opposed to the conditioning approach from J. D. Angrist and Rokkanen (2015) because it requires fewer covariates to meet the assumptions.Dong and Lewbel (2015) reports that the procedure is simple and that researchers estimate this quantity regularly anyway in RD designs but do not report it.

derivative (CPD) is also estimated.[26] The insignificant estimate of CPD suggests that the likelihood of compliance also does not change away from the cutoff.[27] Since the TED and CPD are both statistically insignificant, the compliers and treatment effects appear stable away from the birth cutoff (Cerulli et al., 2017, pg. 325).

Many policy debates involve treatment effects that could change were a qualifying threshold to change. The TED estimate informs whether a different birth cutoff date leads to different effects. The cutoff birth date was moved from September to October in West Virginia for the 1984 birth cohort, so this is an important question. Under a local policy invariance assumption, TED equals the marginal threshold treatment effect (MTTE), the effect of changing the cutoff slightly (Dong & Lewbel, 2015).[28] Local policy invariance means that the expected effect of the scholarship at a given birthdate does not change if the birth cutoff changes marginally (Cerulli et al., 2017, pg. 335). A potential violation, which causes MTTE $\neq$ TED, can occur if mothers who give birth during October are significantly different than those who give birth during September (Buckles & Hungerman, 2013), and students with October mothers somehow benefit differently from having the scholarship. If local policy invariance is met, MTTE = TED. The TED is insignificant, suggesting that moving the cutoff to October does not affect estimates of the scholarship's influence.

VII.2 Non-Compliers

The limitation of using TED from Dong and Lewbel (2015) is that it relaxes only the local to the cutoff condition. Whether estimates are externally valid among other compliance types is important because two ways exist to not be considered a complier—receiving a grade entrance exemption, and not graduating 1 grade per year and then entering college. To investigate this possibility, a test from Bertanha and Imbens (2019) is used that evaluates whether compliance type is independent of the running variable. The test estimates differences in reaching renewal targets

[26] The CPD is the first stage analog of the reduced form TED.

[27] Cerulli et al. (2017) discusses additional conditions on the relative TED and CPD, which are a function of other estimates in the TED estimating equation. The relative TEDs and CPDs are below 1 or 2, which is not ideal for establishing external validity because a relative TED below 1 suggests that the treatment effect could change sign for some subset of people in the range (Cerulli et al., 2017, pg. 325).

[28] Abbring and Heckman (2007) define policy invariance as "an assumption that an agent's outcomes only depends on the treatment assigned to the agent and not separately on the mechanism used to assign treatments," and Dong and Lewbel (2015) extend this to a local policy invariance assumption for those near the cutoff.

between never-takers and untreated compliers, and between treated compliers and always-takers, at the cutoff, and then jointly tests that these differences are zero (Bertanha & Imbens, 2019).

The proportion of compliers, always-takers, and never-takers, among academically eligible students, are summarized in Panel A of Table IV. Nearly half (46.8%) of the students are compliers, meaning that they either were born to the left of the cutoff and did not receive the scholarship, or were born to the right of the cutoff and did receive the scholarship. Always-takers, meaning those born to the left of the cutoff and receiving the scholarship, comprise 19.6% of the students. Always-takers could arise because students could retake grades or be born in a different state and be subject to a different state's cutoff. The never-takers, meaning those born to the right of the cutoff but who did not receive the scholarship, comprise 33.6% of the students. This group could have skipped a grade or could be eligible for other types of aid, such that they do not receive the scholarship.[29]

Panel B of Table IV reports estimated differences between never-takers in comparison to untreated compliers, and treated compliers in comparison to always-takers, for meeting the credits target. The first row shows nearly no difference between never-takers and untreated compliers. The second row shows a noticeable but statistically insignificant difference between always-takers and treated compliers. The joint test that both of these differences are zero is not rejected, suggesting that estimates on credits are externally valid to never-takers and always-takers. Panel C controls for high school GPA and ACT scores, with estimates getting slightly smaller and less statistically significant. This confirms that the estimates are externally valid to non-compliers beyond the cutoff for credits.

Panel D of Table IV reports estimated differences between never-takers in comparison to untreated compliers, and treated compliers in comparison to always-takers, for meeting the GPA target. The never-takers are much more likely (0.352) to reach the GPA target than untreated compliers, and treated compliers do slightly worse than always-takers. The joint F-test fails to reject external validity at conventional levels, with a p-value of 0.104. After controlling for HS GPA and ACT scores, the never-takers do even better than the untreated compliers, which causes the pair of joint restrictions ($F = 7.943$) to be rejected at 95% confidence. Thus, adjusting for academic ability appears to cause the estimates to not be externally valid to non-compliers for GPA. One reason

[29] Table A.VII shows participation, including academically ineligible students.

that never-takers are doing much better than untreated compliers is that one way for a student to become a never-taker is by skipping grades, which would explain their high GPAs.

VIII Conclusion

Despite the proliferation of renewal requirements, there have been few answers as to how strongly students respond to these requirements. Within research on merit-aid, where these requirements appear widely, a primary obstacle to providing this answer has been finding an identification strategy that addresses selection into treatment and provides externally valid estimates. This paper advances new answers on the effectiveness of renewal requirements by leveraging a new insight on how college cohort based variation in scholarship receipt arises. This strategy is internally valid, helps disentangle the seemingly different average responses to various targets, and shows that the response to credits is more externally valid than the response to GPA targets.

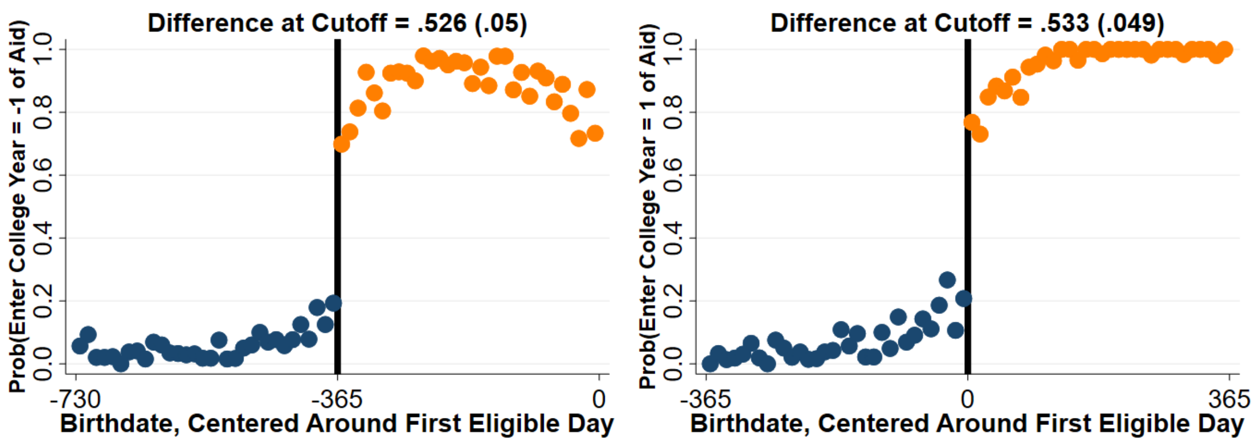
Perhaps the main limitation of the approach is that it requires observing many individuals with birthdates near the birth cutoff. The data used is only 1 university in West Virginia which reduces concerns of heterogeneity bias due to aggregation across universities, but this data requirement combined with the modest sample size does lead to estimates that are sometimes imprecise and the optimal bandwidths are somewhat long. This limitation also extends to estimates by gender or race which are too imprecise, but are nonetheless important to obtain for policy. An additional limitation is that the effect of just the goal cannot be observed, instead it is the combined effect of the requirement and the aid that is awarded for reaching it.

Despite these caveats, the results remain critical for informing renewal requirement policies since institutional scholarships with renewal requirements are ubiquitous. The results say that on average the response to credits or GPA targets are both strong for students who are able to finish 1 grade per year. But, only the credits response is shown to also be strong for never-takers and always-takers who might not finish 1 grade per year. This implies its more likely that credits targets will be effective for scholarships where intended recipients may not finish 1 grade per year during grade school.

There are several fruitful directions for future research. With more comprehensive data, there may be adequate sample sizes to investigate treatment effect heterogeneity across gender, race, and academic ability conditional on being academically eligible. Furthermore, the renewal requirements are effective in other years besides freshman year and carefully accounting for attrition out of the sample and out of the scholarship could reveal insights on renewal requirements in later years. It could also be worthwhile to investigate the effects of these targets separate of one another or with different amounts of financial aid.

Figure I: Kindergarten Entry Birthdate Cutoffs Affect Year of College Entry

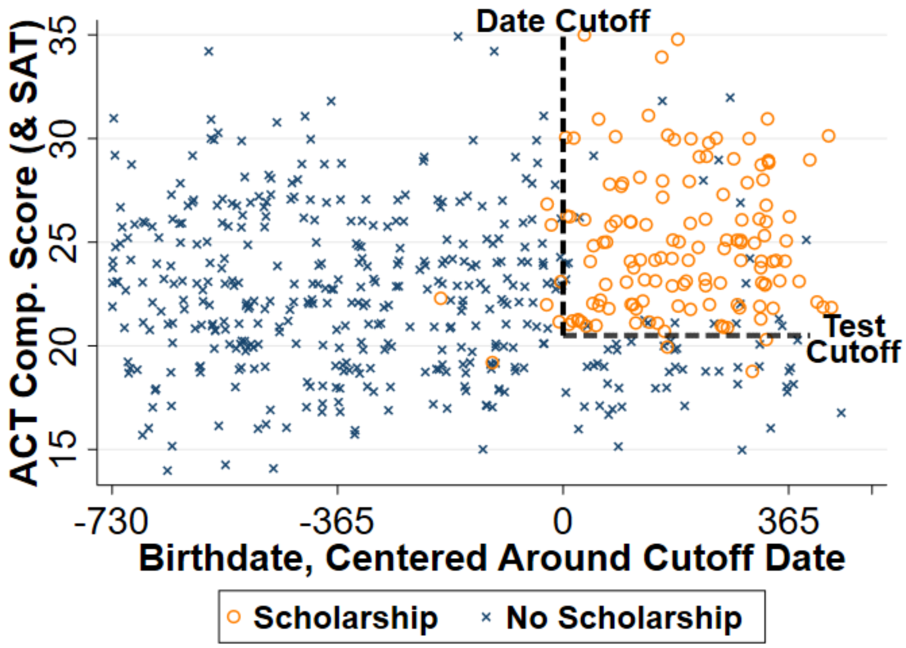
(A) Pre-Scholarship Cutoff on Enter in 2001 (B) Post-Scholarship Cutoff on Enter in 2002



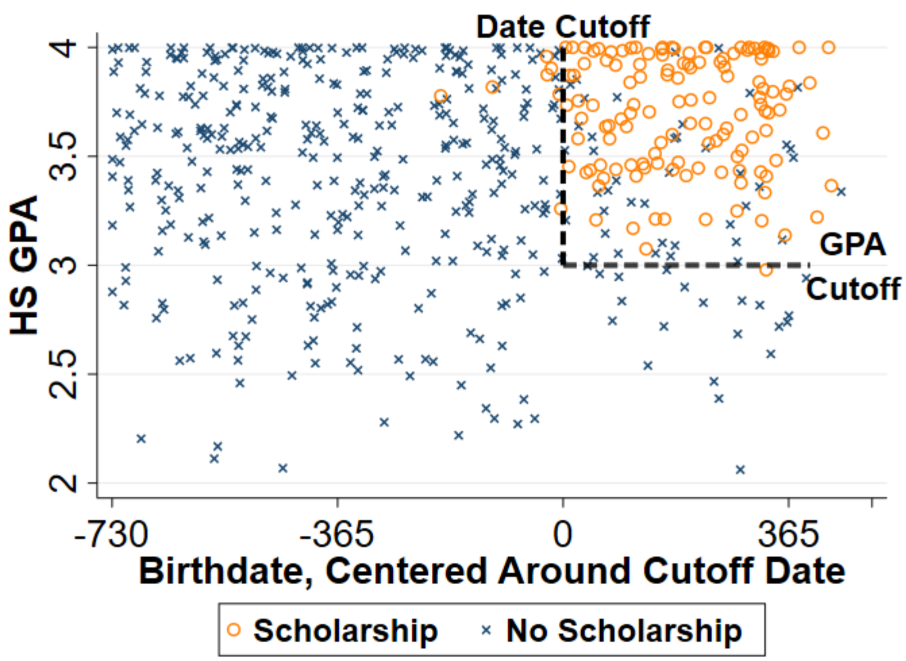
Note: The data are from university administrative records and cover in-state students beginning college in 2000-2002. These figures plot binned averages by re-centered birthdate. Panel A plots the probability of entering college in 2001. Panel B plots the probability of entering college in 2002. The estimates presented above the graphs are for c_{2001} and c_{2002} respectively from Equation 5. The estimates come from local polynomial regressions, fit separately on either side of the cutoff, using MSE-optimal bandwidths and triangular kernels. Equation 7 describes the regression after substituting cohort for scholarship as the dependent variable. No data is observed on the cohorts after 2002 which is why (B) goes to 1 on the right side.The standard errors are shown in parentheses and come from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors (Calonico et al., 2014).

Figure II: Academic Eligibility, Temporal Eligibility, and Scholarship Receipt

(A) Birthdate, Entry Test Score, and Scholarship



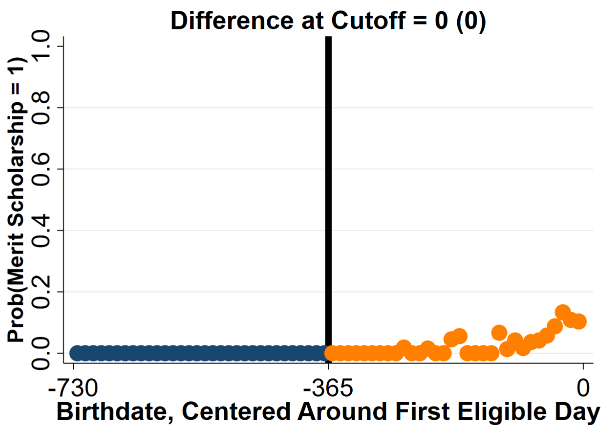
(B) Birthdate, HS GPA, and Scholarship



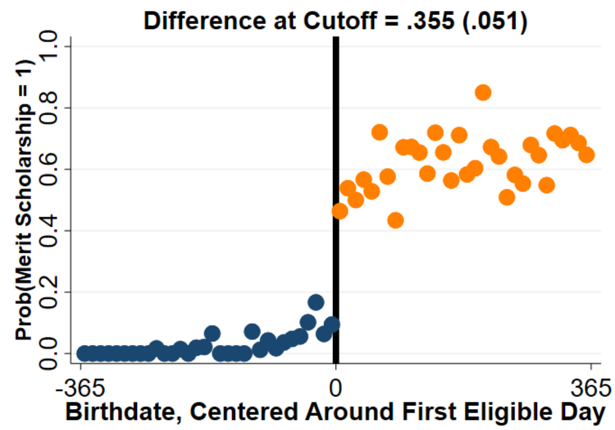
Note: 10% random subsample of data used to avoid overcrowded plots. N = 559. Birthdates centered at September 1, 1983, the first eligible date for the scholarship. See Section II for details as to why this cutoff matters. The cutoff ACT score is 20.5. SAT score converted to ACT using the conversion chart at Edwards (2020). The cutoff HS GPA is 3.0. See Section III.1 for eligibility details. A small amount of spherical random noise is added to avoid exactly overlapping points.

Figure III: The Post-Scholarship Kindergarten Birthdate Cutoff Affects College Scholarship Receipt

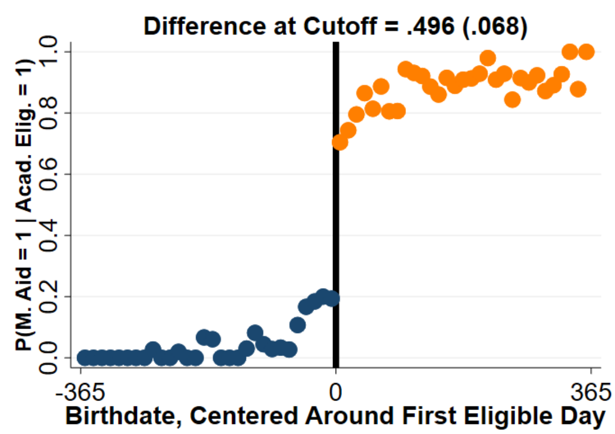
(A) Pre-Scholarship Cutoff on Scholarship Receipt



(B) Post-Scholarship Cutoff on Scholarship Receipt



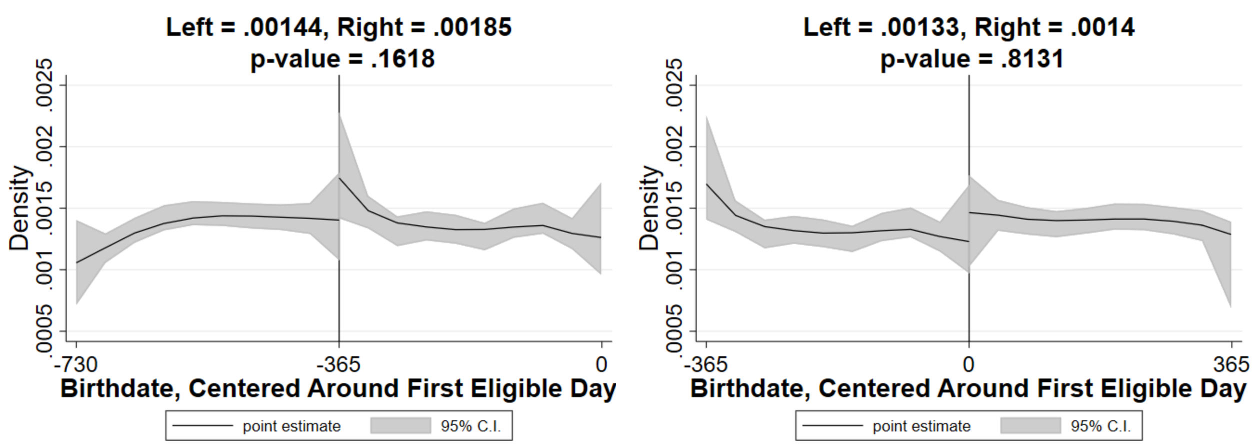
(C) Post-Scholarship Cutoff on Scholarship Received, Academically Eligible Students



Note: The data are from university administrative records and cover in-state students beginning college in 2000-2002. These figures plot binned averages by re-centered birthdate. The dependent variable is Promise scholarship receipt in all panels. Panel A shows students born within 365 days of the pre-scholarship, 1982 birth cutoff. Panel B shows students born within 365 days of the post-scholarship, 1983 birth cutoff. Panel C shows academically eligible students born within 365 days of the post-scholarship, 1983 birth cutoff. For more details on pre and post scholarship cutoffs see Section III.1. The estimates presented above the graphs are for k_{2001} and k_{2002} respectively from Equation 6. The estimates come from local polynomial regressions, fit separately on either side of the cutoff, using MSE-optimal bandwidths and triangular kernels. Equation 7 describes the regressions. The standard errors are shown in parentheses and come from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors (Calonico et al., 2014).

Figure IV: Birthdate Running Variables' Density Continuity Around Cutoffs

(A) McCrary Density Test, Pre-Scholarship (B) McCrary Density Test, First Scholarship Cohort



Note: Panels A and B are density tests at the pre and post-scholarship birthdate cutoffs. Left is the point estimate to the left of the cutoff and right is point estimate to the right. Local quadratic approximation used to construct the density estimator and local cubic approximation used for bias-correct density estimators (Cattaneo et al., 2019). Standard errors are jackknife and also robust to additional uncertainty from bias-correction. The p-value is from a test for whether density to the left equals the density to the right. Failure to reject indicates that the density is smooth across the cutoff, consistent with identifying assumptions (McCrary, 2008).

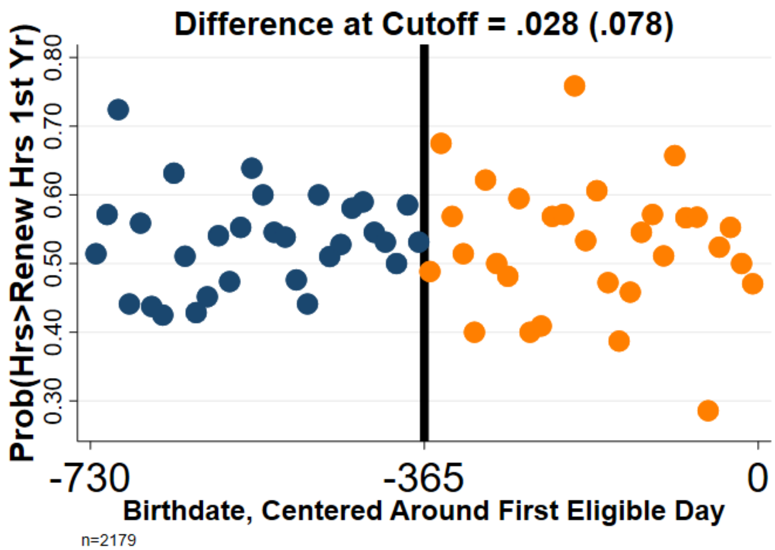
Table I: Pre-Treatment Covariate Balance, Post Cutoff Only

	(1)	(2)	(3)	(4)
Panel A: Gender and Scores				
	Male	ACT Composite	SAT Combined	HS GPA
After Cutoff	0.03	-0.05	40.04	0.13**
	(0.06)	(0.44)	(33.51)	(0.05)
Dep. Var. Mean	.499	22.74	1121.204	3.42
Bandwidth	150	167	192	190
Observations	1489	1362	340	1868
Panel B: ACT Subscores				
	(1)	(2)	(3)	(4)
	ACT English	ACT Math	ACT Reading	ACT Science
After Cutoff	0.00	0.11	-0.31	0.00
	(0.55)	(0.48)	(0.60)	(0.41)
Dep. Var. Mean	22.859	21.23	23.681	22.553
Bandwidth	147	198	179	208
Observations	1210	1616	1446	1683
Panel C: Academic Eligibility				
	(1)	(2)	(3)	(4)
	Academic Eligibility	HS GPA	ACT Composite	SAT Combined
After Cutoff	0.06	0.07*	0.05	0.03
	(0.05)	(0.04)	(0.05)	(0.08)
Dep. Var. Mean	.635	.807	.699	.792
Bandwidth	180	184	216	254
Observations	1760	1803	1748	449

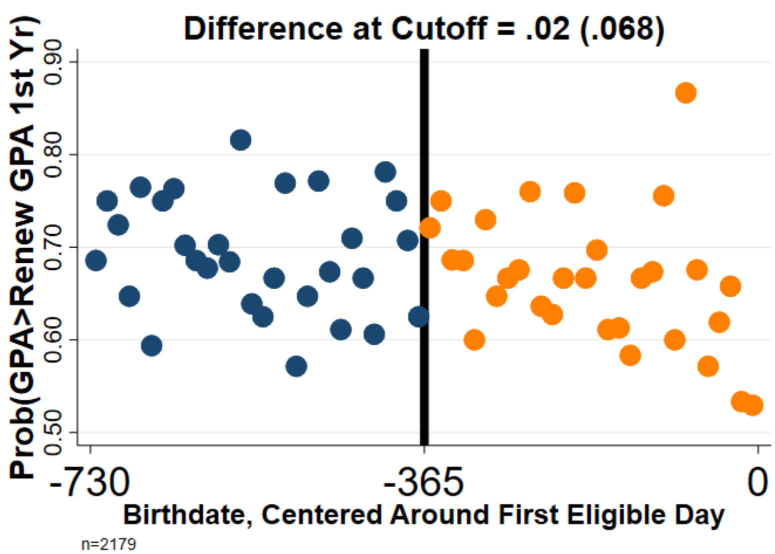
Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors shown in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. Coefficients are for $After83$ from local linear regressions in Equation 8. Triangular kernels and MSE-optimal bandwidths are used. The dependent variable is above each column. Plots for variables in Panel A shown in Figure A.VI. Panel C uses binary variables as outcomes, Figure A.VII shows the plots. Academic eligibility is a binary variable equaling 1 if students exceeded both the HS GPA (3.0) and entry test (ACT $\geq$ 21 or SAT $\geq$ 1000) score requirements to be eligible for the Scholarship. This variable is described in Section III.1. Below each coefficient is the mean of the outcome. Students who enter in the 2001 or 2002 cohort are included. Furthermore, the pre-cutoff estimates and the difference in discontinuities estimates are shown in Table A.II

Figure V: How the Birth Cutoff Affects Reaching Scholarship Renewal Targets When The Cutoff Does Not Affect Scholarship Receipt

(A) Exceed Renewal Credits



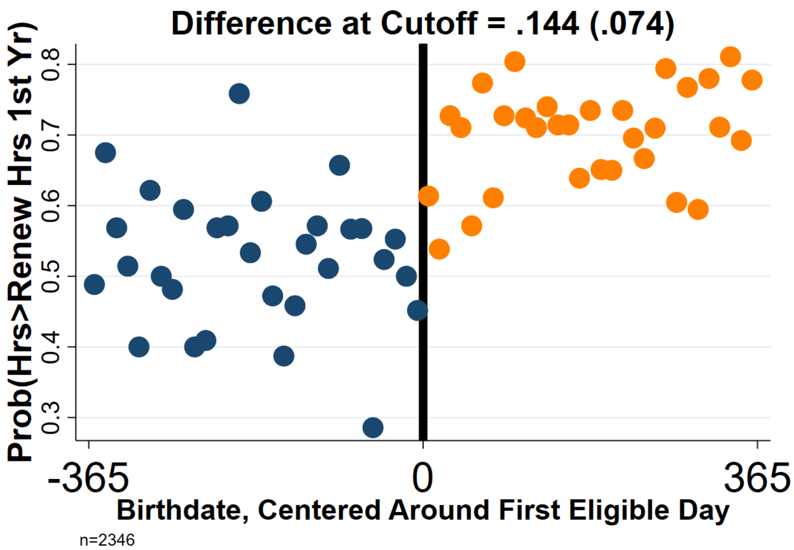
(B) Exceed Renewal GPA



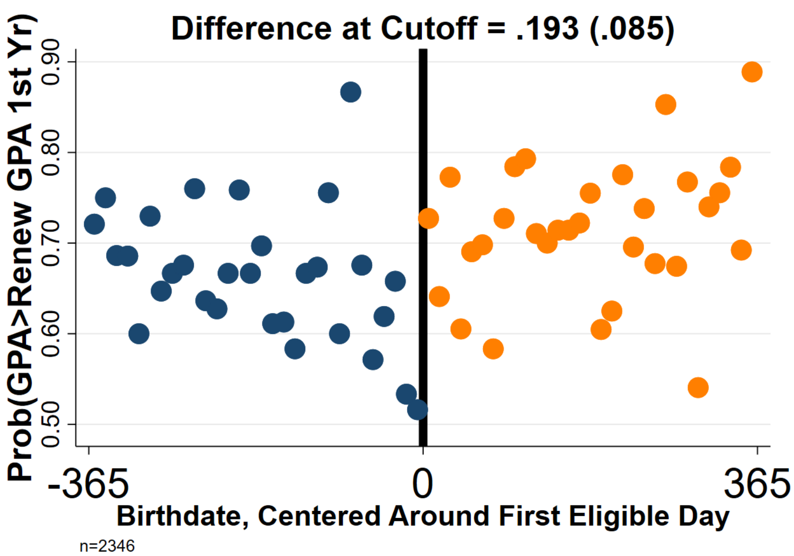
Note: The data are from university administrative records and cover academically eligible, in-state students born within 365 days of the pre-scholarship, 1982 birth cutoff. These figures plot binned averages by re-centered birthdate. In Panel A the dependent variable is whether a student reached the academic renewal threshold for credits (30) in their first year of college. In Panel B the dependent variable is whether a student reached the academic renewal threshold for GPA (2.75) in their first year of college. The scales differ. The estimates presented above the graphs are for $After82$ from Equation 8. The coefficient estimates come from local linear regressions, fit separately on either side of the cutoff, using MSE-optimal bandwidths and triangular kernels. The standard errors are shown in parentheses and come from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors (Calonico et al., 2014).

Figure VI: How the Birth Cutoff Affects Reaching Scholarship Renewal Targets, When The Cutoff Does Affect Scholarship Receipt

(A) Exceed Renewal Credits, Scholarship Birth Cutoff



(B) Exceed Renewal GPA, Scholarship Birth Cutoff



Note: The data are from university administrative records and cover academically eligible, in-state students born within 365 days of the post-scholarship, 1983 birth cutoff. These figures plot binned averages by re-centered birthdate. In Panel A the dependent variable is whether a student reached the academic renewal threshold for credits (30) in their first year of college. In Panel B the dependent variable is whether a student reached the academic renewal threshold for GPA (2.75) in their first year of college. The scales differ. The estimates presented above the graphs are for $After83$ from Equation 8. The estimates come from local polynomial regressions, fit separately on either side of the cutoff, using MSE-optimal bandwidths and triangular kernels. The standard errors are in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors (Calonico et al., 2014).

Table II: Local Polynomial Estimates of Scholarship on Meeting First Year Credits and GPA Renewal Thresholds

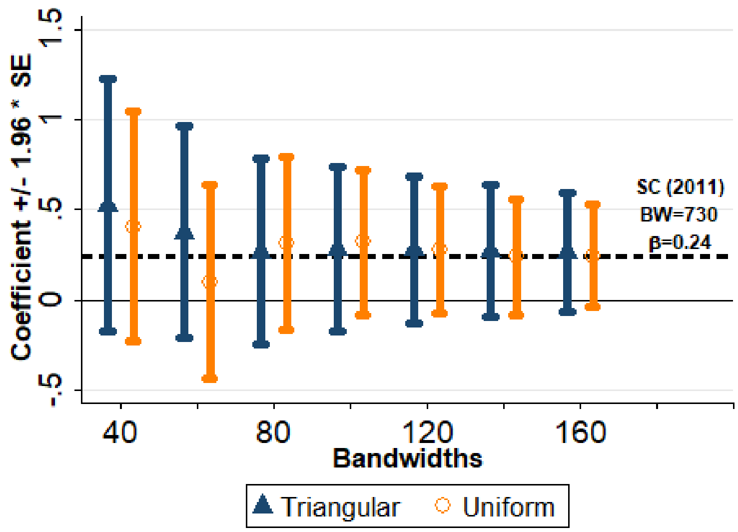
Panel A: Dep. Var.: Renew Credits							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Scholarship	0.36**	0.35**	0.27*	0.24	0.24	0.22*	0.29*
	(0.15)	(0.18)	(0.16)	(0.16)	(0.19)	(0.14)	(0.16)
Bandwidth	141	141	135	134	91	133	90
Effective Obs. Left	680	680	395	395	248	390	246
Effective Obs. Right	740	740	471	471	319	456	316
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.
Academic Controls				X	X	X	X
Bandwidth 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**	0.31**	0.40**	0.20	0.32**
	(0.15)	(0.18)	(0.16)	(0.15)	(0.19)	(0.13)	(0.16)
Bandwidth	151	151	140	142	140	137	93
Effective Obs. Left	720	720	402	406	268	398	261
Effective Obs. Right	777	777	486	497	336	482	327
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
Academic Controls				X	X	X	X
Bandwidth Criterion	MSE	MSE	MSE	MSE	CER	MSE	CER
Kernel	Tri	Tri	Tri	Tri	Tri	Uni	Uni

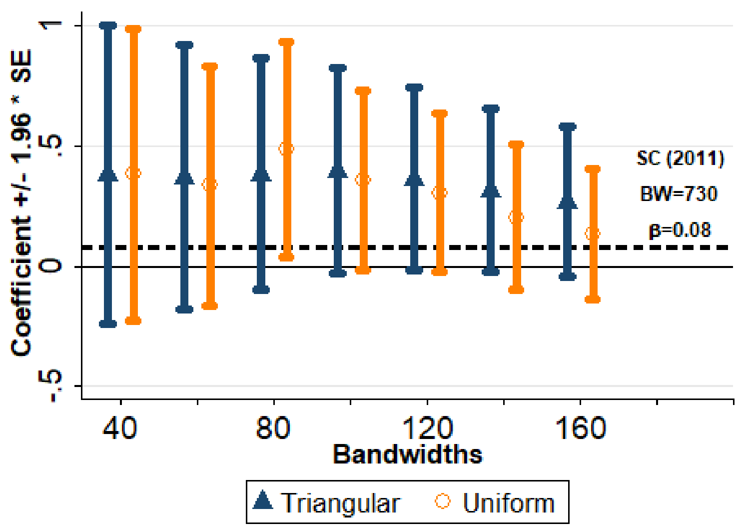
Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The standard errors are in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. The estimate is the effect of receiving the Promise Scholarship. The two estimating equations are Equation 7 and Equation 8. Under the assumptions outlined in Section IV.1.2, the estimate shown ($\tau = \frac{\beta_{After83}}{\psi_{After83}}$) can be interpreted as a LATE estimand for compliers. Estimates of ψ are shown in Figure III and estimates of $\beta_{After83}$ are shown in Figure VI. The equation models $f(Birthdate)$ as local linear regression. In all panels, the birthdate cutoff has a large, statistically significant effect on scholarship receipt. In Panel A the dependent variable is a dummy for finishes 30 credits (the renewal requirement) freshman year. In Panel B the dependent variable is a dummy for exceeds a 2.75 GPA (the renewal requirement) freshman year. In columns 1 and 2, the full sample is used. From column 2 onwards, the bias-correction and standard error adjustment in Calonico, Cattaneo, Farrell, and Titiunik (2017) are used with a local quadratic equation for bias-correction and the standard errors are adjusted for the bias-correction. From column 3 onwards, only students who exceed the academic eligibility cutoffs are included. From column 4 onwards, academic control variables (high school GPA and entry test score) are included. Columns 5 and 7 use coverage error rate optimal bandwidths. Columns 6 and 7 use uniform kernels.

Figure VII: Difference in Discontinuity Estimates of Scholarship's Effect on Exceeding Renewal Targets By Bandwidth and Kernel

(A) Renewal Credits



(B) Renewal GPA



Note: Coefficients represent estimated effect of scholarship from difference-in-discontinuity (Grembi et al., 2016). The point estimates equal $\frac{\beta_{After83}}{\psi_{After83}} - \beta_{After82}$. The corresponding equations are shown in Equation 7 and Equation 8. Whiskers represent 95% confidence intervals. Standard errors are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 6 neighbors. Triangular and uniforms kernels used. The different bandwidths indicate the number of days used from the respective cutoffs. Same bandwidth imposed at both cutoffs. In Panel A the dependent variable is exceed renewal credits in freshman year. In Panel B the dependent variable is exceed renewal GPA in freshman year. See Section III for details. All regressions subsample to academically eligible students. 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.

Table III: Estimates of Treatment Effect Derivative and Complier Probability Derivative for Investigating Validity of Estimates for Compliers Away from Cutoff

Panel A: Dep. Var.: Renew Credits	
	(1)
TED	0.005
	(0.008)
CPD	0.001
	(0.004)
Relative TED	.456
Relative CPD	3.102
Bandwidth	128
Observations	827
Polynomial Degree	2

Panel B: Dep. Var.: Renew GPA	
	(1)
TED	0.004
	(0.014)
CPD	0.007
	(0.007)
Relative TED	1.287
Relative CPD	.871
Bandwidth	86
Observations	539
Polynomial Degree	2

Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors are shown in parentheses and are calculated by the delta method. TED stands for treatment effect derivative, the slope of the treatment effect as one gets farther away from the birth cutoff (Dong & Lewbel, 2015). CPD stands for complier probability derivative, the slope of the first stage as one gets farther away from the birth cutoff (Dong & Lewbel, 2015). The relative TED and CPD are defined in Cerulli et al. (2017). All estimates are based on local quadratic regressions. All regressions use the triangular kernel. Students within 365 days of the post-scholarship, 1983 cutoff are included.

Table IV: External Validity to Always-Takers and Never-Takers

	Estimate	Std. Error	p-Value	Bandwidth
Panel A: Participation				
	Estimate	Std. Error	p-Value	
Compliers	0.468	0.088	0.000	
Always-Takers	0.196	0.059	0.001	
Never-Takers	0.336	0.063	0.000	
Panel B: Renewal Credits				
	Estimate	Std. Error	p-Value	Bandwidth
Never-Takers - Untreated Compliers	-0.032	0.131	0.809	195.860
Treated Compliers - Always-Takers	-0.172	0.166	0.301	214.148
Joint F-Test	1.116		0.572	
Panel C: Renewal Credits, Covar. Adjusted				
	Estimate	Std. Error	p-Value	Bandwidth
Never-Takers - Untreated Compliers	-0.015	0.126	0.907	195.860
Treated Compliers - Always-Takers	-0.144	0.167	0.389	214.148
Joint F-Test	0.749		0.688	
Panel D: Renewal GPA				
	Estimate	Std. Error	p-Value	Bandwidth
Never-Takers - Untreated Compliers	0.352	0.166	0.034	125.706
Treated Compliers - Always-Takers	-0.037	0.218	0.866	158.087
Joint F-Test	4.532		0.104	
Panel E: Renewal GPA, Covar. Adjusted				
	Estimate	Std. Error	p-Value	Bandwidth
Never-Takers - Untreated Compliers	0.420	0.149	0.005	125.706
Treated Compliers - Always-Takers	-0.018	0.229	0.936	158.087
Joint F-Test	7.943		0.019	

Note: In Panels B and C the outcome variable is a binary variable for meeting the credits academic renewal target. In Panels D and E, the outcome variable is a binary variable for meeting the GPA academic renewal target. Nonparametric bias-corrected estimates are obtained by local quadratic regression, triangular kernel, and IK optimal bandwidth and are shown in last column. Panels B and D use no covariates. Panels C and E use high school GPA and ACT score as controls. The standard errors of all estimates are computed using 1,000 simulations of the Wild Bootstrap of Bartalotti, Calhoun, and He (2017) which captures the additional variance from the bias correction. The F-test is a χ^2 statistic computed as a quadratic form of the 2 X 1 vector of the conditional jumps inversely weighted by the covariance matrix of such vector. See Bertanha and Imbens (2019) for details.

References

Abbring, J. H., & Heckman, J. J. (2007). Chapter 72 econometric evaluation of social programs, part iii: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. In J. J. Heckman & E. E. Leamer (Eds.), (Vol. 6, p. 5145-5303). Elsevier. doi: https://doi.org/10.1016/S1573-4412(07)06072-2

Angrist, J., Autor, D., & Pallais, A. (2021, 12). Marginal Effects of Merit Aid for Low-Income Students. *The Quarterly Journal of Economics*. (qjab050) doi: 10.1093/qje/qjab050

Angrist, J. D., & 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.

Barrow, L., Richburg-Hayes, L., Rouse, C. E., & Brock, T. (2014). Paying for performance: The education impacts of a community college scholarship program for low-income adults. *Journal of Labor Economics*, *32*(3), 563–599.

Bartalotti, O., Calhoun, G., & He, Y. (2017). Bootstrap confidence intervals for sharp regression discontinuity designs. In *Regression discontinuity designs.* Emerald Publishing Limited.

Beattie, G., Laliberté, J.-W. P., Michaud-Leclerc, C., & Oreopoulos, P. (2019). What sets college thrivers and divers apart? A contrast in study habits, attitudes, and mental health. *Economics Letters*, *178*, 50–53.

Berlinski, S., Galiani, S., & Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, *93*(1-2), 219–234.

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

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

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2008). Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births. *The Economic Journal*, *118*(530), 1025–1054.

Black, S. E., Devereux, P. J., & 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.

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

Buckles, K. S., & 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., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

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

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

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

Cerulli, G., Dong, Y., Lewbel, A., & Poulsen, A. (2017). Testing stability of regression discontinuity models. In *Regression discontinuity designs.* Emerald Publishing Limited.

Clark, D., Gill, D., Prowse, V., & Rush, M. (2020). Using goals to motivate college students: Theory and evidence from field experiments. *Review of Economics and Statistics*, *102*(4), 648–663.

Cohodes, S. R., Grossman, D. S., Kleiner, S. A., & 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.

Cook, P. J., & Kang, S. (2016). Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation. *American Economic Journal: Applied Economics*, *8*(1), 33–57.

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., Mustard, D. B., & Sridhar, D. J. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program. *Journal of Labor Economics*, *24*(4), 761–786.

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

Dague, L., DeLeire, T., & Leininger, L. (2017, May). The effect of public insurance coverage for childless adults on labor supply. *American Economic Journal: Economic Policy*, *9*(2), 124-54. doi: 10.1257/pol.20150059

Deming, D., & Dynarski, S. (2008). The lengthening of childhood. *Journal of Economic Perspectives*, *22*(3), 71–92.

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.

Depew, B., & Eren, O. (2016). Born on the wrong day? School entry age and juvenile crime. *Journal of Urban Economics*, *96*, 73–90.

Dhuey, E., & Lipscomb, S. (2008). What makes a leader? Relative age and high school leadership. *Economics of Education Review*, *27*(2), 173–183.

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

Du, Q., Gao, H., & Levi, M. D. (2012). The relative-age effect and career success: Evidence from corporate CEOs. *Economics Letters*, *117*(3), 660–662.

Dynarski, S., & Scott-Clayton, J. (2013). *Financial aid policy: Lessons from research* (Tech. Rep.). 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.

Elder, T. E. (2010). The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates. *Journal of Health Economics*, *29*(5), 641–656.

Elder, T. E., & Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement impacts of state policies, family background, and peers. *Journal of Human Resources*, *44*(3), 641–683.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, *25*(4), 191–210.

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

2131.

Goodman, J., Gurantz, O., & Smith, J. (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, *12*(2), 115–158.

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

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

Hemelt, S. W., & Rosen, R. B. (2016). School entry, compulsory schooling, and human capital accumulation: Evidence from Michigan. *The BE Journal of Economic Analysis & Policy*, *16*(4).

Hendren, N., & Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, *135*(3), 1209–1318.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, *91*(4), 717–724.

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, *86*(1), 226–244.

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

Landersø, R. K., Nielsen, H. S., & Simonsen, M. (2019). Effects of school starting age on the family. *Journal of Human Resources*, 1117–9174R1.

Lavecchia, A. M., Liu, H., & Oreopoulos, P. (2016). Behavioral economics of education: Progress and possibilities. In *Handbook of the economics of education* (Vol. 5, pp. 1–74). Elsevier.

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

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

Leuven, E., Oosterbeek, H., & Van der Klaauw, B. (2010). The effect of financial rewards on students' achievement: Evidence from a randomized experiment. *Journal of the European Economic Association*, *8*(6), 1243–1265.

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.

Londono-Velez, J., Rodriguez, C., & Sánchez, F. (2020). Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser Pilo Paga in Colombia. *American Economic Journal: Economic Policy*, *12*(2), 193–227.

Matta, R., Ribas, R. P., Sampaio, B., & 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.

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

McEwan, P. J., & Shapiro, J. S. (2008). The benefits of delayed primary school enrollment discontinuity estimates using exact birth dates. *Journal of Human Resources*, *43*(1), 1–29.

Miller, S., & Wherry, L. R. (2019). The long-term effects of early life medicaid coverage. *Journal of Human Resources*, *54*(3), 785–824.

Pallais, A. (2009). Taking a chance on college: Is the Tennessee education lottery scholarship program a winner? *Journal of Human Resources*, *44*(1), 199–222.

Peña, P. A. (2020). Relative age and investment in human capital. *Economics of Education Review*, *78*, 102039.

Peña, P. A., & Duckworth, A. L. (2018). The effects of relative and absolute age in the measurement of grit from 9th to 12th grade. *Economics of Education Review*, *66*, 183–190.

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., & 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., & Winters, J. V. (2013). The effects of HOPE on post-college retention in the Georgia workforce. *Regional Science and Urban Economics*, *43*(3), 479-490. doi: https://doi.org/10.1016/j.regsciurbeco.2013.02.003

Sjoquist, D. L., & Winters, J. V. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, *80*, 39-50. doi: https://doi.org/10.1016/j.jue.2013.10.003

Takaku, R., & Yokoyama, I. (2021). What the COVID-19 school closure left in its wake: Evidence from a regression discontinuity analysis in Japan. *Journal of Public Economics*, *195*, 104364.

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

West Virginia Higher Education Policy Commission. (2012). *West Virginia financial aid comprehensive report.* https://www.wvhepc.org/downloads/LOCEA/FINANCIAL_2012_11-30-12.pdf.

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

Wherry, L. R., Miller, S., Kaestner, R., & Meyer, B. D. (2018). Childhood Medicaid coverage and later-life health care utilization. *Review of Economics and Statistics*, *100*(2), 287–302.

Winters, J. V. (2020). In-state college enrollment and later life location decisions. *Journal of Human Resources*, *55*(4), 1400–1426.

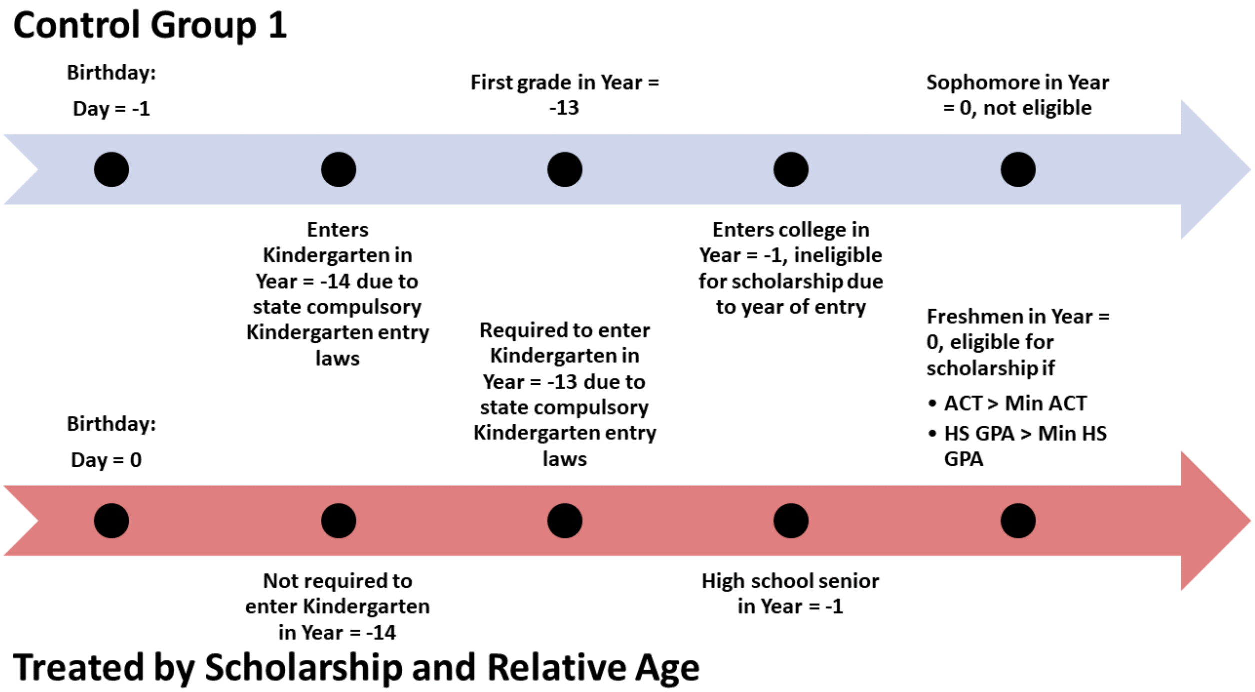
Zhang, L., Hu, S., Sun, L., & Pu, S. (2016). The effect of Florida's Bright Futures Program on college choice: A regression discontinuity approach. *The Journal of Higher Education*, *87*(1), 115–146.

Online Appendix

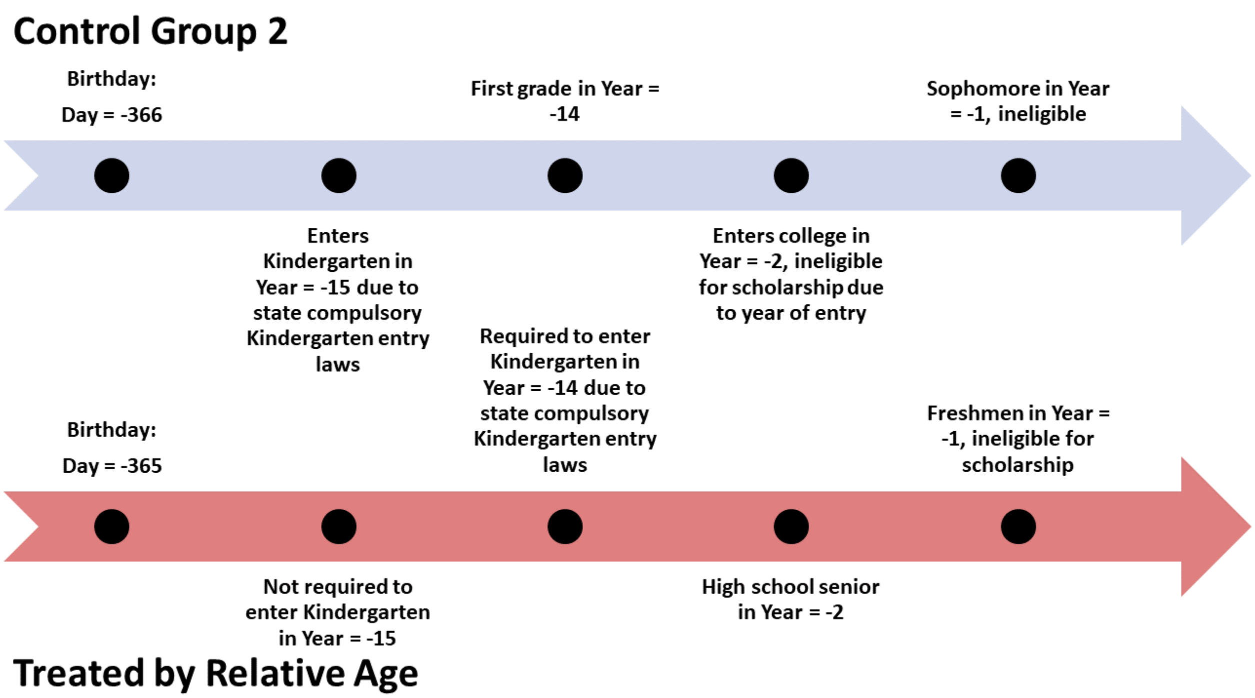
A Additional Tables and Figures

Figure A.I: College Scholarship Assignment By Birthdate Relative to Kindergarten Entry Cutoff

(A): 1983, Post-Scholarship Kindergarten Cutoff



(B): 1982, Pre-Scholarship Kindergarten Cutoff



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 kindergarten cohort that will enter college in the first college eligible cohort. Arrows for progression are valid for compliers who enter in assigned kindergarten cohort and finish 1 grade per year as defined by Equation 2B. See Section II for details.

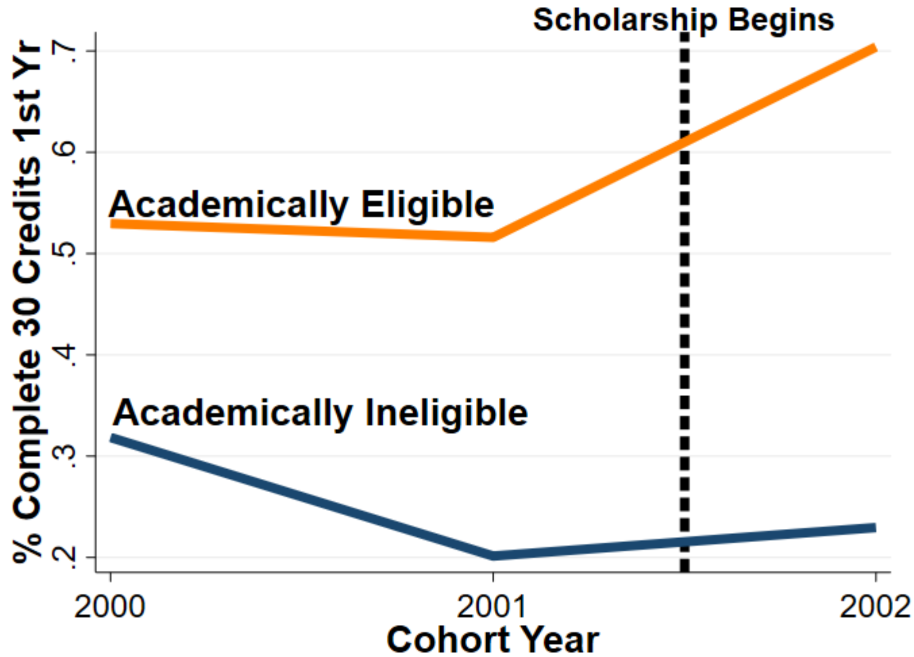
Table A.I: Descriptive Statistics By Academic and Temporal Eligiblity

	Mean	Median	Std Dev	Min	Max	Count
Panel A: Acad Elig., Born Day ≥ 0						
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						
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						
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						
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

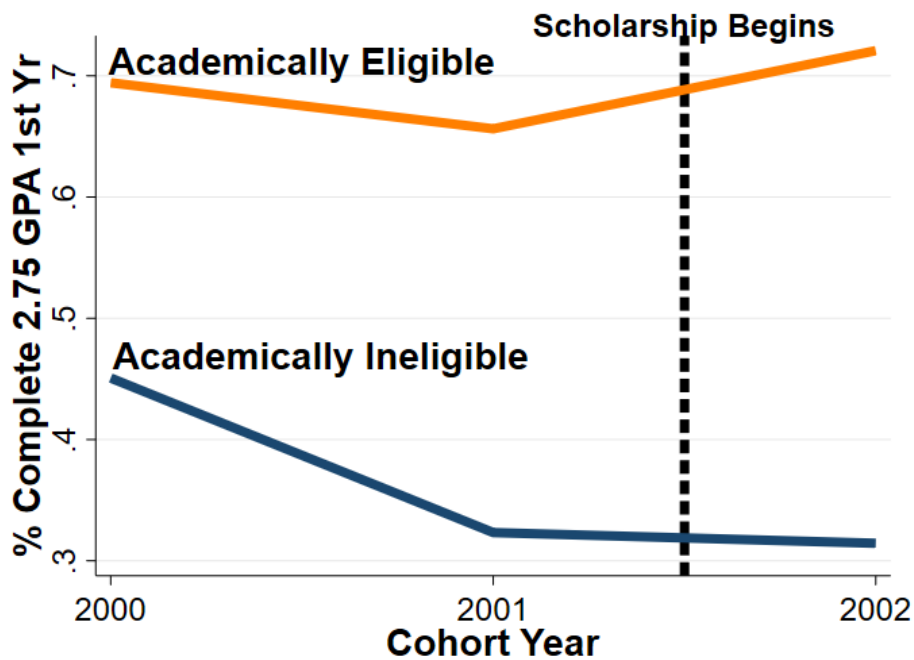
Note: Student-level observations. In-state students at WVU. First-time freshman only. Only cohorts 2001 and 2002. Freshman renewal GPA = 2.75. Freshman renewal credits = 30. Academic eligibility obstained by (HS GPA $>$ 3) AND (ACT Composite $>$ 20 OR SAT Composite $>$ 1000). Temporal eligibility if 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.II: 1st Year Renewal Likelihood by Cohort

(A): First Year Credit Renewal

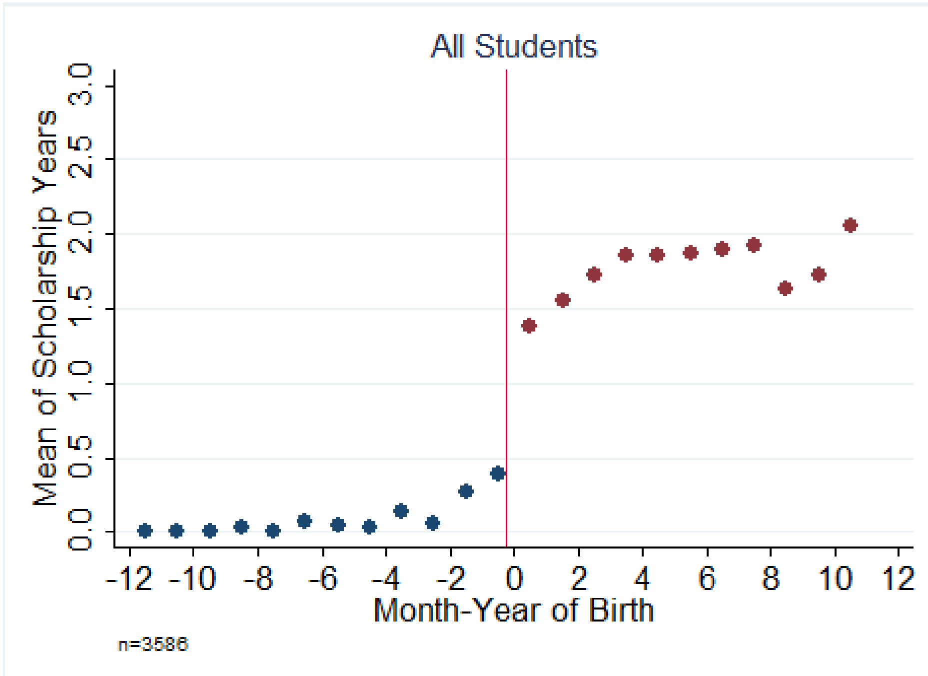


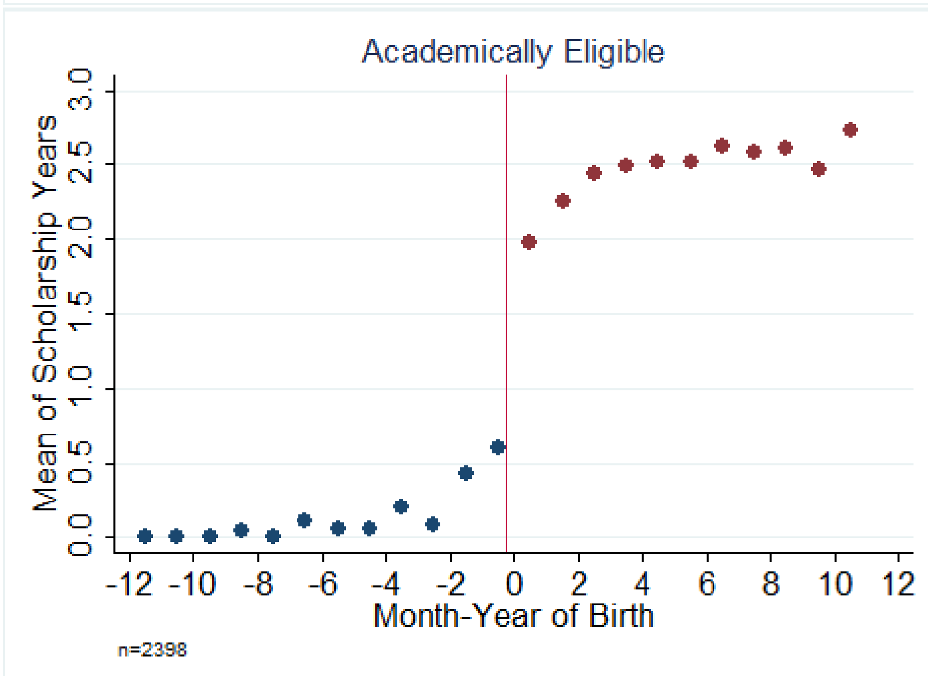
(B): First Year GPA Renewal



Note: The data are from university administrative records and cover in-state students beginning college in 2000-2002. A shows the percent of students who exceed the number of completed credits (30) required for scholarship renewal, once the scholarship begins, by cohort and academic eligibility. B shows the percent of students who exceed the GPA requirement required for scholarship renewal, once the scholarship begins, by cohort and academic eligibility.

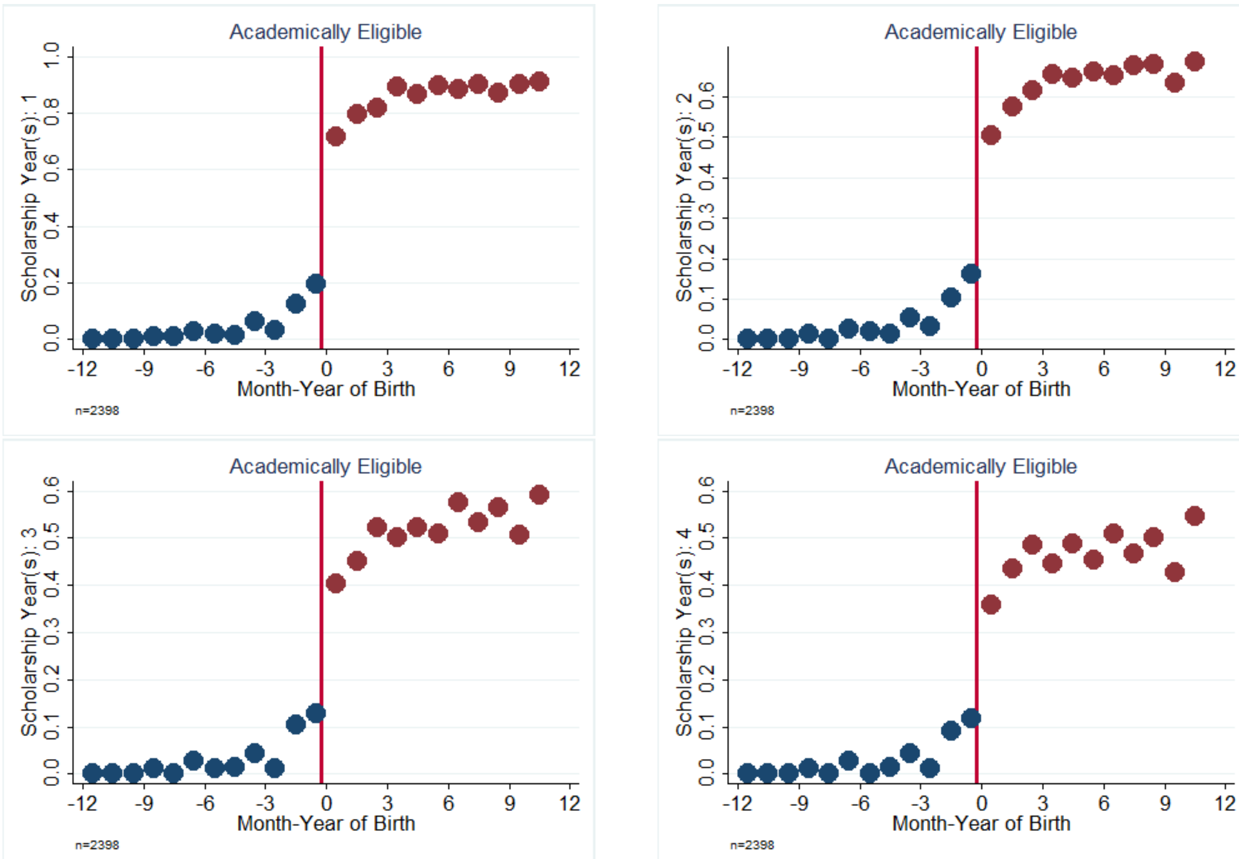
Figure A.III: Effect of Birth Month on Years of Scholarship





Note: The data are from university administrative records and cover in-state students beginning college in 2001-2002. These figures plot binned averages by re-centered birth month. The dependent variable is number of years received the Promise Scholarship. Panel A shows all students and Panel B shows only academically eligible students. This data did not include the exact birthdate.

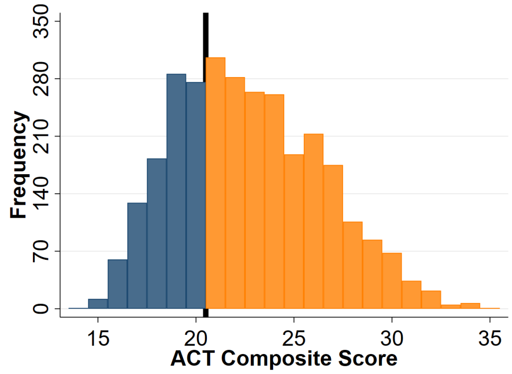
Figure A.IV: Relationship between Birth Date and Years of Scholarship Received, Academically Eligible



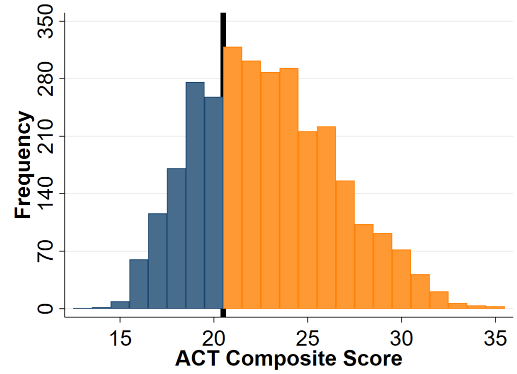
Note: The data are from university administrative records and cover in-state students beginning college in 2001-2002. These figures plot binned averages by re-centered birth month. The dependent variable is number of years received the Promise Scholarship. (a) is receive , at least, 1 year of the scholarship. (b) is receive , at least, 2 years of the scholarship. (c) is receive, at least, 3 years of the scholarship. (d) is receive 4 years of the scholarship. This data did not include the exact birthdate.

Figure A.V: Comparing Academic and Birthdate Running Variables' Continuity Around Cutoffs

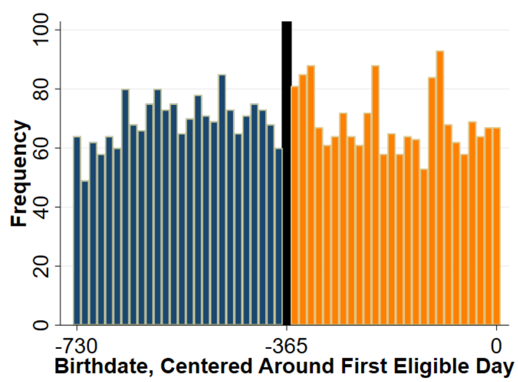
(A): ACT Scores, Two Cohorts Pre-Scholarship



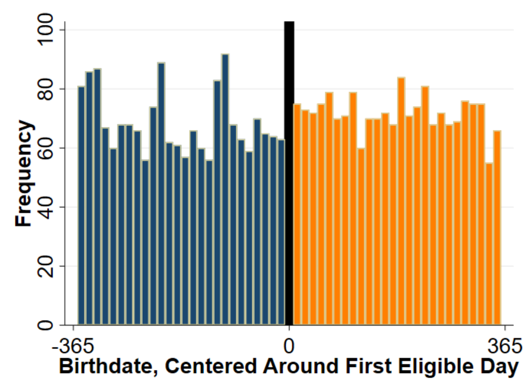
(B): ACT Scores, First Cohort of Scholarship



(C): Birthdate Density Around Cutoff, Before Scholarship



(D): Birthdate Density Around Cutoff, For Scholarship



Note: Panels A and B show the distribution of ACT Composite score. Panel A uses the 2000 and 2001 cohorts. Panel B uses the 2002 cohort. Birthdate 0 is the cutoff for the post-scholarship cutoff and birthdate -365 is the pre-scholarship cutoff. Panel C (n= 3,600) shows the distribution of those born within 1 year of the pre-scholarship cutoff date. Panel D (n= 3,666) shows the distribution of birthdate for those born within a year of the post-scholarship cutoff date. Each bar in Panels C and D represent 2 weeks of birthdates.

Table A.II: Pre-Treatment Covariate Balance, Both Cutoffs

Panel A	Dep. Var.: Male			Dep. Var.: HS GPA		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.03	-0.14**	0.17**	0.13**	0.07	0.06
	(0.06)	(0.06)	(0.08)	(0.05)	(0.05)	(0.07)
Bandwidth	150	154	160	190	166	160
Observations	1489	1569	3252	1868	1657	3231

Panel B	Dep. Var.: ACT Composite			Dep. Var.: SAT Combined		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	-0.05	-0.44	0.47	40.04	83.60**	-42.76
	(0.44)	(0.44)	(0.62)	(33.51)	(36.92)	(53.39)
Bandwidth	167	155	160	192	182	160
Observations	1362	1299	2661	340	324	591

Panel C	Dep. Var.: ACT English			Dep. Var.: ACT Math		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.00	0.05	-0.02	0.11	-0.73	0.90
	(0.55)	(0.53)	(0.72)	(0.48)	(0.52)	(0.77)
Bandwidth	147	146	160	198	173	160
Observations	1210	1211	2661	1616	1427	2661

Panel D	Dep. Var.: ACT Reading			Dep. Var.: ACT Science		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	-0.31	-0.35	0.09	0.00	-0.84*	0.91
	(0.60)	(0.58)	(0.83)	(0.41)	(0.49)	(0.64)
Bandwidth	179	154	160	208	137	160
Observations	1446	1282	2661	1683	1127	2661

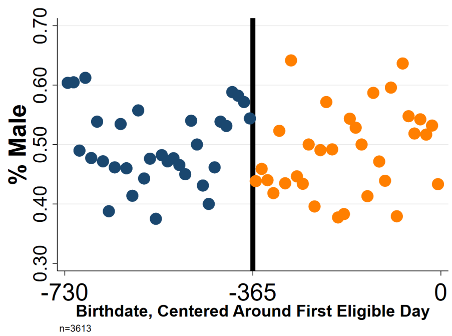
Panel E	Dep. Var.: Acad. Elig.			Dep. Var.: HS GPA Eligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.06	0.02	0.05	0.07*	0.04	0.03
	(0.05)	(0.05)	(0.07)	(0.04)	(0.04)	(0.06)
Bandwidth	180	168	160	184	195	160
Observations	1760	1685	3231	1803	1936	3231

Panel F	Dep. Var.: ACT Comp. Elig.			Dep. Var.: SAT Eligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.05	-0.08	0.15*	0.03	0.23**	-0.14
	(0.05)	(0.05)	(0.08)	(0.08)	(0.11)	(0.16)
Bandwidth	216	152	160	254	191	160
Observations	1748	1258	2661	449	337	591

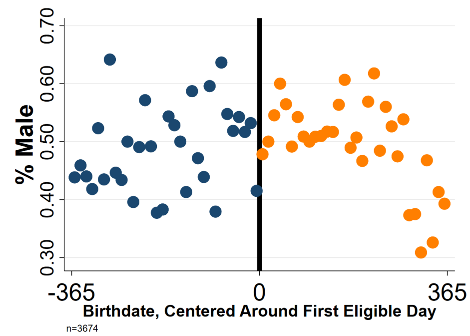
Note: * p<0.1, ** p<0.05, *** p<0.01. Standard errors shown in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. Coefficients are for $After82$ and $After83$ from local linear regressions in Equation 8. Columns 1 and 4 show $After83$, columns 2 and 5 show $After82$, columns 3 and 6 show $After83 - After82$. Triangular kernels and MSE-optimal bandwidths are used. The dependent variable is shown in titles above sets of columns. Students who enter in the 2000, 2001, or 2002 cohorts are included.

Figure A.VI: Test Scores, HS GPA, and Gender Balance at Cutoffs

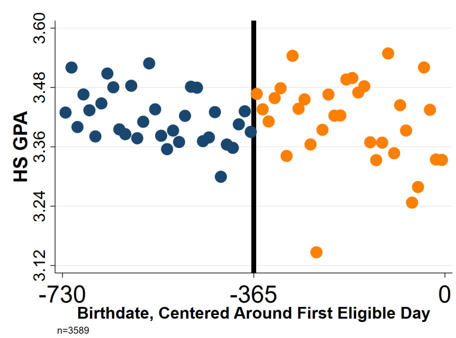
(A): Male, Pre-Scholarship



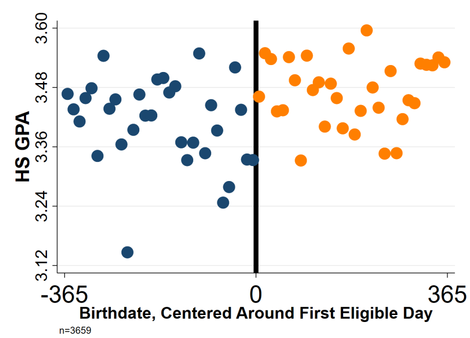
(B): Male, Post-Scholarship



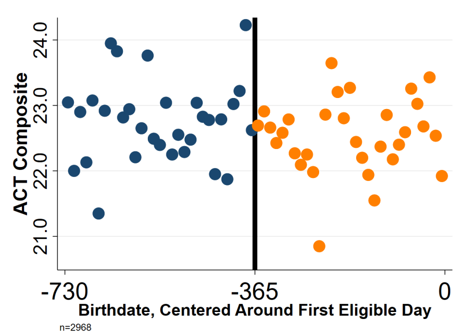
(C): HS GPA, Pre-Scholarship



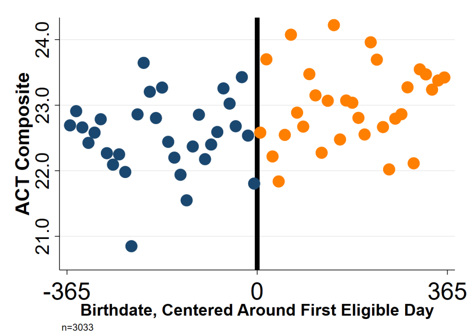
(D): HS GPA, Post-Scholarship



(E): ACT Composite, Pre-Scholarship



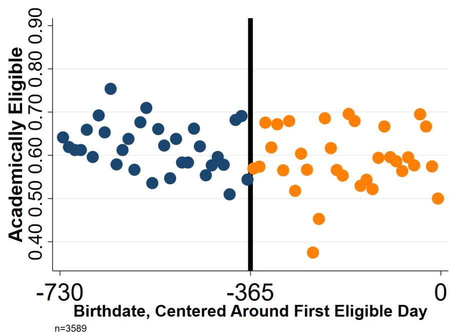
(F): ACT Composite, Post-Scholarship



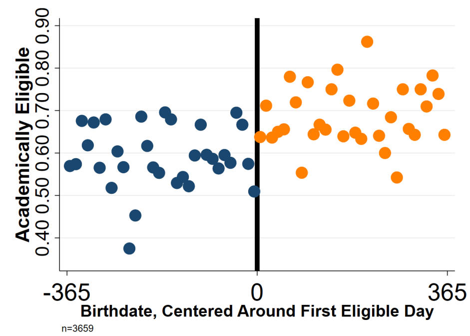
Note: The data are from university administrative records and cover in-state students born within 365 days of the pre-scholarship, 1982 cutoff (left side figures) and born within 365 days of the post-scholarship, 1983 birth cutoff (right side figures). These figures plot binned averages by re-centered birthdate. In Panels A and B the dependent variable is a binary variable equaling 1 if students are male. In Panels C and D the dependent variable is high school GPA. In Panels E and F the dependent variable is ACT composite score. The y-scales differ vertically but are the same horizontally. Table I and Table A.II show the related estimated coefficients for exceeding the cutoffs.

Figure A.VII: Effect of Birthdate Cutoffs on Academic Eligibility

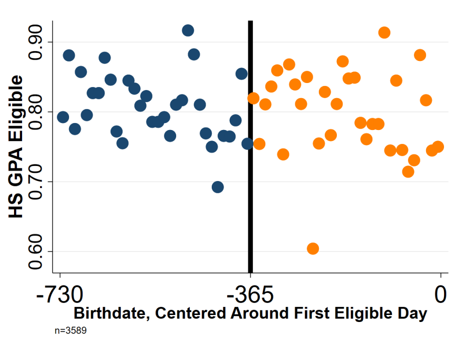
(A): Academic Eligibility, Pre-Scholarship



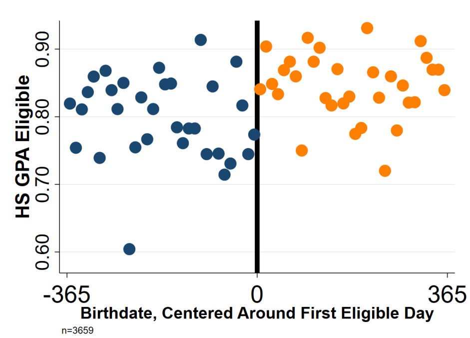
(B): Academic Eligibility, Post-Scholarship



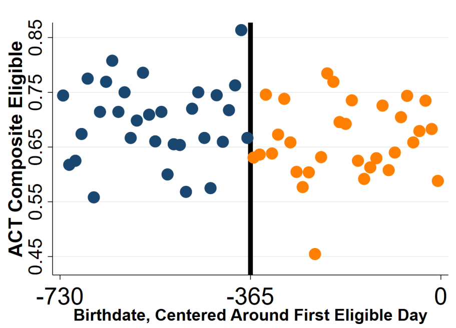
(C): HS GPA Eligibility, Pre-Scholarship



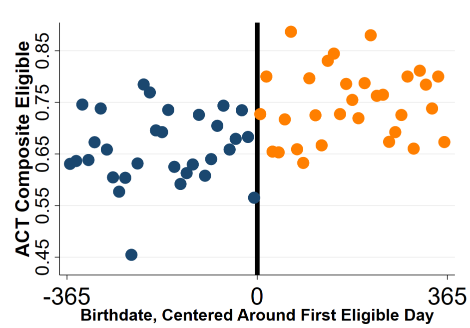
(D): HS GPA Eligibility, Post-Scholarship



(E): ACT Composite Eligibility, Pre-Scholarship



(F): ACT Composite Eligibility, Post-Scholarship



Note: The data are from university administrative records and cover in-state students born within 365 days of the pre-scholarship, 1982 cutoff (left side figures) and born within 365 days of the post-scholarship, 1983 birth cutoff (right side figures). These figures plot binned averages by re-centered birthdate. In Panels A and B the dependent variable is a binary variable equaling 1 if students exceeded both the HS GPA (3.0) and entry test (ACT $\geq$ 21 or SAT $\geq$ 1000) score requirements to be eligible for the Scholarship. This variable is described in Section III.1. In Panels C and D the dependent variable a binary variable equaling 1 if students exceeded the HS GPA requirement to be eligible for the Scholarship.In Panels E and F the dependent variable is ACT composite score. The y-scales differ vertically but are the same horizontally. Table I and Table A.II show the related estimated coefficients for exceeding the cutoffs.

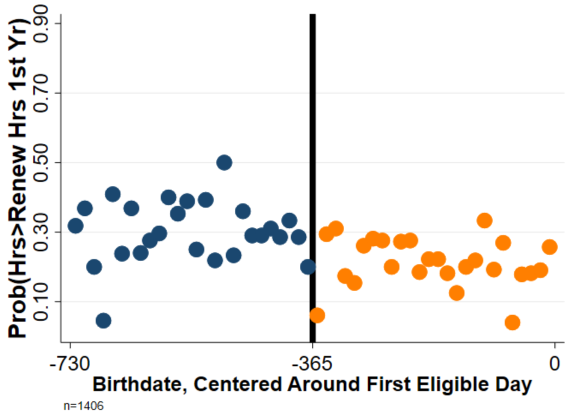
Table A.III: Intent to Treat Effect on Meeting Academic Renewal Threshold By Academic Eligibility

Panel A: Credits, Conventional						
	Academically Eligible			Academically Ineligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.135*	0.011	0.129	0.055	-0.070	0.104
	(0.070)	(0.070)	(0.097)	(0.076)	(0.076)	(0.102)
Bandwidth	161	153	160	168	128	160
Observations	1021	926	1995	599	512	1236
Panel B: Credits, Bias-Corrected and Robust SE's						
	Academically Eligible			Academically Ineligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.122	0.019	0.114	0.051	-0.081	0.087
	(0.084)	(0.084)	(0.148)	(0.091)	(0.089)	(0.143)
Bandwidth	161	153	160	168	128	160
Observations	1021	926	1995	599	512	1236
Panel C: GPA, Conventional						
	Academically Eligible			Academically Ineligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.100*	0.006	0.127	0.049	0.036	0.001
	(0.057)	(0.063)	(0.092)	(0.085)	(0.084)	(0.120)
Bandwidth	231	151	160	180	141	160
Observations	1490	909	1995	629	564	1236
Panel D: GPA, Bias-Corrected and Robust SE's						
	Academically Eligible			Academically Ineligible		
	(1) Post	(2) Pre	(3) Diff	(4) Post	(5) Pre	(6) Diff
After Cutoff	0.116*	0.006	0.196	0.038	0.064	-0.048
	(0.067)	(0.075)	(0.133)	(0.102)	(0.096)	(0.170)
Bandwidth	231	151	160	180	141	160
Observations	1490	909	1995	629	564	1236

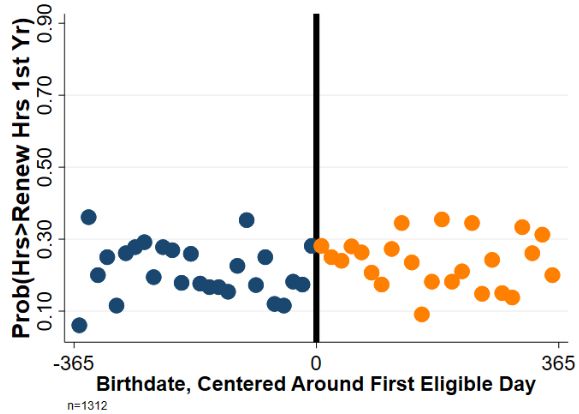
Note: * p<0.1, ** p<0.05, *** p<0.01. Standard errors shown in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. Coefficients are for $After82$ and $After83$ from local linear regressions in Equation 8. Columns 1 and 4 show $After83$, columns 2 and 5 show $After82$, columns 3 and 6 show $After83 - After82$. Triangular kernels and MSE-optimal bandwidths are used. The dependent variable is shown in titles above sets of columns. Students who enter in the 2000, 2001, or 2002 cohorts are included. Plots can be seen in Figures V, VI, and A.VIII.

Figure A.VIII: Reduced Form Effect of Birthdate on Likelihood of Exceeding Renewal Thresholds in Freshman Year, Academically Ineligible Students

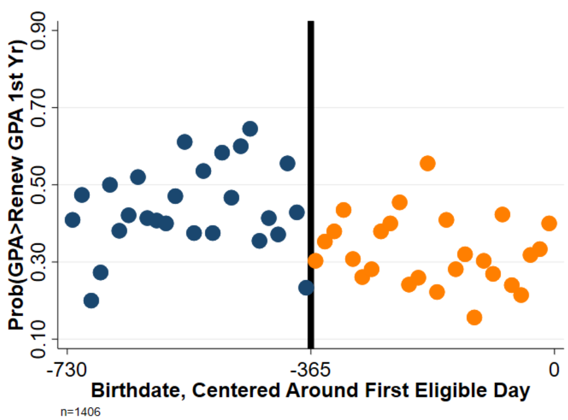
(A): Exceed Renewal Credits, Pre-Scholarship Cutoff



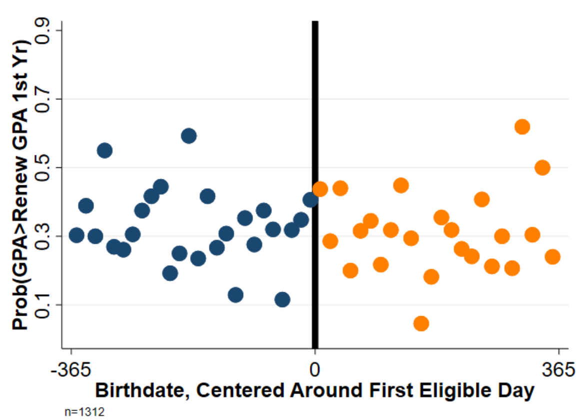
(B): Exceed Renewal Credits, Scholarship Cutoff



(C): Exceed Renewal GPA, Pre-Scholarship Cutoff



(D): Exceed Renewal GPA, Scholarship Cutoff



Note: The data are from university administrative records and cover academically ineligible, in-state students. Left side figures cover students born within 365 days of the pre-scholarship, 1982 birth cutoff. Right side figures show students born within 365 days of the post-scholarship, 1983 birth cutoff. These figures plot binned averages by re-centered birthdate. In Panels A and B the dependent variable is whether a student reached the academic renewal threshold for credits (30) in their first year of college. In Panels C and D the dependent variable is whether a student reached the academic renewal threshold for GPA (2.75) in their first year of college. Table A.III presents estimates of the differences in averages at the cutoff.

Table A.IV: Local Linear Estimates of Scholarship on Meeting First Year Credits and GPA Renewal Thresholds with Donut

Panel A: Dep. Var.: Renew Credits			
	(1)	(2)	(3)
Scholarship	0.20	0.23	0.34
	(0.20)	(0.24)	(0.27)
Dount Size	7	14	21
Bandwidth	126	113	113
Effective Obs. Left	357	294	280
Effective Obs. Right	412	350	323
Panel B: Dep. Var.: Renew GPA			
	(1)	(2)	(3)
Scholarship	0.42*	0.39	0.49
	(0.23)	(0.27)	(0.31)
Dount Size	7	14	21
Bandwidth	110	101	103
Effective Obs. Left	299	245	232
Effective Obs. Right	364	313	288

Note: * p<0.1, ** p<0.05, *** p<0.01. Standard errors shown in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. The estimate is the effect of receiving the Promise Scholarship. The two estimating equations are Equation 7 and Equation 8. All columns use the bias-correction and standard error adjustment in Calonico et al. (2017) with a local quadratic equation for bias-correction and the standard errors are adjusted for the bias-correction. Under the assumptions outlined in Section IV.1.2, the estimate shown ($\tau = \frac{\beta}{\psi}$) can be interpreted as a LATE estimand for compliers. The equation models $f(Birthdate)$ as local linear regression. All columns use MSE-optimal bandwidths and a triangular kernel. In all panels, the birthdate cutoff has a large, statistically significant effect on scholarship receipt. In Panel A the dependent variable is a dummy for finishes 30 credits (the renewal requirement) during freshman year. In Panel B the dependent variable is a dummy for exceeds a 2.75 GPA (the renewal requirement) during freshman year. Column 1 drops individuals born within 7 days of the cutoff on both sides of the cutoff. Column 2 drops individuals born within 14 days of the cutoff on both sides of the cutoff. Column 3 drops individuals born within 21 days of the cutoff on both sides.

Table A.V: Scholarship on Freshman Academic Renewal Threshold Completion, Parametric

	(1) +/- 365 Days	(2) +/- 180 Days
Panel A: Credits, Linear		
	(1)	(2)
At Least 1 Year Scholarship	0.17***	0.27***
	(0.03)	(0.05)
Observations	2437	1140
Kleibergen-Papp F	4273.3	1646.5
Panel B: Credits, Quadratic		
	(1)	(2)
At Least 1 Year Scholarship	0.21***	0.26***
	(0.04)	(0.05)
Observations	2437	1140
Kleibergen-Papp F	2605.5	1924.4
Panel C: GPA, Linear		
	(1)	(2)
At Least 1 Year Scholarship	0.04	0.12**
	(0.03)	(0.05)
Observations	2437	1140
Kleibergen-Papp F	4273.3	1646.5
Panel D: GPA, Quadratic		
	(1)	(2)
At Least 1 Year Scholarship	0.06*	0.12**
	(0.04)	(0.05)
Observations	2437	1140
Kleibergen-Papp F	2605.5	1924.4

Note: * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Standard errors are robust to heteroskedasticity. Coefficients are from 2SLS regressions, representing the difference at the birthdate cutoff when a functional form for birthdate is assumed globally. In Panel A, exceeding the completed credits (30) in freshman year required for renewal is the dependent variable and linear regressions are fit on either side of the cutoff. In Panel B, the same dependent variable is used as A, but the functional form of birthdate is assumed to be quadratic. In Panel C, the dependent variable is exceeding the freshman GPA (2.75) required to renew the scholarship and linear regressions are fit on either side of the birthdate cutoff. Panel D uses the same dependent variable as Panel C and assumes the functional form of birthdate is quadratic. Finally, column 1 uses 365 days on either side of the cutoffs. Column 2 uses 180 days on either side of the cutoffs.

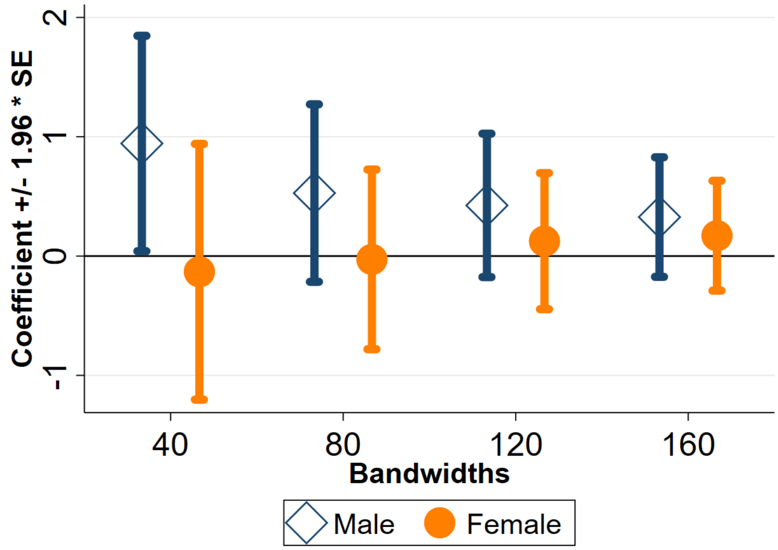
Table A.VI: Local Linear Estimates of Scholarship on First Year Credits and GPA

Panel A: 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
Panel B: 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

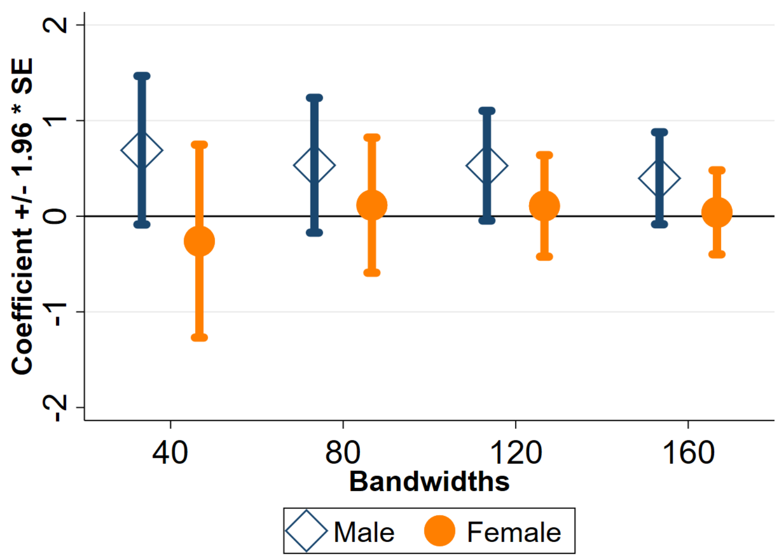
Note: * p<0.1, ** p<0.05, *** p<0.01. The standard errors are shown in parentheses and are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 3 neighbors. The estimate is the effect of receiving the Promise Scholarship. The two estimating equations are Equation 7 and Equation 8. Under the assumptions outlined in Section IV.1.2, the estimate shown ($\tau = \frac{\beta_{After83}}{\psi_{After83}}$) can be interpreted as a LATE estimand for compliers. The equation models $f(Birthdate)$ as local linear regression. All columns use MSE-optimal bandwidths and a triangular kernel. In all panels, the birthdate cutoff has a large, statistically significant effect on scholarship receipt. In Panel A the dependent variable is continuous credits completed in freshman year. In Panel B the dependent variable is continuous GPA freshman year. In columns 1 and 2, the full sample is used. From column 2 onwards, the bias-correction and standard error adjustment in Calonico et al. (2017) are used with a local quadratic equation for bias-correction and the standard errors are adjusted for the bias-correction. In column 3, only students who exceed the academic eligibility cutoffs are included.

Figure A.IX: Difference in Discontinuity Estimates by Gender

(A): Credit Hours



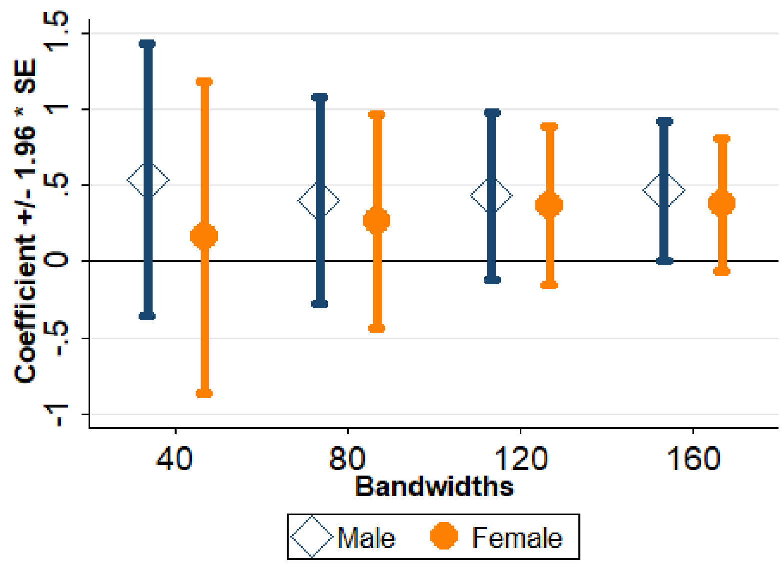
(B): GPA



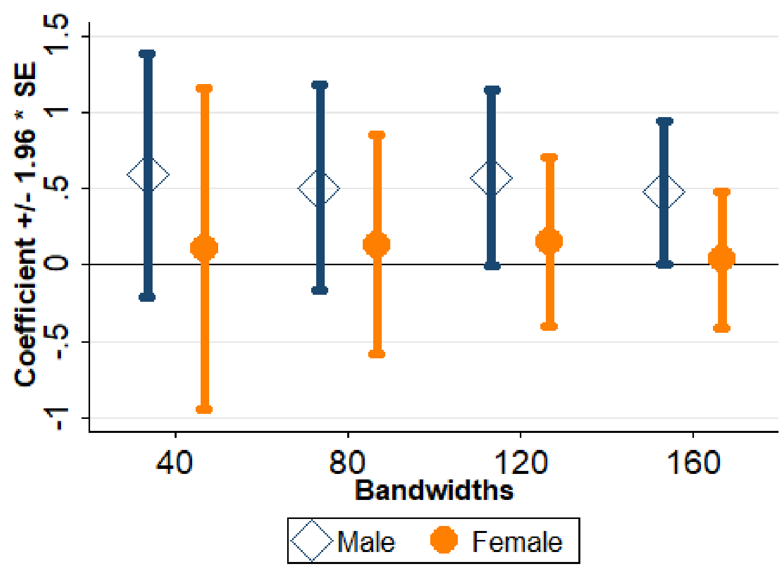
Note: Coefficients represent estimated effect of scholarship from difference-in-discontinuity (Grembi et al., 2016). The point estimates equal $\frac{\beta_{After83}}{\psi_{After83}} - \beta_{After82}$. The corresponding equations are shown in Equation 7 and Equation 8. Whiskers represent 95% confidence intervals. Standard errors are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 6 neighbors. Triangular kernels used. The different bandwidths indicate the number of days used from the respective cutoffs. Same bandwidth imposed at both cutoffs. In Panel A the dependent variable is exceed renewal credits in freshman year. In Panel B the dependent variable is exceed renewal GPA in freshman year. See Section III for details. All regressions subsample to academically eligible students . Blue, open diamonds represent men and orange, full circles represent women.

Figure A.X: Difference in Discontinuity Estimates by Gender, 3 Year Outcomes

(A): Credit Hours



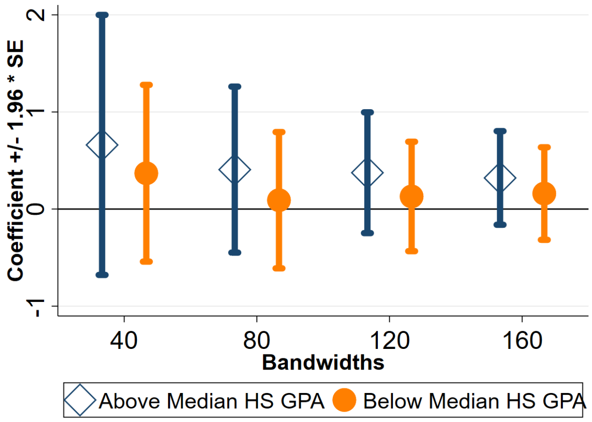
(B): GPA



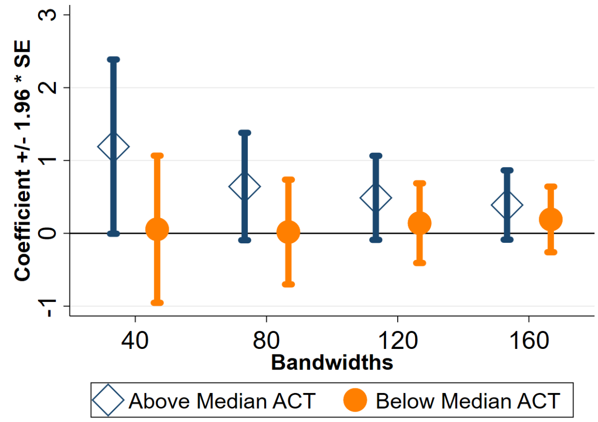
Note: Coefficients represent estimated effect of scholarship from difference-in-discontinuity (Grembi et al., 2016). The point estimates equal $\frac{\beta_{After83}}{\psi_{After83}} - \beta_{After82}$. The corresponding equations are shown in Equation 7 and Equation 8. Whiskers represent 95% confidence intervals. Standard errors are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 6 neighbors. Triangular kernels used. The different bandwidths indicate the number of days used from the respective cutoffs. Same bandwidth imposed at both cutoffs. In Panel A the dependent variable is a binary variable that equals 1 if students exceed renewal credits in freshman, sophomore, and junior years. In Panel B the dependent variable is binary variable that equals 1 if students exceed renewal GPA in freshman, sophomore, and junior years. See Section III for details. All regressions subsample to academically eligible students. Blue, open diamonds represent men and orange, full circles represent women.

Figure A.XI: Difference in Discontinuity Estimates by Academic Ability, Academically Eligible 3

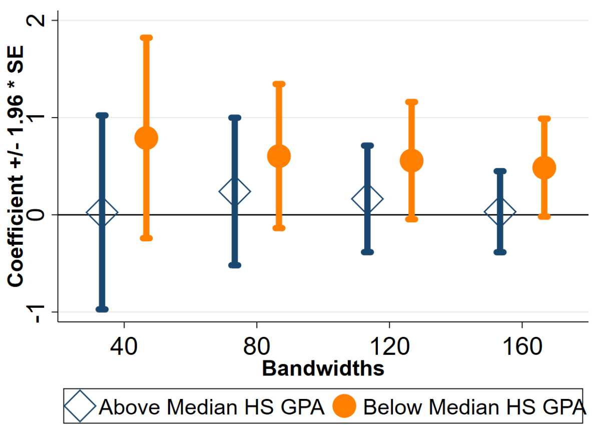
(A): Credits Target, Split by Median HS GPA



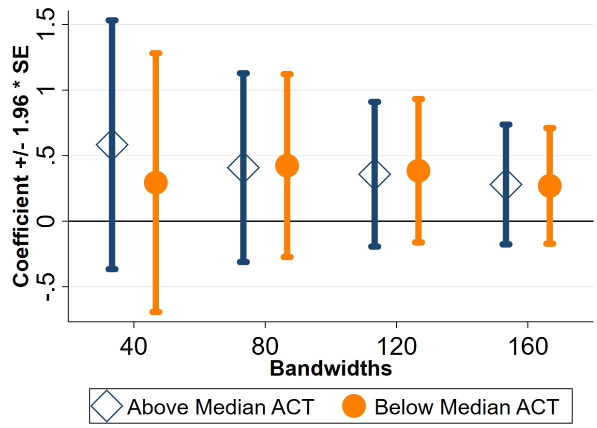
(B): Credits Target, Split by Median ACT



(C): GPA Target, Split by Median HS GPA



(D): GPA Target, Split by Median ACT



Note: Coefficients represent estimated effect of scholarship from difference-in-discontinuity (Grembi et al., 2016). The point estimates equal $\frac{\beta_{After83}}{\psi_{After83}} - \beta_{After82}$. The corresponding equations are shown in Equation 7 and Equation 8. Whiskers represent 95% confidence intervals. Standard errors are from a heteroskedasticity-robust nearest neighbor variance estimator with a minimum of 6 neighbors. Triangular kernels used. The different bandwidths indicate the number of days used from the respective cutoffs. Same bandwidth imposed at both cutoffs. In Panel A and B the dependent variable is a binary variable that equals 1 if students exceed renewal credits in freshman year. In Panel C and D the dependent variable is binary variable that equals 1 if students exceed renewal GPA in freshman year. See Section III for details. All regressions subsample to academically eligible students. In Panel A and C, blue, open diamonds represent student who students who do better than the, conditional on being academically eligible, median HS GPA (3.71) and orange, filled in dots represent those with lower than median HS GPA. In Panel B and D, blue, open diamonds represent student who students who do better than the, conditional on being academically eligible, median ACT composite score (26) and orange, filled in dots represent those with lower than median composite ACT.

Table A.VII: Compliance Types, Academically Eligible and Ineligible

Panel A: Participation	
	Estimate
Compliers	0.339
Always-Takers	0.105
Never-Takers	0.555

Note: This table shows the compliance groups when academically ineligible students are included. To see the compliance groups without academically ineligible students, see Table IV.

B 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.

There was change in the entry date set by West Virginia's laws over time. In 1982, the rule changed to students who were 5 by September 1 of the current school year were supposed to enter kindergarten. Additionally, in 1988 the law was changed again which means that students in the year after are sorted at a different date cutoff, November 1. Thus, the 1988/1989 beginning cohorts are the last time that September 1 is the cutoff.

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).[30] (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. Furthermore, West Virginia compulsory age of kindergarten entry does not apply to students that enter kindergarten in another state, then move into West Virginia.

B.1 Other Cohort-Based Policies

A threat to identification is that there is some other confounding policy that begins for children born just beyond the cutoff point (i.e. on or after September 1, 1983). This could be a federal government policy or some other specific policy that affects children in the 1989 kindergarten cohort in West Virginia. To investigate these possibilities, I investigate age/birthdate requirements for eligibility for all programs considered in Hendren and Sprung-Keyser (2020).

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

Policies from a Unified Welfare Analysis (Hendren & Sprung-Keyser, 2020) There are 133 government policies that are compared in Hendren and Sprung-Keyser (2020). These policies are searched for policies that change discontinuously for students on either side of the birth cutoff, September 1, 1983, for entering Kindergarten in 1989 in West Virginia. While many policies affect students in the 1980's, most do not discontinuously affect those who are born on opposite sides of the cutoff.

Throughout the 1980's there was significant expansions of federal health insurance programs, particularly Medicaid. One policy change was the extension of Medicaid to low-income children who were born after September 30, 1983. This expansion increased education and health care utilization (Wherry, Miller, Kaestner, & Meyer, 2018); however, this eligibility only changes discontinuously for students who are born 30 days from the cutoff (Cohodes, Grossman, Kleiner, & Lovenheim, 2016). Second, there was an extension of Medicaid to pregnant mothers which raised high school completion for babies who were in utero at the time (Miller & Wherry, 2019). Nevertheless, while this program changed in the 1980's, it did not affect those on either side of the 9/1/1983 birthdate cutoff differently. This is true for all other health insurance expansions at this time.