Fall 12-2021

AN EXPLORATION OF UNINTENDED EFFECTS OF EDUCATION POLICIES IN THE UNITED STATES

Kevin Estes
University of New Mexico - Main Campus

Kevin Estes
Candidate

Economics
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Xiaoxue Li, Co-Chair

Melissa Binder, Co-Chair

Robert Berrens

Becky Kilburn

Sergio Ascencio

AN EXPLORATION OF UNINTENDED EFFECTS OF EDUCATION POLICIES IN THE UNITED STATES

BY

KEVIN ESTES

B.B.A, Finance, New Mexico State University, 2013
M.A., Economics, University of New Mexico, 2017

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Economics

The University of New Mexico
Albuquerque, New Mexico

December, 2021

DEDICATION

This work is dedicated to my family: Lainey, Mayvin, and Trevin.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Melissa Binder and Dr. Xiaoxue Li, who have guided me throughout my time in the graduate program. My development as a researcher can be mostly attributed to the time I spent in your classes and your careful feedback on classwork, research ideas, and manuscripts.

I also thank my committee members, Dr. Robert Berrens and Dr. Becky Kilburn, for their guidance, confidence, and mentorship in important research projects during my time in graduate school. The projects I was able to work on with both of you are among my fondest memories of graduate school. I also would like to thank Chris Doss, Andie Phillips, and the IDI team for their guidance and assistance on these two projects.

I am also very thankful to Dr. Sergio Ascencio for serving as the external member of my committee. I would not have completed in Fall 2021 without his kindness and willingness to serve as external member on short notice.

THREE ESSAYS ON BARRIERS TO HUMAN CAPITAL ACQUISITION

by

Kevin Estes

B.B.A, Finance, New Mexico State University, 2013
M.A., Economics, University of New Mexico, 2017
PhD, Economics, University of New Mexico, 2021

ABSTRACT

The work in this dissertation examines unintended consequences from various public policies in education. The first policy examined is the adoption of the four-day school week schedule by public-school districts across the United States. Concerns over the additional weekend night for students are explored by examining teen traffic safety within the school district before and after adoption of the four-day schedule. The second policy examined is the usage of academic probation by universities. Student responses to being placed on academic probation vary, and financial implications for the student are a potential mechanism behind these responses. Student responses segmented by Pell status is explored throughout this chapter. The final policy examined is the introduction of charter schools throughout the United States. The effect of charter school competition on public-school district finances is explored and important differences are found depending on who establishes the charter school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 – Introduction: Education Policy in the United States and Unintended Consequences

Most people around the world, and particularly in more affluent countries, spend the first quarter of their lives acquiring human capital. Human capital increases the wellbeing of a society and the wellbeing of the individual, thus it is both subsidized by governments and pursued independently in many cases. Individuals in the pursuit of human capital face a number of challenges, or barriers, throughout their education years. Public policy around education thus attempts to address these barriers, whether the barriers are financial (in the case of Pell grants for college-going students who demonstrate need), geographical (additional funding for students in hard to reach, rural areas), or even in how education is delivered (student teacher ratios, common education standards for school districts, etc.). Despite the good intentions of many policies, some do not have their intended effects. This work uses multiple empirical strategies to identify the causal impact of some of these unintended consequences of policies that could lead to additional challenges for students attempting to pursue an education.

Although the history of public education in the United States extends back to the first national system of education developed in the 19th century, unintended consequences of education policy dates back further. For nearly 200 years before that national system local communities and religious groups were the primary forces behind education. With no oversight, this led to "public" education at the time being highly segregated by wealth and race (Thattai 2001). This unfortunate consequence, from a lack of regulation, would not change for many years, and even today is still a challenge across the United States despite the increased involvement of government. This involvement has led to the public education system today bearing little resemblance to that early system. For example, attendance is now compulsory for all school-age children (i.e. required) in most states till at least the age of 16 (Bush 2010). This policy, with the intended effect of increasing education within society, also had an unintended consequence. Older students (those born in the first quarter of the year) could drop out at lower levels of attained education than younger students in the same grade. This small difference in age is then tied to significantly less future earnings for these older students (Angrist & Krueger 1991).

This pattern of regulation, response to regulation, and unintended effects continues today. In chapter 2 the adoption of a four-day school week policy, and its unintended effect on teen traffic safety, is examined. The four-day school week arose out of a need by small, rural school districts to reduce transportation costs during a spike in global oil prices. Once the schedule was adopted by some school districts, communication between education leaders from other communities facilitated the spread of the reduced schedule to other school districts. Today school districts in at least 24 states have adopted the schedule, and its usage is expanding to larger urban school districts. To ensure the same quantity of education between four-day and five-day school districts, most states have implemented a requirement on the minimum number of school-year instructional hours. Four-day school week districts thus go longer each day. This small change has a large impact on the daily driving patterns of students in these districts.

Using a detailed national dataset of fatal traffic accidents over an eight-year period, I examine how the adoption of a four-day school week affects driving age teen traffic safety in nine U.S. states (these states provided longitudinal information on public school district schedules). With longer school days where driving occurs at different hours, and an additional “weekend” night, for four-day school week students the question becomes empirical. To answer this question I use the latest difference-in-differences methods (Callaway & Sant’Anna’s (2020) doubly robust estimator) to estimate the causal impact of the four-day school week on teen traffic safety. Results do not show a significant change in fatal accident involvement for teens after a district adopts a four-day school week. In sub-analyses these findings are confirmed (with some evidence suggesting female driving age teens benefit from the policy). Concerns over the increased risk of an extra weekend night do not seem to materialize across all analyses, including a days of the week analysis.

Chapter 3 turns to an example of education policy in higher education, academic probation. Academic probation is a notification from a university that is given to students following a semester where their cumulative academic performance was below a GPA standard upheld by the university. This policy has the intended effect of focusing students on their academic studies but has been shown in some settings to have the unintended effect of causing students to leave the university. To understand this unintended consequence, and confirm its presence, I examine academic probation by a student's financial aid status. Ignoring financial aid implications of academic probation makes interpretation of why students respond to academic probation incomplete. I use a regression discontinuity design to examine Pell and non-Pell students' responses to being placed on academic probation. Due to satisfactory academic progress (SAP) requirements of the Pell grant, Pell students have additional financial consequences of being placed on academic probation (which has the same GPA requirement as SAP for upperclassmen, but a different GPA cutoff for freshmen). I find that despite the financial implications for freshmen Pell students, both non-Pell and Pell freshmen respond to academic probation by increasing their next semester GPA by nearly identical amounts of 0.2 GPA units. Upperclassmen do not respond to academic probation. A sub-analysis examining eventual graduation shows non-Pell student graduation rates (four, five, and six years) are negatively affected by the academic probation GPA cutoff.

Chapter 4 returns to pre-postsecondary education and considers how public district school finances can be affected by charter school competition. Charter schools have a history of providing students with an alternative option for education from their assigned public school. While additional options theoretically improve the wellbeing of a student, it could come at the cost of reduced resources for all students when redirected from public schools. An additional consideration on how charter school competition affects public schools is who establishes the charter school. Charter schools can be established by traditional public schools or by other parties (i.e. the state education department or a non-profit organization). This distinction is important, because if established by an outside party the charter school is a direct competitor for students within the geographic area. Direct competition could lead school districts to redirect resources to categories that students and parents more highly value if the threat of disenrollment exists. This analysis investigates the impact of the implementation of charter schools on public-school district finances across the United States for both kinds of charters, charters established by the local school district and from charter schools established within the same county by another group. Using difference-in-differences and event study methods I find increases enrollment the years following a charter being established by the district. Charter schools in competition with public-school districts results in a decrease in staff employed by the district and financial assets per student, but do not otherwise affect school district finances. Additional evidence is presented that suggests that as more students within a county are enrolled in a competing charter the number of staff employed by the district declines.

In chapter 5, the main conclusions from each chapter are presented. Contributions of each analysis, along with policy prescriptions, limitations, and the research frontier are also discussed.

The summary point of this dissertation is that unintended consequences of policies in education are neither new, nor unique to only specific policies. Each choice made by a policymaker when crafting policy to address problems in the education system comes with the cost of unintended effects. Research in this dissertation has mostly identified potential unintended effects that did not materialize (increased teen traffic risk from four-day school week policy, disenrollment effects from academic probation, impact on public school finances from charter competition). The absence of an unintended effect is also important to understand, and be documented, when policymakers are directly addressing concerns like these from their constituents. Research into education policy can also uncover positive unintended effects, as in the case of potentially improving the traffic safety of driving age female teens in four-day school week districts or showing that financial penalties might not be driving increased GPA performance in academic probation. These types of findings can motivate better future education policy (with hopefully smaller unintended effects).

Chapter 2 - Is Thursday the new Friday? The Four-Day School Week and Teen Traffic Safety

2.1 Introduction

Going to school five days a week is no longer the norm for some students in the United States. Increasingly over the past two decades small, rural school districts in 24 states have adopted a four-day school week (NCSL 2020; Thompson 2021). This simply means that instruction occurs within four days in a week instead of the traditional five; with the trade-off often being that each of those four days of the week are lengthened to make up the lost instructional time and to satisfy state instructional hours mandates. While districts often adopt the schedule for financial reasons (NCSL 2020), this inadvertently creates an additional weekend night for students. Concerns thus have been raised by teachers, administrators, policymakers, and parents due to these changes in students' schedules. Exacerbated weekend learning loss, food insecurity issues over the longer weekend, and elementary students mentally drained from the longer days are just a few of these concerns. The effect of the four-day school week schedule even extends to other activities. For high school students, a longer school day keeps students at school even later into the evening hours if they choose to participate in extra-curricular activities (such as athletic teams or clubs) which typically meet after school for practice, meetings, and competitions.

The four-day school week also has the indirect effect of changing the traffic patterns of students. The shift in schedule increases the amount of night time driving for extra-curricular active high school students, which is considerably more dangerous than daytime driving (Rice, Peek-Asa, & Kraus 2003). The additional "weekend" night added by the four-day school week schedule is also a potential concern. Weekends and nights are the most dangerous times for teens to drive (Doherty, Andrey, & MacGregor 1998). The dangers with weekends are most often associated with risky driving behaviors, such as drinking and speeding, that are influenced by extended curfews on non-school nights, especially when other peer passengers are present (Doherty, Andrey, & MacGregor 1998).

This chapter investigates the impact of four-day school week adoption by a school district on teen traffic safety. Data from the Fatality Accident Reporting System (FARS), which is produced by the National Highway Traffic Safety Administration, is used with school district schedule information gathered from eight state education departments. Additional information is collected from the National Center for Education Statistics to create an eight-year school district-level panel. A difference-in-differences approach using the Callaway & Sant'Anna (2020) doubly robust estimator is used, which while accounting for differential treatment timing in this setting, also allows for reporting of multiple treatment types (including overall treatment effect and dynamic treatment effects). I find no evidence that adopting a four-day school week schedule affects driving age teens' involvement in fatal accidents. Additional analyses investigating the day of the week the accident occurred, the time of day, and the gender of the individual involved also do not indicate increased risks for driving age teens. As a robustness check an analysis of summer months only, which are unaffected by school schedule, is done. A final analysis of all traffic accidents within a county where a school district adopts a four-day school week also does not indicate increased traffic accident risk from the four-day schedule.

The remainder of the paper is organized as follows: Section 2 discusses the literature on teen traffic fatalities and four-day school week policy. Section 3 provides data sources and descriptive statistics. Section 4 reviews the empirical method employed. Section 5 presents results from the analyses and robustness checks. Section 6 provides a discussion and concluding remarks.

2.2 Teen Traffic Fatalities and Four-Day School Week Policy

Past research into teen traffic safety has shown the success of public policy in reducing the total number of teen traffic fatalities over the past 30 years. It has been well established that night time driving and alcohol involvement, particularly for males, are leading contributors to teen traffic fatalities (Keall, Frith, & Preston 2005; Shope & Bingham 2008). In particular, increasing the minimum legal drinking age (MLDA) from 18 years of age to 21 years of age has been shown to be among the most effective of policies in decreasing teen traffic fatalities (Dee, 1999; Dee & Evans, 2001). This is further supported by research showing an increase in morbidity risk at age 21 attributed to MLDA laws (Carpenter & Dobkin 2017). Another policy attributed to the decline in teen traffic fatalities is graduated licensing programs (Dee, Grabowski, & Morrisey 2005; Karaca-Mandic & Ridgeway 2010; Gilpin 2019). These programs limit the exposure of teens to some of the most identified risk factors in fatal accidents; limits to the number of passengers and night time driving. The relationship between other factors, such as beer taxes and minimum wage laws are not as clear, with well-conducted studies finding both null and positive effects (Dee 1999; Adams, Blackburn, & Cotti 2012; Sabia, Pitts, & Argys 2019).

Evidence does exist of the impact of four-day school week policy in other contexts, with mostly positive findings associated with the switch. Early studies reported no negative academic effects after making the switch (Grau & Shaughnessy 1987; Richards 1990; Sagness & Salzman 1993) Donis-Keller & Silvernail (2009) provides a detailed summary of this policy through the mid 2000's. Since that time though, many more school districts have made the switch to a four-day school week. Switches are often attributed to the economic downturn in 2008 (Hill & Heyward 2015), as a cost-saving measure. The reduced schedule does not save a proportional amount in spending, i.e. a savings of 20%, but savings have been identified in: transportation costs, building utilities, and service/support staff (Morton 2020; Kilburn *et al.* 2021; Thomspon 2021). These savings are then redirected to other categories by many school districts switching to the four-day school week schedule to avoid other financial cuts such as teacher layoffs. The schedule change appears to be embraced by most communities (Amys 2016), with few examples of districts switching back to a traditional schedule after funding returned to past levels.

Recent empirical evidence shows mixed results. Academic outcomes, such as achievement scores, shown to have increased or stayed approximately the same in Colorado (Anderson & Walker 2015; Hewitt & Denny 2011), stayed the same in Missouri and Oklahoma (Gower 2017; Morton 2021), and fell in Oregon and Montana for some groups of students (Thompson 2021; Tharp, Matt, & O'Reilly 2016). In the largest analysis of student outcomes to date, across five states (Idaho, Missouri, New Mexico, Oklahoma, and South Dakota), Kilburn *et al*. (2021) find evidence that despite having similar academic outcomes to comparable five-day districts at the time of a adopting the reduced schedule, over time four-day school week districts do not have the same growth in achievement that five-day peers do. Unintended consequences of four-day school week policy have also been examined recently in the literature. Morton (2021) examines district level finances after the adoption of a four-day school week in Oklahoma and finds four-day school week districts have a significant decline in federal revenues and spending on non-instructional and support services. Ward (2019) examines female labor force participation and finds that married mothers have lower labor force participation as four-day school week schedules are introduced. Single mothers however do not change their participation.

Fischer & Argyle (2018) examine juvenile crime in Colorado after the adoption of four-day school week schedules. They find that counties with higher percentages of students on four-day school week schedules have higher rates of larceny after adoption of the schedule than their non-treated peers. An important finding is the increase is not singularly attributed to Fridays, but instead is spread across the week. This provides further evidence that the shift in student schedules should not only focus on the day off. Israel *et al.* (2020) examine differences in adolescent health behaviors between four-day and five-day school week districts and notes a mix of positive and negative health associations with a four-day school week including: higher levels of sexual activity, skipping breakfast, and less sleep. Positives associated with the four-day school week though included: higher participation in extra-curriculars, lower rates of skipping school, and lower rates of using marijuana and cigarettes. While intriguing, these findings must be considered non-causal since the analysis was cross-sectional (i.e. not accounting over time for district-specific characteristics).

2.3 Data and Descriptive Statistics

I constructed an eight-year panel of data from the 2010-2011 school year (starting in August of 2010) to the 2017-2018 school year (finishing in May of 2018). Summer months are excluded since school schedules would not have an impact on those months. The data on involvement in fatal accidents come from the Fatality Accident Reporting System (FARS) dataset. The FARS dataset, published annually by the National Highway Traffic Administration, is considered a near universal-level reporting of all traffic fatalities in the United States. Importantly for this analysis, geolocation information for each accident is reported in FARS, this can be matched to school district boundaries as reported by the National Center for Education Statistics (NCES) to create district-specific fatal crash statistics. The analysis focuses on involvement in a fatal accident (i.e. not just fatalities, but all passengers within a vehicle where a fatality occurred). This is an outcome variable regularly used in the literature (for example see Karaca-Mandic & Ridgeway 2010) due to the low incidence of traffic fatalities. The low incidence of traffic fatalities is especially true of rural school districts, the primary focus of the analysis because of their higher preferences for the four-day school week. In sub-analyses additional information from the FARS dataset (day of the week the accident occurred, gender of those involved, and time of the accident) are used to explore the overall results in more detail.

School year district-level traffic fatal accident involvement counts are linked to state provided information that longitudinally details the adopted schedule (either five-day or four-day) by year for school districts within eight U.S. states (Colorado, Idaho, Kansas, Missouri, New Mexico, Oklahoma, South Dakota, and Wyoming). More information on the number of districts and the years districts within these states adopt a four-day school week can be found in the appendix. These states either publish school schedule information online on their state education department websites or responded to inquiries requesting that information. In addition to district boundaries, school district characteristics come from the NCES. The NCES data includes include district level student teacher ratio, free or reduced lunch percentage, and total student enrollment. Statistical information by district schedule is reported in Table 2.1. Clear differences exist between four-day school week districts and five-day school week districts. Four-day school week districts have nearly half the students of five-day districts in the analysis (1,229 students vs. 2,130 students) and higher levels of students who qualify for free/reduced lunch (63% vs. 54%). Student-teacher ratios are nearly identical (13.42 vs. 13.34), likely attributed to state statutes mandating student teacher ratios kept below a certain threshold. Despite having nearly half as many students enrolled within a four-day school week district as compared to a five-day school week district, involvement in fatal accidents is approximately two-thirds to three-quarters of five-day school week districts depending on the group considered. The number of driving age teens that are involved in fatal accidents is 0.196 individuals per district school year in four-day districts, whereas in five-day districts it is 0.257 individuals. An older group of 26-31 year olds follow this pattern (.269 in four-day vs. .348 in five-day districts). The overall number of individuals involved in fatal accidents is 2.582 in four-day school week districts and 3.175 in five-day school week districts.

2.4 Empirical Model

This analysis uses a difference-in-differences research design to evaluate the causal effect of the adoption of a four-day school week by a school district and involvement of teens (and in robustness checks other age groups) in fatal traffic accidents. The classic two-way fixed effect (TWFE) model generally used to estimate the treatment effect of adopting a four-day school week schedule, denoted by β_1, is shown in the below equation:

$$Y_{it} = \beta_o + \beta_1 FourDay_{it} + \theta X_{it} + \gamma_i + \tau_t + \varepsilon_{it}$$

where Y_{it} is the number of individuals involved in a fatal accident in school district i in year t. $FourDay_{it}$ is the variable of interest and is an indicator for a district adopting a four-day school week, and X_{it} is a vector of time varying district characteristics. These include district enrolment in thousands, free-reduced lunch percentage, and student teacher ratios. γ is a district-level fixed effect, τ is a year-specific effect, and ε is a random error term. Using district-level and year fixed-effects provides a within variation, i.e. a before and after effect interpretation for the adoption of a four-day school week. Reverse causality, or endogeneity, is a common concern with this type of analysis but is unlikely to affect this specific analysis because of no examples found of school districts adjusting their academic schedule to affect teen traffic safety.

Districts switched to a four-day school week schedule at different times throughout our panel years, which causes a slight variation from the traditional difference-in-differences approach where a clear before and after period exists for all observations. Goodman-Bacon (2019) discusses this variation of the classic difference-in-differences approach and demonstrates that it is a weighted-average of the individual TWFE 2x2 matrices. The issue that arises with TWFE models with differential timing lies in the comparison of non-relevant 2x2 matrices. Some of the weights of the matrices can even be negative, which can lead to estimates of treatment effects being negative, even though the effect of treatment is positive. This can also occur when the treatment effect changes over time, i.e. the dosage of the treatment increases.

A second issue that arises with the traditional TWFE difference-in-differences is the common practice of researchers satisfying the parallel trends assumption only once conditioned on time-varying covariates. Additional assumptions are imposed on the data generating process that are often not considered. The first of those assumptions is that treatment effects are homogeneous (in X). The second assumption is that X-specific trends in both treated and comparison groups do not exist. If these two assumptions are not satisfied, then the estimated TWFE treatment effect is generally different from the average treatment effect on the treated.

To address these issues a doubly robust (DR) estimand proposed by Sant'Anna and Zhao (2020) and Callaway & Sant'Anna (2021) is used in this analysis. DR flexibly incorporates covariates into a multiple time period difference-in-differences setup with multiple groups, and provides transparent aggregate treatment effects (overall treatment effect and dynamic treatment effects, i.e. event study estimates are both presented in the results). The two-step estimation strategy uses a bootstrap procedure to conduct asymptotically valid inference which can adjust for autocorrelation and clustering. Using the potential outcomes framework, let $Y_t(1)$ and $Y_t(0)$

be the potential outcomes at time t with and without treatment (adopting a four-day school week schedule). The observed outcome in each period is then expressed as $Y_t = D_t Y_t(1) + (1 - D_t)Y_t(0)$. From this, the average treatment effect on the treated (ATT) is then calculated:

$$ATT_{dr}^{ny}(g, t;\ \delta) = E\left[\left(\frac{G_g}{E\left[G_g\right]} - \frac{\dfrac{p_{g,t+\delta}(X)(1 - D_{t+\delta})(1 - G_g)}{1 - p_{g,t+\delta}(X)}}{E\left[\dfrac{p_{g,t+\delta}(X)(1 - D_{t+\delta})(1 - G_g)}{1 - p_{g,t+\delta}(X)}\right]}\right)\left(Y_t - Y_{g-\delta-1} - m_{g,t,\delta}^{ny}(X)\right)\right] \quad (1)$$

where (using the author's notation) the $ATT_{dr}^{ny}(g, t;\ \delta)$ is a simple weighted average of the difference of the outcome variable considered (count of driving age teens involved in a fatal accident for a district school year for example). The innovation in the DR estimator is to take observations from the control group, which consists of never treated and not yet treated observations in this analysis, omitting non-relevant groups (i.e. comparisons between early treated and later treated groups), and then increasing the weight on observations from the control group that have characteristics similar to those frequently found in the treatment group and reducing the weighting on control observations that are not similar to treatment observations (first part of the above equation). This is what the inverse probability weighting DiD estimator proposed by Abadie (2005) does. The second part of the above equation is directly from the outcome regression model proposed by Heckman *et al.* (1997, 1998). Each of these methods relies on different underlying assumptions, which makes a comparison of the robustness of to model misspecification dependent on the pursued analysis. The DR estimator thus takes properties from each model to create a consistent estimator of the ATT. Double robustness means that the estimand identifies the average treatment effect on the treated even if either (but not both) the inverse probability weighting (Abadie 2005) or outcome regression model (Heckman et al. 1997, 1998) are misspecified. Estimates are calculated using the CSDID package in Stata (Rios-Avila, Sant'Anna, & Callaway 2021).

An important assumption for difference-in-differences strategies is the parallel trends assumption, which assumes that treatment and control groups would have the same trends as they did before a policy intervention and thus the difference in the post policy intervention period is truly the response to the policy. To test this assumption, an event study analysis can be done to estimate the effect of the policy change on the outcome variables for the years before and the years after the change. This is shown in the below specification:

$$Y_{dt} = BX_{dt} + \sum_{\tau=1}^{5} B_{-t}\, Charter_{d,t-\tau} + \sum_{\tau=0}^{5} B_{+t}\, Charter_{d,t+\tau} + \delta_d + \omega_t + \varepsilon_{dt} \qquad (2)$$

where B_{-t} is the effect of charter school establishment on public-school district expenditures in the years leading to the charter school (i.e. how public-school district expenditures were in districts who eventually establish a charter school before implementation). If the coefficients on the years leading to implementation are near zero, then the parallel trends assumption holds. If not, then other explanations could exist for the continual change of the outcome variable. Due to the importance of the parallel trends assumption, which causal interpretation of the coefficients relies on, event study results are first discussed in the following section. As stated above, the flexibility of the DR estimator allows for multiple calculations of treatment effects, including a dynamic treatment effect (event study). The traditional event study presented above has also come under closer scrutiny lately (Sun & Abraham 2020), and thus dynamic treatment effects from Callaway & Sant'Anna (2020) are presented. For all analyses standard errors are clustered at the state level.

2.5 Results

Estimates for the average treatment effect on the treated, where treatment is the adoption of a four-day schedule by a school district, are reported in Table 2.2. In column one the ATT estimate for 15-18 year olds is reported. A district adopting a four-day school week is estimated to reduce involvement in fatal traffic accidents for 15-18 year olds by a statistically significant 0.186 individuals. This translates into an approximate 95% reduction in fatal accident involvement for driving age teens from the sample mean of 0.196 individuals involved in a fatal accident per district-year. As a robustness check done in other traffic accident literature, an alternative treatment group is considered to determine if outside phenomena could be influencing overall traffic safety (and thus groups with no theoretical impact would have the same response). Following previous literature that examines traffic policy changes on driving age teens, an alternative "untreated" group is created of 26-31 year olds. In column 2 this group has an estimated ATT of nearly zero that is insignificant. A final analysis estimates the ATT for all age ranges of individuals involved in a fatal accident and a significant decrease of 0.702 individuals is estimated. While the magnitude appears larger than for driving age teens, it is important to note that approximately 2.582 individuals of all age ranges are involved in a fatal accident per district year in the sample (an approximate 25% decline).

Dynamic treatment effects for driving age teens are shown in Figure 2.1. Darker coefficient estimates represent average treatment effects before the adoption of a four-day school week by a district, whereas lighter coefficient estimates are after a district adopts the schedule. The vertical bands are 95% confidence intervals for each estimate. Before treatment, we would expect average treatment effects to be approximately zero. If estimates are not zero, then endogeneity of the policy or treatment anticipation would violate the parallel trends assumption necessary for causal interpretation of difference-in-differences estimators. In this figure we see significant pre-treatment estimates in the three years before a district adopts a four-day school week. These estimates, while significant, are interesting in that the ATT flips between positive and negative in each of those three years. Despite these mixed results in the period prior to adoption, after adoption of the four-day school week ATT's are consistently negative, and significant or nearly significant for the next five years after adoption.   Figure 2.2., which reports dynamic treatment effects for the older group of 26-31 year olds, shows a consistent near zero estimate throughout the panel. Figure 2.3. which reports dynamic treatment effects for all ages of individuals in the school district, while a bit noisier, is similar to Figure 2.1 showing initial declines in fatal accident involvement (which become insignificant in year 4).

While the above evidence is suggestive of a decrease in driving age teens fatal accident involvement after the adoption of a four-day school week schedule, and no other groups of individuals within the district being clearly affected, an overall causal interpretation is not possible due to the dynamic treatment effects analyses. It is clear though that the adoption of a four-day school week cannot be contributed to any increased risk of fatal accident involvement for driving age teens, despite concerns of the additional potentially dangerous weekend night. Instead, the evidence would suggest a potential decline in risk for driving age teens. A decline in fatal accident involvement for driving age teens is challenging to explain though with the known potential mechanisms that affect traffic safety for teens. Unlike effects found in the Graduated Drivers License literature (Gilpin, 2019), where the mechanism behind improved teen traffic safety is the limitation of risky driving opportunities, the opportunity for driving and risky-driving in particular would have increased under the schedule change. Two mechanisms do stand out in their potential to explain the reductions. The first is an overall reduction in driving that occurs by students staying home, or at least not driving to school regularly, on the fifth-day of the week (this is explored in the next analysis). The second mechanism would be due to a favorable change in the times that students are driving on the road. While this is somewhat unexpected since the existing evidence suggests that earlier start times is associated with more danger for driving age teens (Vorona *et al.* 2015), which four-day school week districts tend to start earlier than their five-day counterparts.

An additional robustness check is conducted using summer months. As described above in the data section summer months are not included in the analysis since school schedule would not affect these months. This presents an opportunity to find support for the overall treatment estimates found during the school year despite the interpretation of pre-treatment coefficient estimates that would indicate concerns the parallel trends hypothesis holds. This analysis is included in the appendix, but findings support an effect of the four-day school week on fatal accident involvement for teens. Overall estimates are null, and event studies do not consistently show evidence of an impact during summer months.

2.5.1 Days of the Week Analysis

If previously mentioned risky behaviors are reduced, then the decline in accident involvement would be attributed to a reduction in risky driving times (Friday and Saturdays, particularly at night). If some other mechanism is in play, such as a change in daily driving patterns or an overall reduction in driving time, then effects could be found during the school week. Using the same specifications, separate analyses are conducted by the day of the week and reported in Table 2.3. Standard errors are relatively large for each sub-analysis due to splitting the data by days, but for most days of the week there does not appear to be an effect of the four-day school week schedule on involvement in fatal accidents for driving age teens. The first part of the school week, Monday through Wednesday, has notably small insignificant estimates. The new weekend days (Thursday through Saturday) do have larger estimated treatment effects, yet each day is insignificant. Sunday has a marginally significant overall decline in fatal accident involvement for teens.

Dynamic treatment effects for each day are reported in appendix Figures 2A.1 through 2A.7. Estimated dynamic treatment effects mostly support the null overall estimates reported in Table 2.3. Figure 2A.1, which reports the dynamic effects for driving age teens on Sundays, does show some evidence of a decline occurring in the years following adoption of a four-day school week. An effect on Sunday would again be hard to explain by a potential mechanism. Additional rest and recuperation for students has been found in qualitative work conducted by Kilburn *et al*. (2021), which could explain safer driving by teens. In author discussions with parents in four-day school week districts numerous anecdotal experiences could attribute to the decline (with the caveat that these can not be tested with the current dataset). These suggestions include: increased family travel on weekends with additional weekend day, parent-custody agreements for children that allow for children to spend additional time under the four-day school week schedule with the non-primary custody parent, and a shift in sports scheduling to more non-school night competitions.

2.5.2 Additional Analyses – Gender, Time of Day, and County Level

Due to the richness of the FARS dataset, additional analyses can be done to pinpoint the groups of driving age teens "driving" the main results. An analysis that splits the group of driving age teens into two groups of males and females is reported in Table 2.4. In the first column, the average treatment effect of adopting a four-day school week on driving age teen males is reported. Males are not measurably affected by the change to a four-day school week. In column 2, the ATT is reported for females, a significant decline of 0.100 driving age teen females involved in fatal accidents. Of the approximately 0.20 driving age teens involved in fatal accidents within these school districts during a school year, males comprise the majority of this group (0.12). This is not surprising considering that teen male crash risk is well established as being higher (Keall, Frith, & Preston, 2005; Shope & Bingham, 2008). Event studies reported in the appendix (Figures 2A.8 and 2A.9). The pre-treatment figures are similar to the overall event study figure, where there is considerable noise in the three years leading to treatment. Females though do show significant declines in the two years post adoption of the four-day school week, with consistent negative but insignificant estimates in the following years. While anticipatory driving effects, or endogenous policy adoption to combat traffic safety, are both unlikely in this situation the failure of parallel trends assumption does lead to caution in interpreting this effect as causal.

Table 2.5 investigates the time of day that driving age teens are involved in a fatal accident. As mentioned before, past literature notes the exceptionally dangerous evening hours for young drivers. Following Dee (1999), night is defined as the hours from midnight to 4:59 am, and daytime is defined as the hours 7:00 am to 3:59 pm. Results from examining driving age teens involvement in fatal accidents at night are reported in column 1. The ATT in this case is nearly zero, which further would suggest that concerns about risky behaviors increasing from the schedule change are not substantiated by the data. An event study in the appendix, Figure 2A.10, shows further evidence of no change in fatal accident involvement during night hours for driving age teens after adopting a four-day school week. In column 2 fatal accident involvement for driving age teens during the day is examined, which again is non-significant (negative sign). The event study analysis reported in the appendix, Figure 2A.11.

In a final analysis, a secondary dataset that contains all traffic accidents (not just fatal accidents) that occur in a county during the school years 2010-2011 to 2017-2018 is explored.

This analysis is limited to only Idaho and Colorado due to data availability. While information for all counties is included in the dataset, the analysis is limited to only 34 counties that had a school district adopt a four-day school week within the county after the beginning of the panel school year (2010-2011). A considerable number of counties had a school district adopt a four-day school in this year or before (73 counties), which due to the popularity of the schedule within these two states is unavoidable when aggregating to the county level. Summary statistics for the analytic sample are reported in Table 2.6. There are considerably more reported traffic accidents within a county that do not include a fatality, the average number of accidents in a county during the school year is 1,409 accidents. Covariates including the percentage of students within the county that qualify for free/reduced lunch (43%) and the number of students enrolled within the county (13,354) suggest that these counties contain more affluent and larger school districts than the districts included in the primary analysis. This follows the reported trend of more urban school districts exploring the four-day school week schedule in recent years.

Table 2.7. reports the average treatment effect on the treated for counties that have a school district first adopt a four-day school week during the panel years, which is a significant decline of 144 traffic accidents (approximately a 10% decline from the sample mean). Dynamic effects of the policy are presented in the event study in Figure 2.4. Limitations of the dataset become clearer when examining these dynamic effects. Pre-treatment periods cannot be calculated for two periods in the years leading to a county school district adopting a four-day school week (which also contributes to parallel trends assumption not holding). More time needs to pass (and additional data for other states included) before a clear causal interpretation can be made. The limited evidence presented here though does suggest that the schedule change could

be influencing a decline in traffic accidents within these counties and further research is warranted.

2.6 Summary and Concluding Remarks

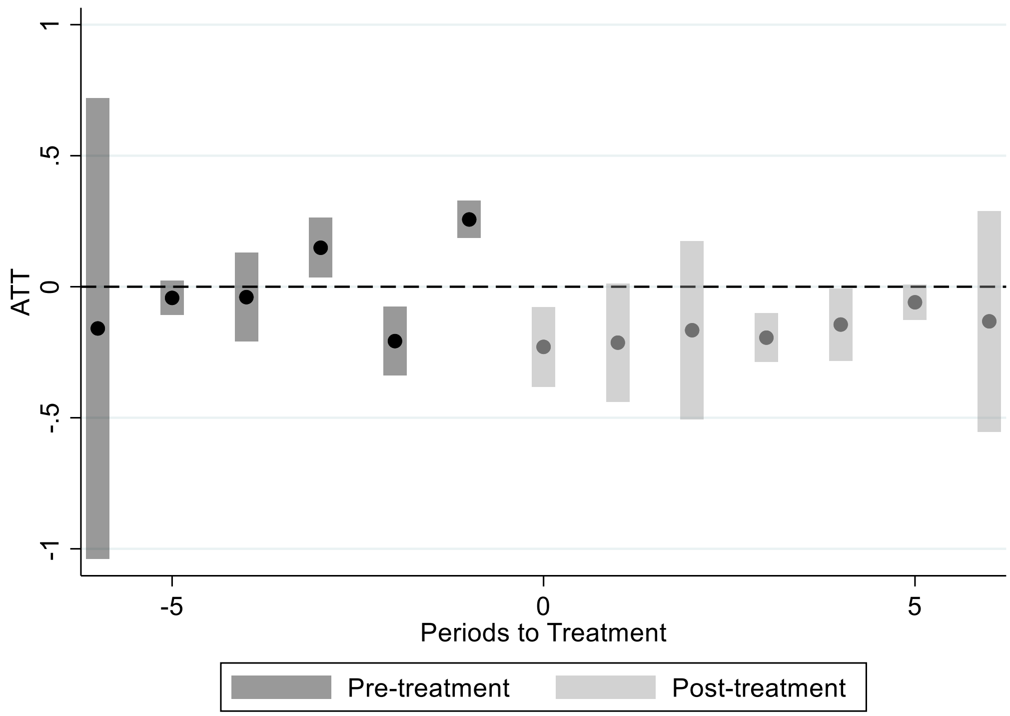
The adoption of a four-day school week by a school district impacts students, parents, and communities in numerous ways. While the immediate impacts on student outcomes and school finance are the primary focus of past research, impacts on local communities have recently came under focus by researchers. The four-day school week schedule has had measurable impacts on juvenile crime (Fischer & Argyle 2018) and female labor market outcomes (Ward 2019). While this analysis cannot clearly state that the four-day school week schedule lowers teen traffic risks after implementation, concerns over increased teen traffic risks on the extended weekends are not substantiated by this analysis. Thursday, Friday, and Saturday nights each show no evidence of increases in fatal accident involvement.

Limitations of this analysis include the primary use of fatal accident involvement to assess potential changes in student risky behaviors or driving risks. Fatal accidents, by definition, do not include the significantly more numerous non-fatal accidents that occur each year. While an attempt is made to explore the effects on non-fatal accidents in two states, data limitations are abundant. Furthermore these non-fatal accidents also suffer from a severity bias, where less severe accidents ("fender benders") are sometimes not reported to police. Ideally direct information on risky behaviors, such as those included in the Youth Risk Behavior Risk Surveillance System (YRBSS), would supplement the analysis. This is what is done in Israel *et al.* (2020), but the analysis is cross-sectional and thus non-causal. Using the YRBSS could be challenging though due to the state-specific methods of data collection and reporting, which means information might not exist for this population of students across the study states.

An additional limitation is that states in the analysis were chosen based on their state education department's willingness to respond to data requests and if public information was available detailing four-day school week adoption. Some states with the most prolific use of the four-day school week schedule, such as Oregon, are not included in the analysis. This was due to the heterogeneity in what makes a district a four-day school week district. In many school districts within Oregon, students attend school five-days a week for part of the year and then switch to a four-day school week near the end of the semester. Some districts even attend school every other Friday. Although these differences are likely impacting traffic patterns of students within these districts, they are a complication to the more homogenous treatment that is found in the states included in this research. While the analysis would be improved by the inclusion of more states that each have similar four-day school week schedules as the states already included, this analysis (to the author's knowledge) is the largest (in terms of states included) in the literature.

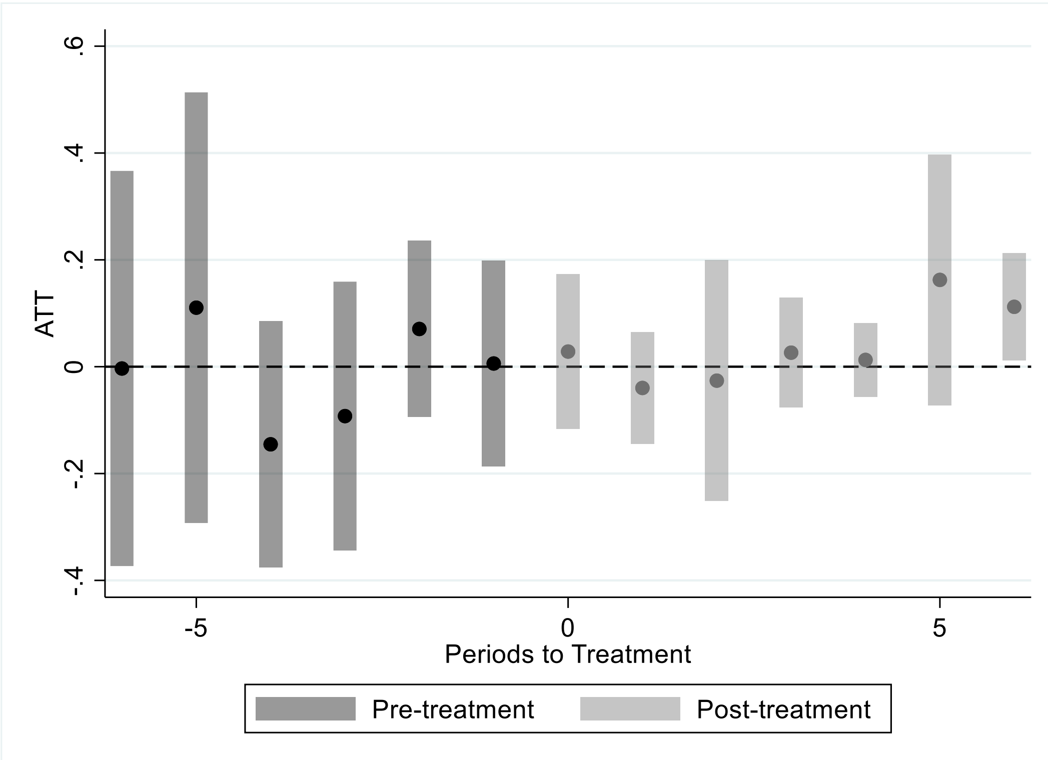
Future research opportunities include assessment of other forms of risky behaviors in these communities after the adoption of the four-day school week schedule. While the YRBSS is not available for all states, individual state analysis is a possibility. Individual state analysis also offers the opportunity to examine the effects on other measures of traffic safety. Furthermore, Fischer & Argyle's analysis on juvenile crime could be extended to other states (again, where data is available) applying this same district-level methodology (instead of their county-level percentage treatment).

Figure 2.1 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident



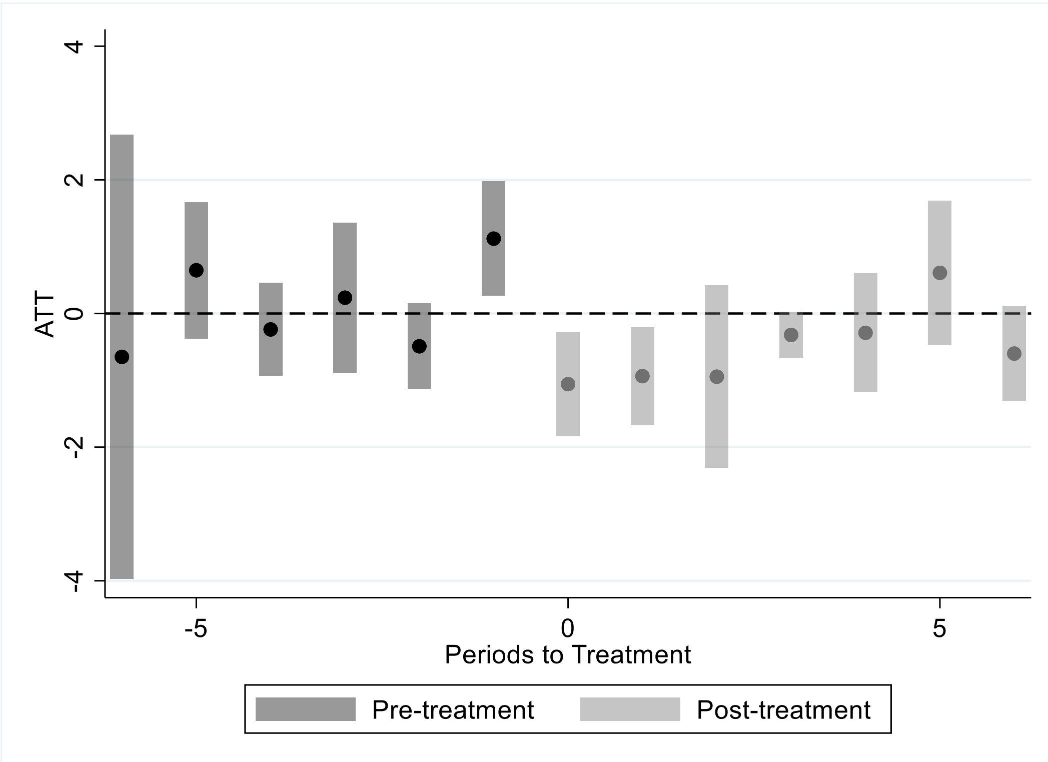
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2.2 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 26-31 Year Olds Involved in Fatal Accident in District



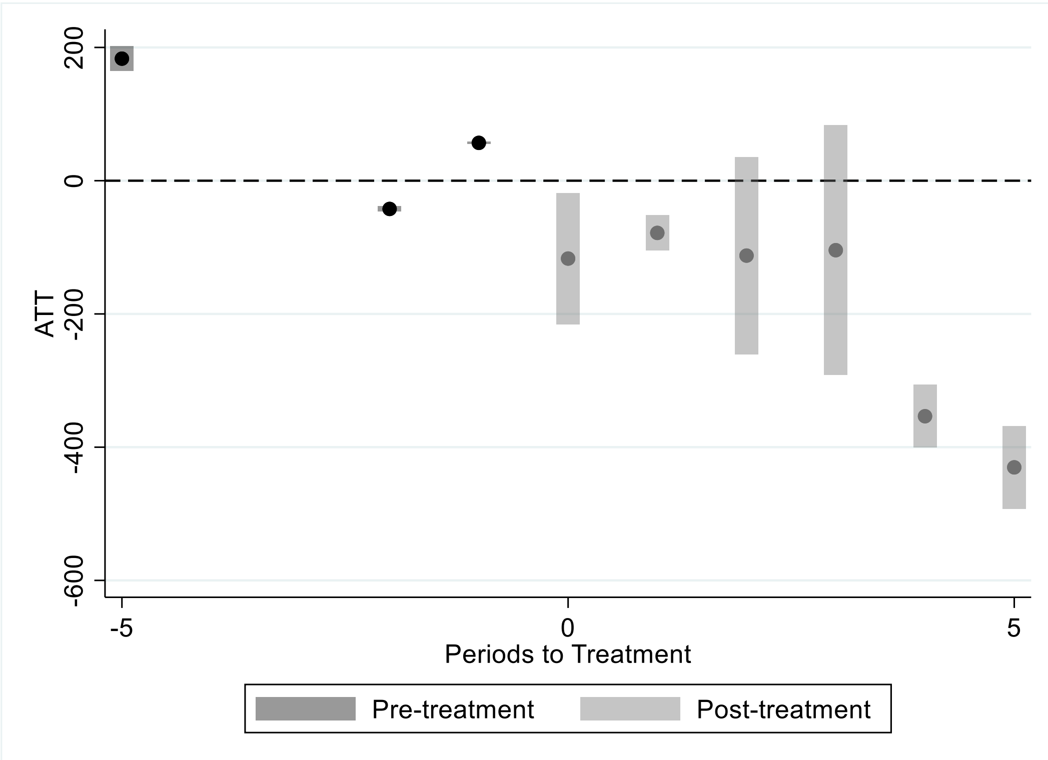
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2.3 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated All Individuals Involved in Fatal Accident in District



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2.4 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week Present in County) on the Treated – Total Traffic Accidents in County During School Year



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Table 2.1 Summary Statistics

	Four-Day District		Five-Day District	
	Mean	SD	Mean	SD
District Characteristics				
Enrollment	1,229	6,786	2,130*	5,377
Free/Reduced Lunch %	0.632	0.186	0.542*	0.199
Student Teacher Ratio	13.42	3.408	13.34	3.702
Outcomes – Involved in Fatal Accident				
15-18 Year Olds	0.196	0.769	0.257*	0.929
26-31 Year Olds	0.269	1.225	0.348*	1.094
All Ages	2.582	7.459	3.175*	7.594
Number of Districts	164		1,490	
Number of Observed Years	1,312		11,920	

Notes: District characteristics are from the National Center for Education Statistics (NCES). Involvement in fatal accident counts are author's calculations from Fatality Accident Reporting System (FARS) data from the National Highway Transportation Administration.
*Significantly different at 5% level in t-test analysis

Table 2.2 Average Treatment Effect on Treated of Four-Day School Week Schedule on Number of Individuals Involved in Fatal Accident for School Year

	15-18 Year Olds	26-31 Year Olds	All in District
ATT – Four Day School Week	-0.186**	0.017	-0.702***
	(0.086)	(0.036)	(0.235)
Mean Outcome for Four-Day District	0.196	0.269	2.582
	(0.769)	(1.225)	(7.459)
Observations	13,232	13,232	13,232

Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios. Standard errors robust to heteroskedasticity in parentheses, clustered at state.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.3 Average Treatment Effect on Treated of Four-Day School Week Schedule on 15-18 Year Olds Involved in Fatal Accident by Days of the Week

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
ATT	-0.051*	-0.003	-0.028	0.003	-0.042	-0.002	-0.063
	(0.030)	(0.010)	(0.026)	(0.013)	(0.026)	(0.017)	(0.063)
Mean Outcome	0.040	0.018	0.030	0.025	0.017	0.033	0.035
for Four-Day District	(0.287)	(0.153)	(0.283)	(0.199)	(0.169)	(0.261)	(0.274)
Observations	13,232	13,232	13,232	13,232	13,232	13,232	13,232

Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios. Standard errors robust to heteroskedasticity in parentheses, clustered at state.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.4 Average Treatment Effect on Treated of Four-Day School Week Schedule on Number of 15-18 Year Olds Involved in Fatal Accident for School Year by Gender

	Males	Females
ATT	-0.087	-0.100**
	(0.064)	(0.044)
Mean Outcome for Four-Day District	0.115	0.081
	(0.501)	(0.370)
Observations	13,232	13,232

Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios. Standard errors robust to heteroskedasticity in parentheses, clustered at state.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.5 Average Treatment Effect on Treated of Four-Day School Week Schedule on Number of 15-18 Year Olds Involved in Fatal Accident for School Year by Time of Day

	Night	Day
ATT	-0.006	-0.073
	(0.013)	(0.057)
Observations	13,232	13,232

Time of day is split into day and night based on previous literature (Dee, 1999). Night is from the hours of midnight to 4:59 am and day is from the hours 7:00 am to 3:59 pm.
Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios. Standard errors robust to heteroskedasticity in parentheses, clustered at state.
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6 Summary Statistics for County Traffic Accident Analysis

	Mean	SD
School Year Traffic Accidents	1,409.691	(3,079.84)
Free/Reduced Lunch Fraction in County	0.431	(0.165)
Student Enrollment in County	13,354.82	(22,771.88)

Notes: Dataset are for the sample that consists of 34 counties in Idaho and Colorado from the 2010-2011 to 2017-2018 school years. School year traffic accidents in the county came from each states motor vehicle department. Free/reduced lunch fraction and enrollment in the county are calculated from district information from the National Center for Education Statistics (NCES).

Table 2.7 Average Treatment Effect on Treated of Four-Day School Week Schedule in County on School Year Traffic Accidents

ATT	-144.249*
	(87.718)
Observations	272

Doubly Robust difference-in-differences estimator with not yet treated observations as control group. Covariates include: free reduced lunch fraction of students in county, county enrollment. Standard errors robust to heteroskedasticity in parentheses, clustered at state.
*** p<0.01, ** p<0.05, * p<0.1

Chapter 3 - Academic Probation & Financial Aid: Financial Aid Implications of Probation

3.1 Introduction

Why do students receive grades for assignments and overall performance in a class? Many teachers would say they grade to provide feedback to students on performance, and to give cues on how to improve future performance. College bound students in the United States are familiar with receiving feedback on their performance from their years in high school. Many at this point though have not experienced the same level of independence and responsibility that is expected of a college student, which can be a challenge for many students. Unfortunately, some students will be placed on academic probation during their college years because of poor performance. Academic probation can be a major shock for students who have not previously underperformed (otherwise they might not have made it to college).

The mechanics of academic probation varies by university, but all share some common features. Students who have a GPA below a certain threshold are notified of their less-than-desired performance and placement on academic probation. Depending on the university academic probation could be just a simple notification and encouragement to increase their GPA in the future to meet graduation requirements, or academic probation could lead to suspension and loss of institutional financial aid if the GPA is not increased in a designated period. Regardless of each institution's specific academic probation features, academic probation is a shock to many students. Student responses to this shock are important to understand. If students respond positively to academic probation and improve their subsequent academic performance then the policy has achieved its intended effects. If students respond by disenrolling from the university, then the policy could be responsible for disenfranchising students who otherwise would have graduated. Underlying these responses are important student characteristics that influence their responses to academic probation. This paper seeks to answer the question of how a student's financial aid status, as a Pell grant recipient or not, affects their response to being placed on academic probation.

Despite the good intentions of the intervention, the existing evidence suggests heterogeneous responses occur to being placed on academic probation. Positive impacts on a student's next semester GPA are common (Lindo *et al.* 2010; Yeaton & Moss 2018; Bowman *et al.* 2019; Casey *et al.* 2018; Wright 2020), but numerous studies have also documented deterrent effects of academic probation on retention (Lindo *et al.* 2010; Ost *et al.* 2018), obtaining a bachelor degree (Ost *et al.* 2018), and long-term earnings (Ost *et al.* 2018). Thus the positive effect on GPA is unlikely to outweigh the potential negatives from disenrollment in the university unless disenrollment mostly occurs within students unlikely to graduate and thus save future college costs (including the cost of lost income during those years). These effects, and different interpretations of those effects on student sub-populations, makes understanding the impacts of academic probation of utmost importance for policymakers.

This analysis expands on the research by further investigating the impacts of academic probation by financial aid status, something that is missing from previous analyses. Recipients of the Pell Grant, a subsidy to students who demonstrate financial need from the U.S. government, must meet satisfactory academic progress requirements. These requirements include maintaining a cumulative GPA of 2.0 or higher for upperclassmen and a GPA of 1.7 for freshmen. While non-Pell students must also meet satisfactory academic progress requirements to receive financial aid, merit aid GPA requirements are higher (2.5) and thus satisfactory academic progress does not have the same financial consequences at the lower GPA requirement. In work focused on Pell grant students, Schudde & Scott-Clayton (2016, 2020) has found disenrollment effects from satisfactory academic progress requirements and positive impacts on remaining Pell students' GPAs. This coincidence of GPA cutoff though has the potential to confound estimates of the effect of academic probation on academic outcomes if Pell status is ignored. Using longitudinal data from a large public university in the U.S. Southwest this paper estimates the causal impact of being below the GPA cutoff for those who receive the Pell Grant and those who do not using a regression discontinuity (RD) design (the causal impact must be interpreted as the combined effect of satisfactory academic progress and academic probation for Pell students, but only academic probation for non-Pell students). The impact of the cutoff on the size and composition of the financial aid package is also investigated. Results indicate that despite implied financial consequences for freshmen Pell recipients, both groups of students respond similarly and have a positive increase in next semester GPA after being placed on academic probation. Upperclassmen do not respond to placement on academic probation and satisfactory academic progress. When examining next semester enrollment and the eventually receiving a bachelor degree (four-year, five-year, and six-year graduation are included) evidence is found that non-Pell students are negatively affected by the academic probation GPA cutoff.

The remainder of this paper is as follows. Section two describes the existing literature on academic probation in detail. Data and descriptive statistics are discussed in section three. The empirical strategy employed is reviewed in section four. Results and robustness checks are presented in section five, and section six concludes.

3.2 Academic Probation Literature

An early analysis of academic probation came from Lindo, Sanders, & Oreopoulos's (2010) study in the Canadian university system. Employing a regression discontinuity design, they find that first-year students on academic probation who persist into the next semester significantly increase their GPAs. Further research in a midwestern U.S. university by Bowman *et al.* (2019) finds that sophomore and junior engineering students respond to academic probation by increasing next semester GPA, while first-year students do not increase subsequent semester GPA. Bowman *et al.* (2019) suggest that the structure of academic probation, with additional meetings with academic advisors and more contact with the student, are more beneficial to these upper-classmen since support services are traditionally more prevalent for first-year students. Although not explored in this analysis, the mechanisms behind the GPA increase have also been investigated. Casey *et al*. (2018) explores changes in the courses students enroll in once placed on academic probation. They find that students engage in "strategic course-taking", meaning that they enroll in easier courses to improve GPA and are more likely to drop courses than non-academic probation peers. This behavior is shown by non-minority students only though, suggesting that minority students do not have access to the same helpful institutional knowledge.

When considering measures of college persistence, Lindo, Sanders & Oreopolous (2010) found that some subgroups, for example males, those who were higher performers in high school, and native speakers are more likely to drop out once placed on academic probation. These effects are not explained by differences in how probation is administered between groups, making these findings concerning and not easily explained. It is hypothesized that academic probation could serve as more of a "shock" to previously high performers since this a "new" assessment of their abilities. More research is needed to explain why some groups respond differently, but not all see a deterrent effect on enrollment as a negative outcome. An alternative view is that if students placed on academic probation decide to drop out and enroll in another university with more suitable programs or coursework, then this could be a positive outcome. Wright (2020) does consider this exact outcome, where students transfer to another university but does not find a significant effect. Likewise, if a student who otherwise would not graduate is notified at an early stage of a lower likelihood of graduating due to current performance, then that student can reallocate their time to other training or work before incurring additional debt or not receiving income from full-time work. It is important to note that in a U.S. university setting Casey *et al.* (2018) are not able to replicate this negative effect on future enrollment. Ost *et al.* (2018) examine credit hours completed and finds a significant decline for those who are placed on academic probation. In this same analysis they also find an 11-percentage point decline in the likelihood of obtaining a bachelor degree[1].

A separate line of literature is also important for this analysis, and that is the existing research on satisfactory academic progress for Pell Grant recipients. The mechanisms behind not meeting satisfactory academic progress for the Pell Grant are more complicated than academic probation. National requirements for satisfactory academic progress include the GPA requirement, a course completion rate requirement, and a program completion requirement. These requirements are set at the university level, which creates variation between university settings. Schudde & Scott-Clayton's (2016) statewide analysis of community college students finds inconclusive effects of satisfactory academic progress on education outcomes when employing regression discontinuity design. Difference-in-differences analyses though suggests that Pell recipients are more likely to drop out than similar non-Pell receiving peers, but the different method employed does question how comparable the results would be to these other analyses. In a different state, but also examining community college students, Scott-Clayton & Schudde (2020) do find discouragement effects on enrollment consistently between both

[1] Ost *et al.* (2018) extends the analysis to future earnings. They find that while those who dropped out after being placed on academic probation earn more in the following 3 years than those who persisted, students who persist catch up and earn more for approximately the next decade. They estimate an internal rate of return (IRR) of 4.1%. of college persistence.

regression discontinuity and difference-in-differences methods. They also find increases in GPA from satisfactory academic progress requirements.

The literature on the effects of losing financial aid (or in the case of regression discontinuity analyses, being below a financial aid cutoff) has been mixed with some evidence showing a decrease in next semester credit hours and the probability of next semester enrollment (Carruthers & Özek 2016). While others have shown no effect on college persistence (Welch 2014; Jones *et al.* 2020). The loss of financial aid has also been shown to increase student's labor market time (Carruthers & Özek, 2016), which can be interpreted as time allocated separately from the time spent on the production of human capital (Becker 2009). This is in line with research that shows receiving additional financial aid, in the form of performance-based scholarships, reduces allocated time to work or leisure and increases engagement with studies (Barrow & Rouse 2018) Receiving merit-based financial aid has been positively linked to time-to-degree and positive socioeconomic outcomes after college (Scott-Clayton & Zafar 2020).

This research contributes to the previous literature by examining both freshman and upperclassmen, while adding an additional study setting in a large public southwestern university where the majority of students enrolled are minority. The research conducted in this analysis combines the academic probation and satisfactory academic progress lines of literature. Previous research on academic probation has thus far ignored the potential financial implications of the policy for students. This detail is important to consider when investigating if pecuniary penalties are appropriate for under-performing students.

3.3 Data and Descriptive Statistics

I use an administrative dataset of undergraduate students at a public university in the U.S. Southwest. Observations are at the student-semester level and cover the academic years 2006-

2007 to 2018-2019. Students enrolled in this university are placed on academic probation at the end of any semester in which their cumulative grade point average at the university falls below 2.00. This is most likely to occur after the first semester in this analysis. Of all instances of having a GPA below the academic probation threshold 36% of occurrences happen in semester one. An additional 24% happens in semester two. Academic probation could also occur in later semesters if a GPA that was previously above the cutoff is brought below the cutoff by a poor semester[2]. Once a student is placed on probation they are required to periodically meet with their academic advisor during the probation semester and comply with terms of their probation status. Students on academic probation will be academically suspended at the end of the probation semester if the cumulative GPA does not rise to 2.00 or better, or if "reasonable progress" has not been made in meeting that GPA. Reasonable progress is defined as earning at least a 2.5 GPA, having at least one-half of the student's course load applying toward a student's degree requirements, and passing at least two-thirds of the course credits for that semester. If suspension does occur students may not enroll for classes for at least one semester from the date of the suspension. This extends to two semesters for a second time being suspended, and for five academic years for a third offense. Students who are placed on suspension can appeal the decision if they provide documented extenuating circumstances (the extent to which this occurs is not documented in the obtained administrative data).

Satisfactory academic progress has the same GPA cutoff for upperclassmen, but a lower GPA cutoff of 1.7 for freshmen (those below 31 cumulative credit hours). In addition to the GPA requirements, students at the university are required to complete 2/3 of attempted credit hours

[2] In research segmenting by student progress in the program, Bowman *et al.* (2019) finds that sophomore and juniors respond with increased GPAs in the subsequent semester but freshman do not

each semester and complete their program of study within 150% of the published length of the program (measured in credit hours). To receive institutional financial aid these requirements must all be met, but programs and merit aid are likely to have other more stringent requirements. For example, institutional aid from the university has GPA requirements between 2.5 and 3.3 GPA that varies between cumulative (like academic probation) and term specific. A student thus can lose some forms of financial aid while still meeting academic probation and satisfactory academic progress requirements. Importantly for this analysis though, is that students receiving the Pell grant are required to meet satisfactory academic progress's requirements. This creates financial consequences non-Pell students do not suffer from when not meeting satisfactory academic progress.

The data includes student-level data measured by a semester (such as: semester GPA, financial aid award package, loans offered through financial aid), cumulative (cumulative GPA), and fixed (student race, ethnicity, age at matriculation, and high school GPA). For each analysis the sample of students is observed for two consecutive semesters, where in the previous semester the student would be placed on academic probation due to falling below the cumulative GPA threshold and the subsequent semester would be their response (enrollment, GPA, etc.) to academic probation placement. Students are omitted if they are missing any of the covariates or outcome variables described below. Missing covariates most often occur with high school GPA (11% of students missing high school GPA). These students with missing high school GPA are slightly older and are more likely to report a race and ethnicity of unknown, two or more races, or non-resident alien (these categories are grouped together for statistical analysis as "Other"). In the case of missing outcome variables, this only occurs for next semester GPA when the student does not enroll in the following semester. This is a regular challenge in the academic probation literature and its implications are explored in the results section further. For now, the bias from missing this information has an unclear direction due to multiple possible interpretations.

The primary outcome variables in this analysis are: next semester enrollment, next semester GPA, next semester financial aid package, next semester loans, and receiving a bachelor degree. Next semester enrollment is an indicator if the student enrolled at the university the semester following being placed on academic probation. Academic probation's goal is to serve as a "warning" to students that their academic performance is below university goals and jeopardizes their future ability to complete their degree. Nowhere is it stated that academic probation is intended to encourage students to disenroll from the university (otherwise the consequences of falling below the GPA threshold would be more severe, such as suspension). Next semester GPA is the semester GPA earned in the semester following being placed on academic probation. This measure ignores the composition of classes, or the number of credits attempted, which have been shown to be affected by academic probation but neither of which are included in academic probation policies. Strategic course taking (i.e. taking easier courses) and reducing credit loads are two ways that students could try to increase their GPA, but each of these could result in a longer time to degree. The next outcome considered is next semester financial aid package. This is a measure of all institutional, state, and federal financial aid awarded (and accepted) by the student for the semester. This also includes private financial aid administered through the university (such as foundation scholarships). The variable next semester loans is only the loans component of the financial aid package (i.e. it is a subset of the total package), including direct subsidized and unsubsidized loans. No distinction is made between the two in this analysis.

The final outcome considered is receiving a bachelor degree. Four-year, five-year, and six-year graduation is considered separately. For four-year degree receipt, only students who are observed for eight fall-spring semesters are included in the analysis, which includes all cohorts who started in Fall 2015 and before (Five-year, and six-year follow with Fall 2014 and Fall 2013 respectively). Considering that a student can be placed on academic probation at any point in their academic career, a challenge of the analysis is determining the appropriate control group (i.e. a student who is placed on academic probation in their 2nd year Fall semester should be compared to a different group than a student placed on academic probation after their 3rd year Fall semester). Small sample sizes limit the ability to focus on later groups in separate analyses, so the analysis included here is simplified to students who are placed on academic probation after their first semester enrolled at the university. This is the most likely time to be placed on academic probation as mentioned above, with 36% of all probation instances occurring at this point. This makes the analysis comparable to Lindo, Sanders & Oreopolous' (2010) analysis which examined graduation outcomes for students placed on academic probation after their first full year. Length of treatment, i.e. the number of semesters that a student is on academic probation after the first semester, is also a potential concern in this type of analysis. For the full sample of students, approximately 9% of total students are placed on academic probation once during their academic careers (82% are never placed on academic probation). An additional 5% is placed on academic probation twice due to their cumulative GPA. Some students are on academic probation for three semesters or more during their academic career due to cumulative GPA being below the cutoff, but this is rare. It is important to note that for a student to be placed on academic probation multiple semesters, "reasonable" progress as described above needs to have been met in the subsequent semester (a semester GPA of 2.5 or higher), thus disenrollment effects for these multiple treated students are less likely to be a concern.

Table 3.1 shows descriptive statistics by Pell grant recipient status. Columns two and three in Table 3.1 shows the descriptive statistics for students who did not receive the Pell grant. Columns four and five show descriptive statistics for students who did receive the Pell grant. Pell recipients are statistically different from their non-Pell peers in nearly all of the observed characteristics shown. Pell recipients are more likely to be females and a minority. Differences in high school GPA and semester GPA between Pell recipients and non-recipients, while statistically different, do not appear meaningfully different for most interpretations (3.4 vs. 3.3 high school GPA and 3.0 vs. 2.9 collegiate GPA). Financial aid packages are significantly larger for Pell recipients, who receive $7,715 compared to $4,439 for non-Pell students. Pell recipients also receive more in loans through their financial aid packages, $1,770 compared to $974 for non-Pell students[3].

While the differences between Pell and non-Pell students at the university are clear, the university is also unique compared to other public universities in the United States. In the most recent IPEDS 2020 report representing students enrolled in the Fall 2019 semester (U.S. Department of Education 2020) the percentage of Native American and Hispanic students are significantly higher at the university than twenty comparison universities (comparison universities include nearby neighboring state universities and the other major university located in the same state). Comparison universities have a median of 21% of Hispanic students, compared to 44% at the study university, and 0% Native American students compared to 5%.

[3] This analysis does not include information for private loans offered to students

These differences are mostly explained by the fact that the university is located in a minority-majority state. The university is also recognized as a Hispanic Serving Institution. The university is also considerably less expensive when considering posted tuition and fees ($7,875 at the study university vs. $10,042 at comparison universities) and total cost of attendance minus financial aid ($11,368 at the study university vs. $15,676 at comparison universities).

3.4 Empirical Strategy

As noted in previous analyses of academic probation, a simple comparison between students on academic probation and students not on academic probation would likely lead to incorrect interpretations of the effect of academic probation due to unobserved differences between students. Family background characteristics such as household income, parent education, and ability, are likely to be considerably different between those placed on academic probation and those not placed on academic probation. A regression discontinuity design is used to exploit the predetermined GPA threshold that defines assignment to academic probation. This type of analysis relies on the assumption that students cannot manipulate their GPA within a local area around the cutoff (also referred to as the threshold). With that assumption, analyses examining students that fall just on either side of the threshold should generate quasi-experimental variation in assignment to academic probation. Considering the difficulty of manipulation of this kind – where final grades are released independently at the end of the semester by numerous professors in multiple departments, and a student would then need to convince one or more of these professors for a grade change once other grades are known – the theoretical assumption of non-manipulation is likely to hold.

In this setting at the end of a semester the decision to be placed on academic probation for student i in semester t is a deterministic function of their GPA, which can be expressed as

$$PROB_{it} = 1(GPANORM_i < 0), \quad (1)$$

where $GPANORM_i$ is the distance between student i's GPA and the probationary cutoff. As long as the discontinuity status is "sharp" (i.e. students with a GPA below the cutoff cannot avoid academic probation and students above the cutoff cannot be placed on academic probation for other reasons, thus making the likelihood of being on academic probation less than 100% if a GPA is below the cutoff), then student characteristics and outcomes should be continuous through the threshold, and the treatment effect for students near the threshold can be calculated by comparing outcomes of students just below the threshold to those just above the threshold.

The following equation can be used to estimate the effects of academic probation on subsequent student outcomes:

$$Y_i = \beta(GPANORM_i < 0) + m(GPANORM_i) + \rho(GPANORM_i) \times (GPANORM_i < 0) + \beta_n X_i \ u_i) \quad (2)$$

where Y_{it} is an outcome for student i in semester t, $(GPANORM_i < 0)$ is an indicator equal to one if the student's GPA is below the academic probation cutoff, $(GPANORM_i)$ is a continuous function of a student's standardized GPA (GPA is standardized to the probation cutoff's 2.0 GPA requirement, thus any number above (below) zero is the distance in GPA units to the cutoff), X_i is a vector of individual-level covariates (including: gender, race, ethnicity, high school GPA, and age at admission), and u_i is a random error term. The coefficient of interest is β, the estimated impact of being placed on academic probation.

As stated above, the GPA cutoffs for freshmen are different for satisfactory academic progress and academic probation. An analysis that does not account for these two different treatments could misattribute the effect of one policy to the other unless conducted separately. Following Cattaneo *et al.'s* (2016) guidance when conducting regression discontinuity analysis with cumulative cutoffs each cutoff is analyzed separately with bandwidths chosen to not overlap between the two cutoffs (bandwidth is 0.29 for each cutoff). For upperclassmen the bandwidths were chosen based on prior academic literature using similar RD analyses (Casey *et al.* 2018), where the preferred specification uses a data-driven bandwidth selection process as described by Calonico, Cattaneo, & Titiunik (2014), henceforth CCT. This analysis is conducted separately by Pell status to ensure confounding of satisfactory academic progress is not driving overall estimates of the effect on academic probation. The analysis is also limited to those students who meet the other requirement of satisfactory academic progress (course completion ratio) to ensure that GPA is the mechanism by which a student would not be meeting satisfactory academic progress. The sample is then limited to only students who are within defined bandwidths of the academic probation cutoff. Heaping, which occurs when many observations occur at the same point, is a concern with GPA data. In this instance, heaping occurs at the 2.0 GPA cutoff because of the way GPA is calculated as weighted discrete units. With heaping occurring to the right of a cutoff, estimates of the treatment effect can be biased upwards (Barreca, Lindo, & Waddell 2015). Following Barreca, Lindo, & Waddell's recommendations, and recent literature (Casey *et al.* 2018), a "donut" regression is employed around the cutoff and observations with a 2.0 GPA are removed from the analysis.

3.4.1 Validity of RD Design

As discussed in the RD literature, the primary threat to identification is manipulation of the running variable. In this context, students near the academic probation cutoff would need the ability to influence final grades in a manner to avoid being placed on probation. This seems unlikely due to the complexity of the task, but if students were able to manipulate their GPA in this manner it would be observable in the data by the density of the GPAs near the cutoff. If students are able to manipulate their GPA, we would expect a large "mountain" of students just to the right of the cutoff and a small "valley" of students just to the left. If this pattern existed, we would make the conclusion that some unobservable difference between students near the cutoffs (motivation, institutional knowledge, etc.) partially explains differences in outcomes between students on probation and students not on probation, confounding the research design.

To investigate this, the top panel of Figure 3.1 shows the histogram of normalized semester GPA. Some evidence of heaping exists at the lowest whole numbers (translating to a GPA of 0.0 and 1.0). The overall distribution is clearly normal, with students having GPAs throughout the 0 to 4.33 (a grade of A+ at the university translates to a 4.33) range in a continuous fashion. The bottom panel of Figure 3.1 presents the same distribution of semester GPA excluding students who earned a whole number GPA. No heap or valley appears to occur near the cutoff in either panel, reassuring concerns of GPA manipulation in this area.

Another common test in the literature to verify if GPA manipulation could occur is to estimate the RD equation presented above for each of the observable characteristics as the dependent variable. Table 3.2 presents results from these regressions for freshmen by Pell status and Table 3.3 presents results from these regressions for upperclassmen by Pell status. If there are substantial changes at the cutoff for multiple observable characteristics then the validity of the research design would again be called into question as certain groups at students might be able to influence their grades to avoid being placed on academic probation. For upperclassmen in Table 3.3 multiple bandwidths are presented across columns 1-3 to investigate the sensitivity of the estimate of the covariate's impact on being placed on academic probation. In RD designs, a tradeoff occurs between the ability to precisely estimate the true effect of the treatment and the size of the bandwidths considered. The preferred data-driven bandwidth by CCT is reported in column 1, with additional bandwidths reported in columns 2 and 3.

Estimates across covariates in Table 3.2 show no evidence of manipulation of the cutoff across most observable characteristics for freshmen. Pell students (columns 3 and 4) do not appear to manipulate either the satisfactory academic progress or academic probation cutoff, but some non-Pell students seem to. Those from other race categories, younger students, and those with a higher high school GPA are more likely to be below the academic probation cutoff. Older students successfully navigating the cutoff could be explained by prior college experiences not at this university. Those with higher high school GPAs could possibly not have gained experience marginally improving grades while in high school, thus are unable to respond like other students in college. The likely direction of bias if students can successfully manipulate GPA would be to bias estimates in a negative direction, leading to an interpretation that academic probation as hurting student outcomes. Table 3.3 reports results from the same analysis with upperclassmen. Less concern over manipulation exists for upperclassmen, with no observable characteristic consistently showing manipulation across bandwidths. As has been done in previous analyses in the literature these covariates will be included in the RD analysis (Calonico *et al.* 2019). Results without their inclusion (the theoretically traditional RD design) are available upon request.

3.5 Results

The impact of satisfactory academic progress and academic probation GPA cutoffs on freshmen student outcomes is first investigated in Table 3.4. Column 1 reports the estimate for the effect of the satisfactory academic progress GPA cutoff on non-Pell students. In panel A. next semester outcomes are shown for enrollment, financial aid, and loans. In column 2 the effect of academic probation is presented for non-Pell students. First considering next semester enrollment in row 1, no evidence is found that either cutoff impacts non-Pell students. Examining columns 3 and 4, which report the same estimates for Pell students, shows similar non-significant findings. Scott-Clayton & Schudde (2020) do find discouragement effects for Pell students from satisfactory academic progress on enrollment in a community college setting. The differences between the two types of students, and institutional supports within a university, are important to consider when interpreting these different findings. In row 2 next semester financial aid is examined. Again, neither GPA cutoff measurably affects both groups of students' next semester financial aid. The magnitude of the satisfactory academic progress cutoff is considerably larger than the academic probation cutoff, which is expected since academic probation does not in itself have a financial consequence. In unreported analyses the pooled effect of the two cutoffs does result in a significant decline in financial aid for Pell recipients, but neither policy individually has a measurable effect.

Moving to panel B, the effect of both policies on next semester GPA for non-Pell students who enroll the next semester are shown in columns 1 and 2. The satisfactory academic progress cutoff does not appear to affect next semester GPA but being placed on academic probation increases next semester GPA by 0.213 GPA units for non-Pell students. Non-Pell students do not have financial consequences at the lower GPA cutoff since all institutional aid would have been lost once falling below a 2.5 GPA, and thus are not measurably responsive to that cutoff This same effect is found in columns 3 and 4, with Pell students responding to academic probation by increasing next semester GPA by a marginally significant .202 GPA units. A small sample size could explain why despite financial implications for Pell students at the satisfactory academic progress cutoff they do not significantly increase GPA at that cutoff. The increase in next semester GPA for both groups of students is in line with other analyses of academic probation (Lindo, Sanders and Oreopoulos 2010; Casey *et al.* 2018). Bias does potentially exist in this estimate of the effect of academic probation (and satisfactory academic progress) on next semester GPA due to change in the composition of students who continue to enroll in the university. The findings in panel A. should reassure some concerns of a change in composition since a significant disenrollment effect was not found but considering the direction of potential bias is a helpful exercise. If academic probation results in lower-ability students disenrolling from the university, then we would tend to find positive estimated impacts on next semester GPA even if being placed on academic probation does not encourage better subsequent performance. Alternatively, if being placed on academic probation is particularly discouraging for the higher-ability students and causes disenrollment of this group, the estimated impact on subsequent GPA will be biased downward.

Panel C reports results for an analysis looking into the effect of early academic probation on a student's likelihood of receiving a bachelor's degree. As described above the semester GPA examined is limited to after the first semester of college attendance (i.e. freshmen near the satisfactory academic progress and academic probation cutoffs in semester 1), which is 36% of all occurrences of academic probation at the university. In the first row of Panel C., the effect of academic probation on receiving a bachelor degree in four years is reported. To observe at least 8 semesters, the latest cohort of students included in this analysis started at the university in the fall of 2015. The earliest cohort of students began in the fall of 2006. An assumption is made to determine the semester a student received their bachelor degree since only an indicator if they received a degree or not is included in the dataset. This assumption is a student who receives a bachelor degree received the degree in the last semester they enrolled in the university (i.e. a student does not attend additional semesters after receiving a bachelor degree). In columns 1 and

2 the results for non-Pell recipients are reported. Being below the satisfactory academic progress cutoff does not impact four-year graduation, but being placed on academic probation associated with a 6.5% lower likelihood of receiving a bachelor's degree within four-years for non-Pell students. When expanding the analysis to include five-year and six-year recipients, the negative effect of academic probation for non-Pell students continues (13.7% for five-year graduation, and 13.0% for six-year). Pell student results are reported in columns 3 and 4 and across each cutoff and graduation time frame Pell students are not measurably affected by either policy. Sample sizes are small though, potentially masking an effect. Considering the lack of a disenrollment effect from Panel A, negative effects on non-Pell students graduation rates points to other mechanisms by which graduation can be affected. Strategic course-taking, or reducing credit loads, are potential explanations for the reduced graduation rates.

Estimates for upperclassmen are presented in Table 3.5. Upperclassmen do not have the lower satisfactory academic progress GPA cutoff, thus each analysis is only conducted at the 2.0 combined GPA cutoff. Thus estimates are to be considered the combined treatment effect of each policy. Panel A. presents results for next semester outcomes of enrollment, financial aid, and loans for both non-Pell and Pell students. Across each outcome neither group is measurably affected by the combined cutoff. Standard errors are considerably larger for each outcome, providing evidence that there is considerably more variation in outcomes for upperclassmen then freshmen. This is the same when examining next semester GPA for those who enroll the next semester in Panel B. Neither group appears to respond to being placed on academic probation and not meeting satisfactory academic progress. An explanation for the non-responsiveness of upperclassmen is potentially higher student college and departmental standards. Upperclassmen are typically expected to have a higher GPA within their major, i.e. a 2.7 GPA or 3.0 GPA, and thus lower GPA requirements might not be relevant for most students at this point.

In summary, despite financial implications for freshmen Pell recipients, both Pell recipients and non-Pell students respond similarly in next semester GPA after being placed on academic probation.

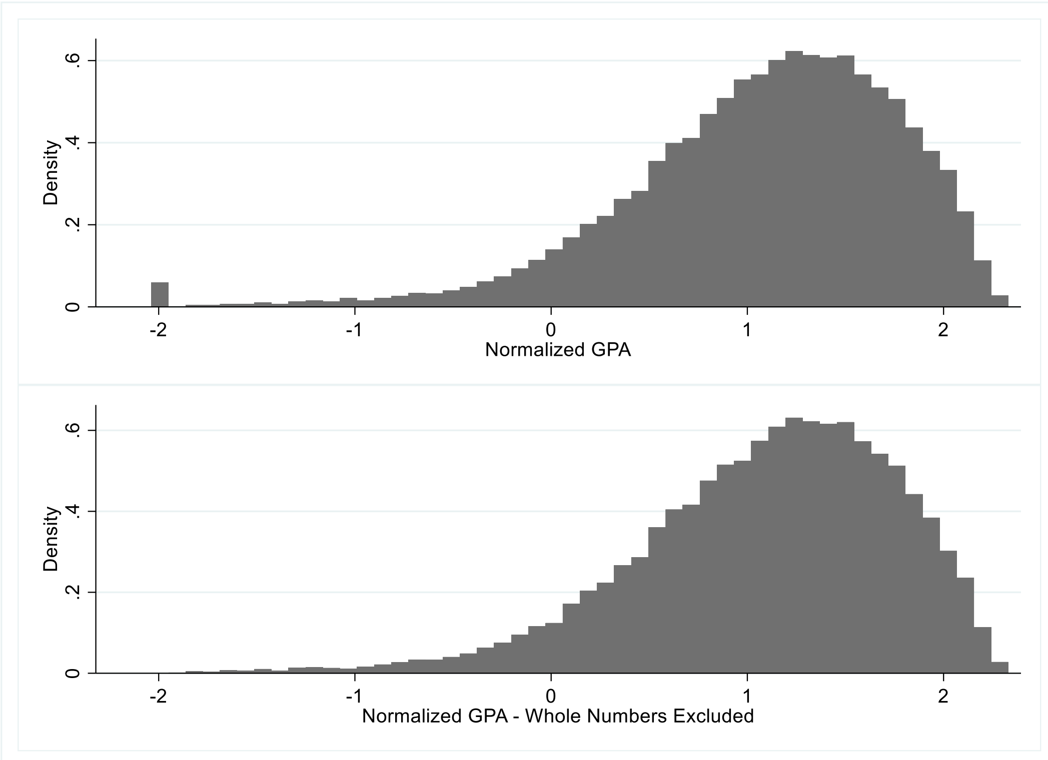
3.6 Discussion

Using a regression discontinuity design, I investigate the impact of academic probation on students in a large, public university in the U.S. southwest. Academic probation's effect on two student sub-groups, those who receive the Pell grant and those who do not, are considered in the analysis. I find that both non-Pell students and Pell students respond to being placed on academic probation by increasing their next semester GPA by approximately 0.2 GPA units. Upperclassmen do not measurably respond to the combined satisfactory academic progress and academic probation GPA cutoff. In a sub-analysis examining graduation outcomes non-Pell students' four, five, and six-year graduation rates are negatively impacted by the academic probation GPA cutoff.

This evidence applies only to students near the margin of being assigned to probation and does not take into account the effect of an academic probation policy for students not near the cutoff. While academic probation does have significant penalties if a student continually performs below the cutoff, with suspension or expulsion occurring after multiple semesters of not meeting academic progress standards, the implications of being placed on academic probation for a non-Pell student are not much. Students receive the notification that they have been placed on probation, are required to meet with an academic advisor throughout the semester, and if they increase their GPA past the cutoff that is the end of consequences. Performance-based scholarship GPA cutoffs, which often occur at higher GPAs, are likely to be more salient to the average student (in this setting a GPA of approximately 3.0) than academic probation is.

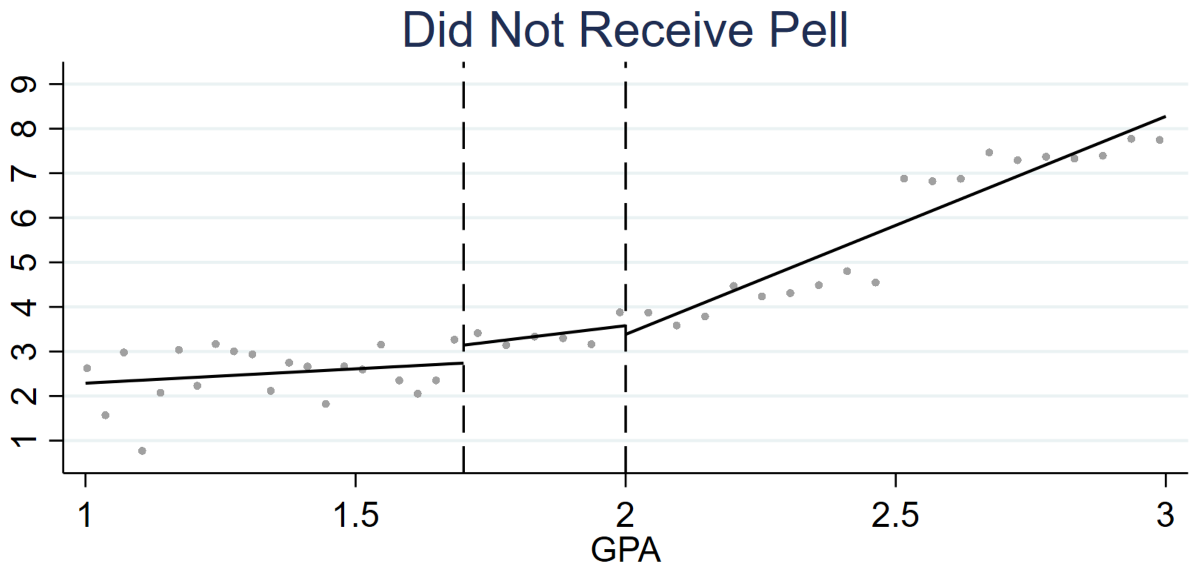
Institutional policy implications come from the main findings, where both Pell and non-Pell freshmen students responded similarly in academic outcomes despite implied financial penalties for Pell students. This would suggest that universities who currently do have financial penalties associated with academic probation are not affecting how a student responds academically in the following semester, and thus the financial penalty is not necessary for GPA improvement. In an extension of this finding, performance-based scholarships that have GPA requirements to continue receiving the scholarship could be relaxed in favor a policy that does not have financial implications for the following semester for students who do not meet the GPA cutoff. Administrators of performance-based scholarships, or financial aid, should consider the goals of a GPA requirement and why financial support should be changed if below that cutoff. If the cutoff is meant to refocus students on academic outcome, this appears to be achievable without a change in financial burden on the student. Alternatively, if the analysis had shown that non-Pell students did not increase their GPA or that Pell students had increased their GPA more than their peers, then an argument could be made to keep financial penalties for low performance (of course limited resources does justify the re-allocation of money to other students who are not on academic probation or not meeting SAP). Upperclassmen are not responsive to the combined GPA cutoff, and thus need to be motivated in other ways. Opportunities for future research include further investigation of institutional financial aid cutoffs, and their effects on student performance. From this analysis a considerable discontinuity can be seen at the 2.5 GPA cutoff for receiving the state supported legislative lottery scholarship, which as mentioned earlier could be more salient for resident students at the university than academic probation GPA cutoffs. A final area of interest not examined here, but similar in policy, is the course grade requirements that exist for upper-division courses within a student's major. Often for these courses students are required to pass with a grade of C, or even B in some cases, to satisfy department graduation requirements. With limited course offerings (such as each year in the fall only for example), these requirements could potentially have large impacts on students' degree choices and time-to-degree for students who are near the grade cutoff requirement.

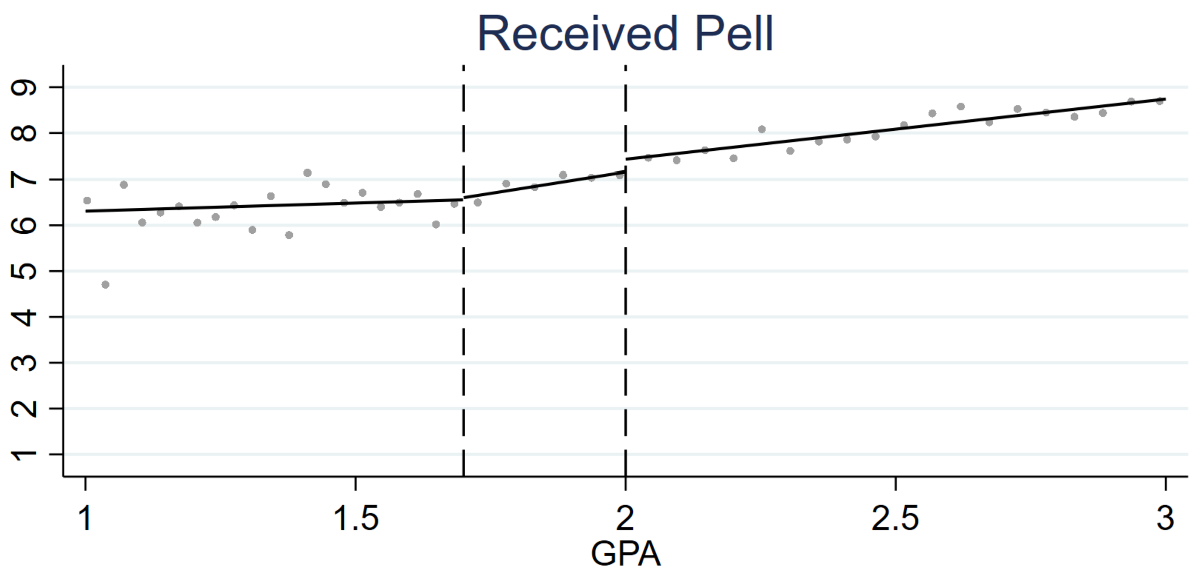
Figure 0.1 Histogram of Normalized Semester GPA



Notes: Top Histogram: Shows distribution of GPA (2.0 is centered at 0.0) for all students first enrolled at the university from Fall 2006 to Fall 2013. Bottom histogram: Shows distribution of GPA with whole numbers excluded.

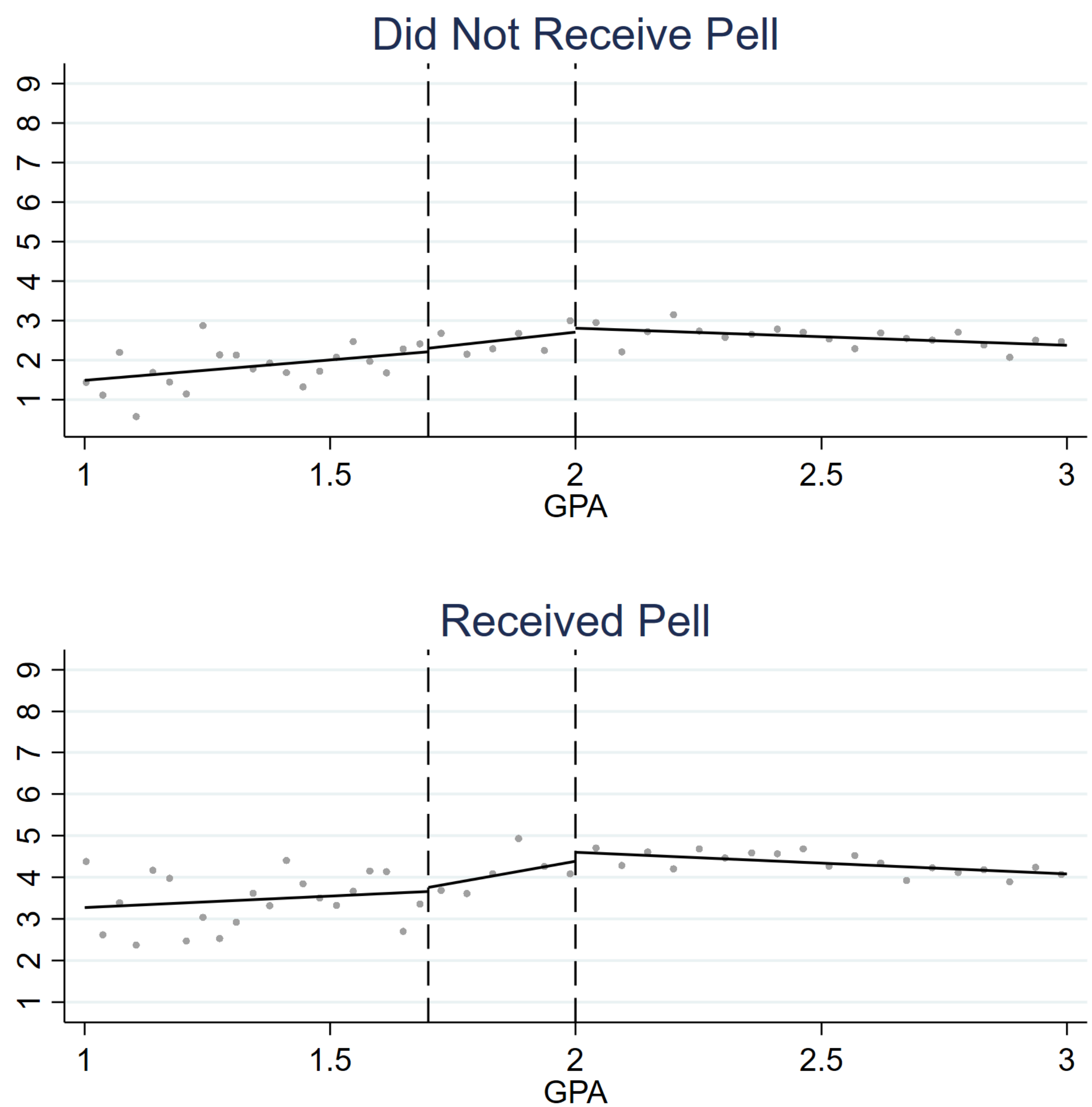
Figure 0.2 Discontinuities in Outcomes for Freshmen – Next Semester Financial Aid (log)





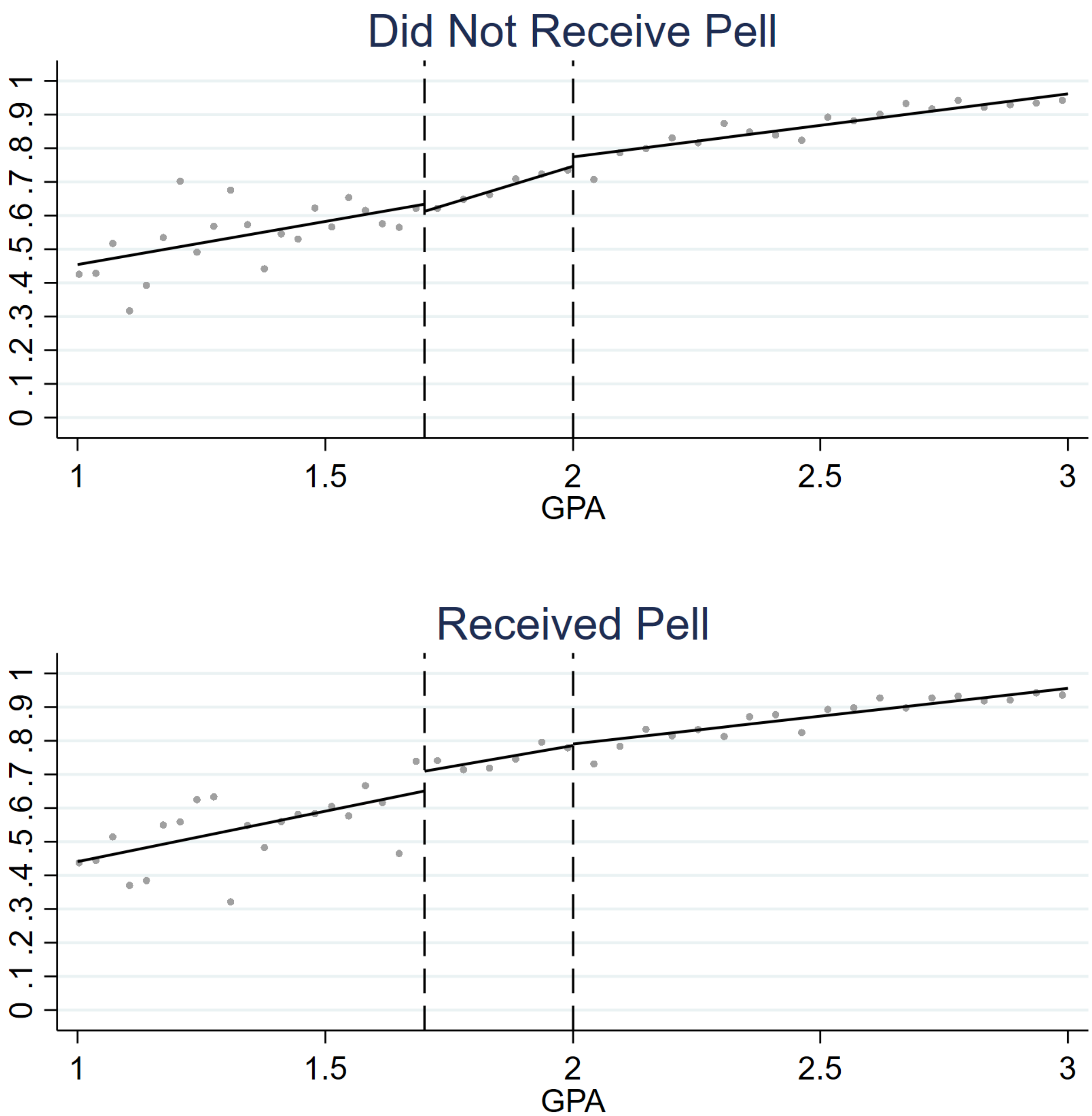
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.3 Discontinuities in Outcomes for Freshmen – Next Semester Loans (log)



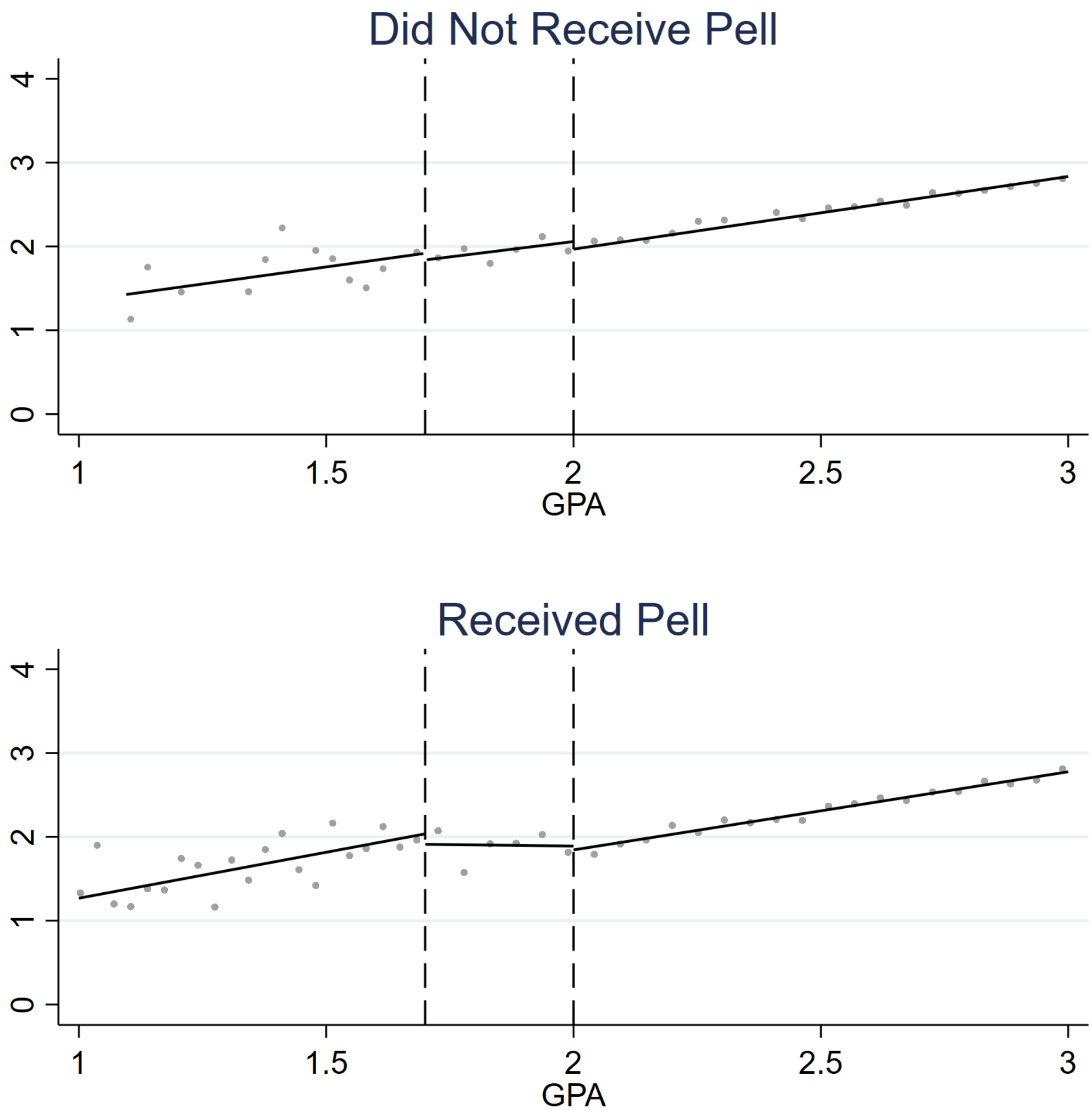
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.4 Discontinuities in Outcomes for Freshmen – Next Semester Enrollment



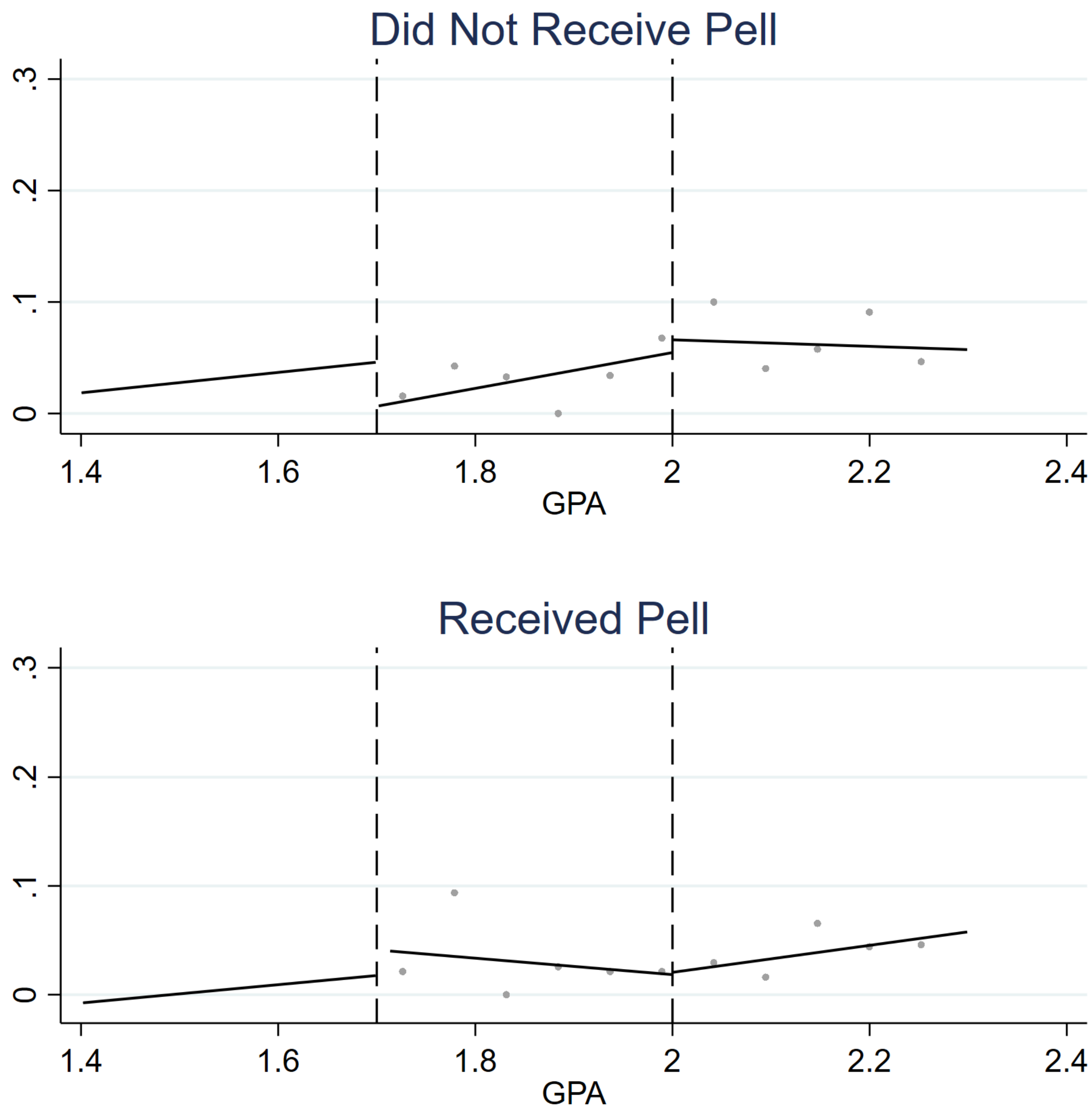
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.5 Discontinuities in Outcomes for Freshmen – Next Semester GPA



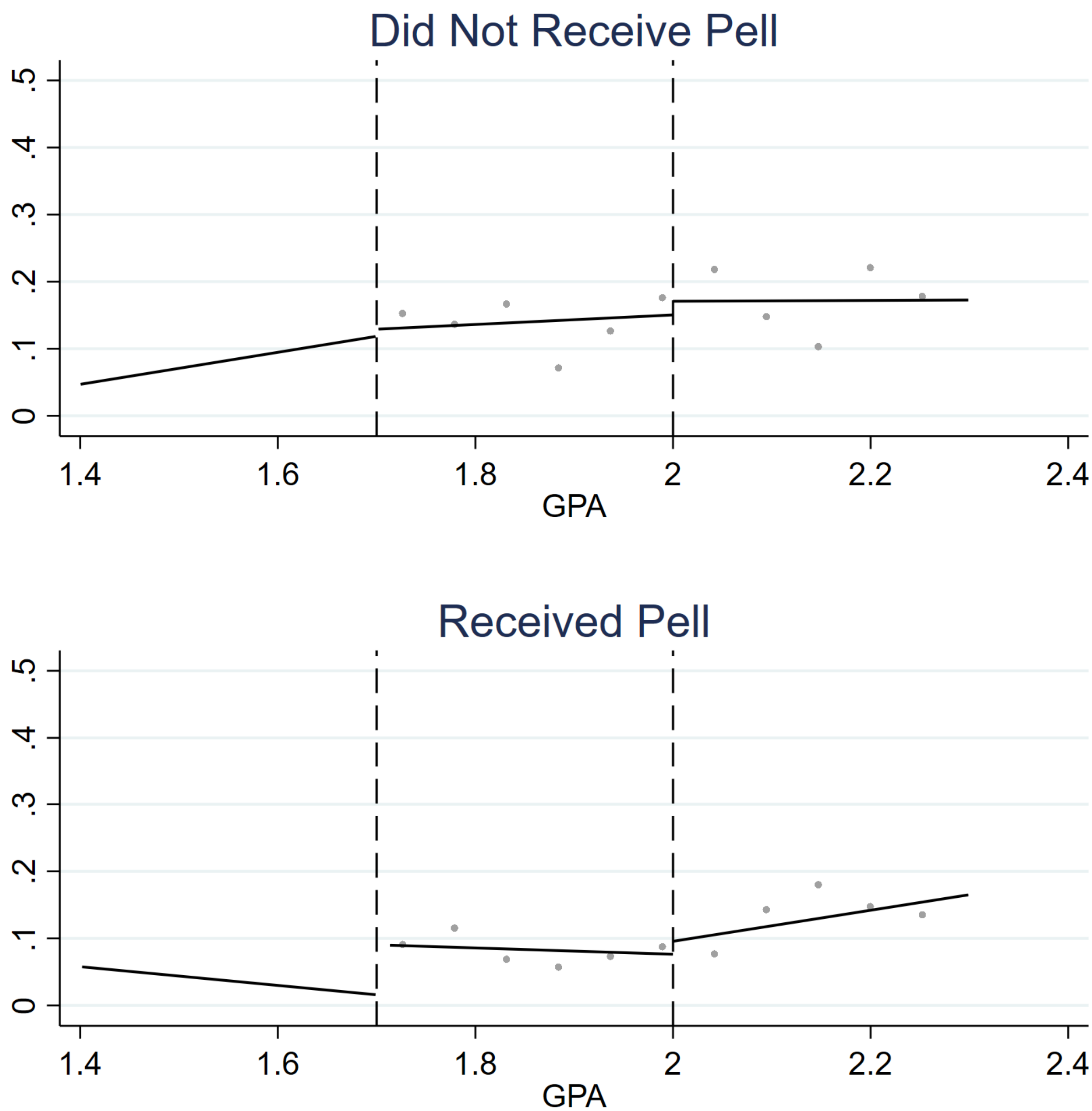
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.6 Discontinuities in Outcomes – Receive Bachelor Degree in Four Years



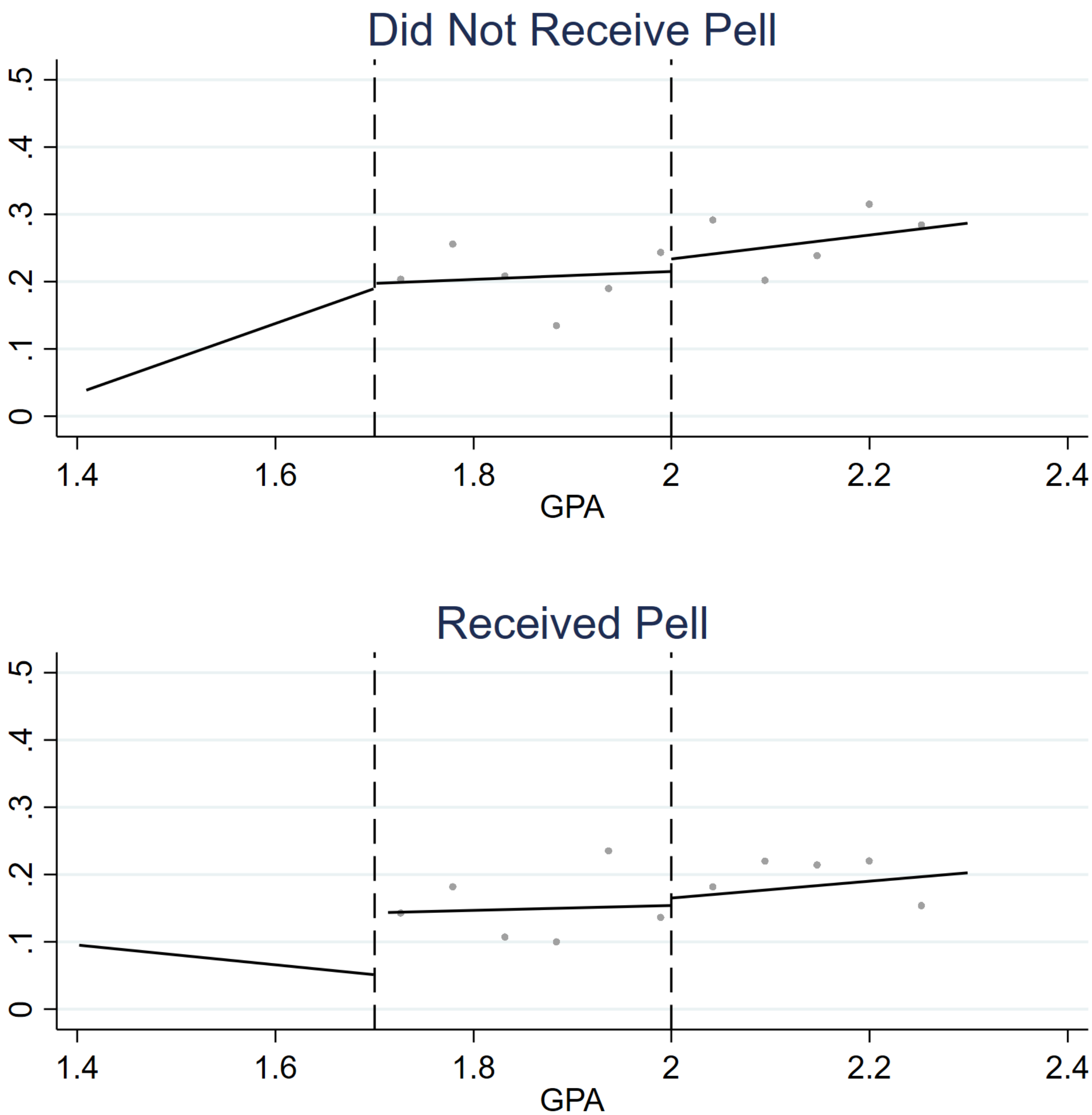
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.7 Discontinuities in Outcomes – Receive Bachelor Degree in Five Years



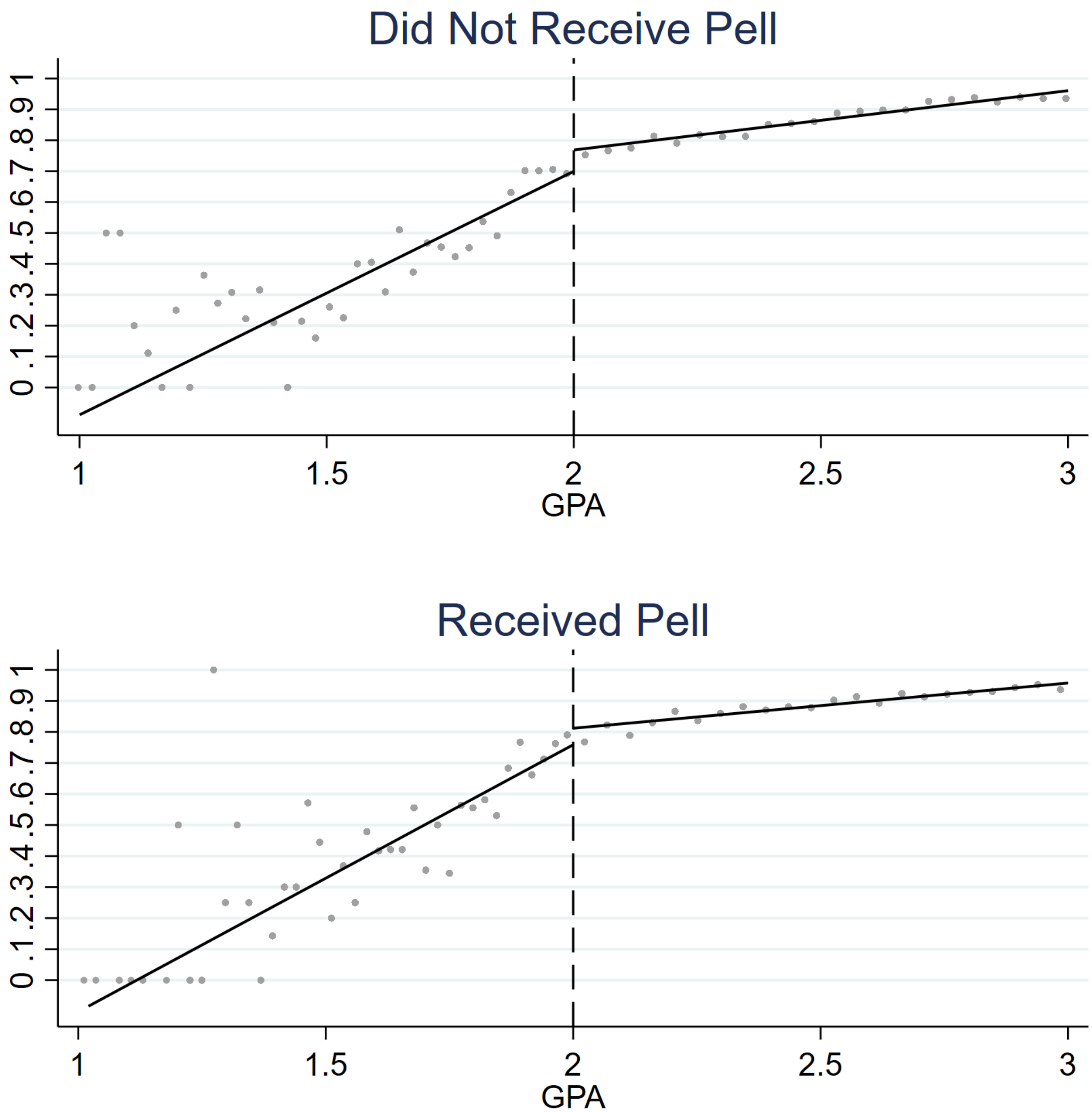
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.8 Discontinuities in Outcomes – Receive Bachelor Degree in Six Years



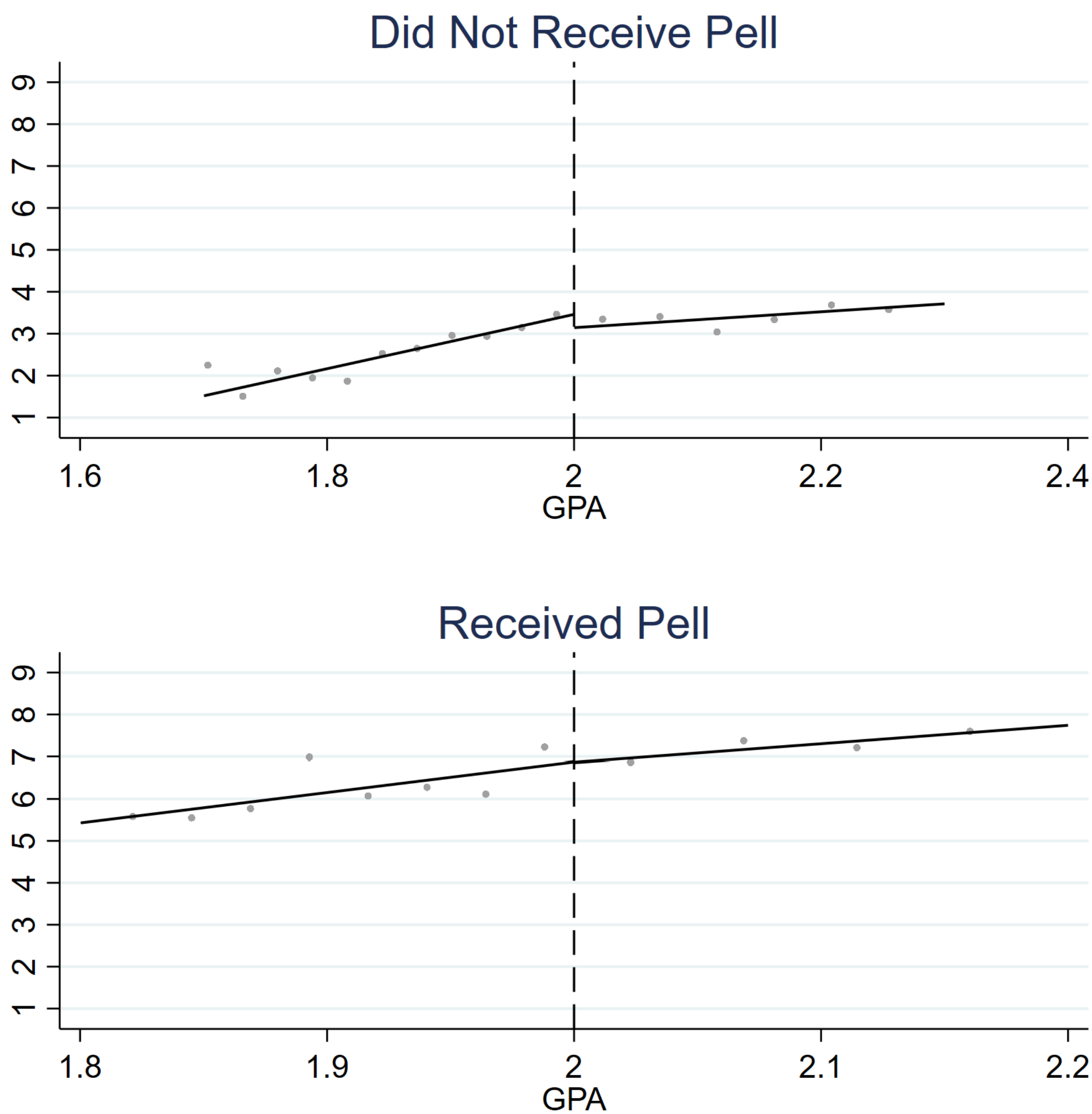
Regression discontinuity results for next semester financial aid offer (log) the semester they were placed on academic probation *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirement, at 1.7 GPA and after being placed on academic probation, which occurs at a 2.0 GPA threshold. Each cutoff is shown with a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using a bandwidth of 0.29.

Figure 0.9 Discontinuities in Outcomes for Upperclassmen – Next Semester Enrollment



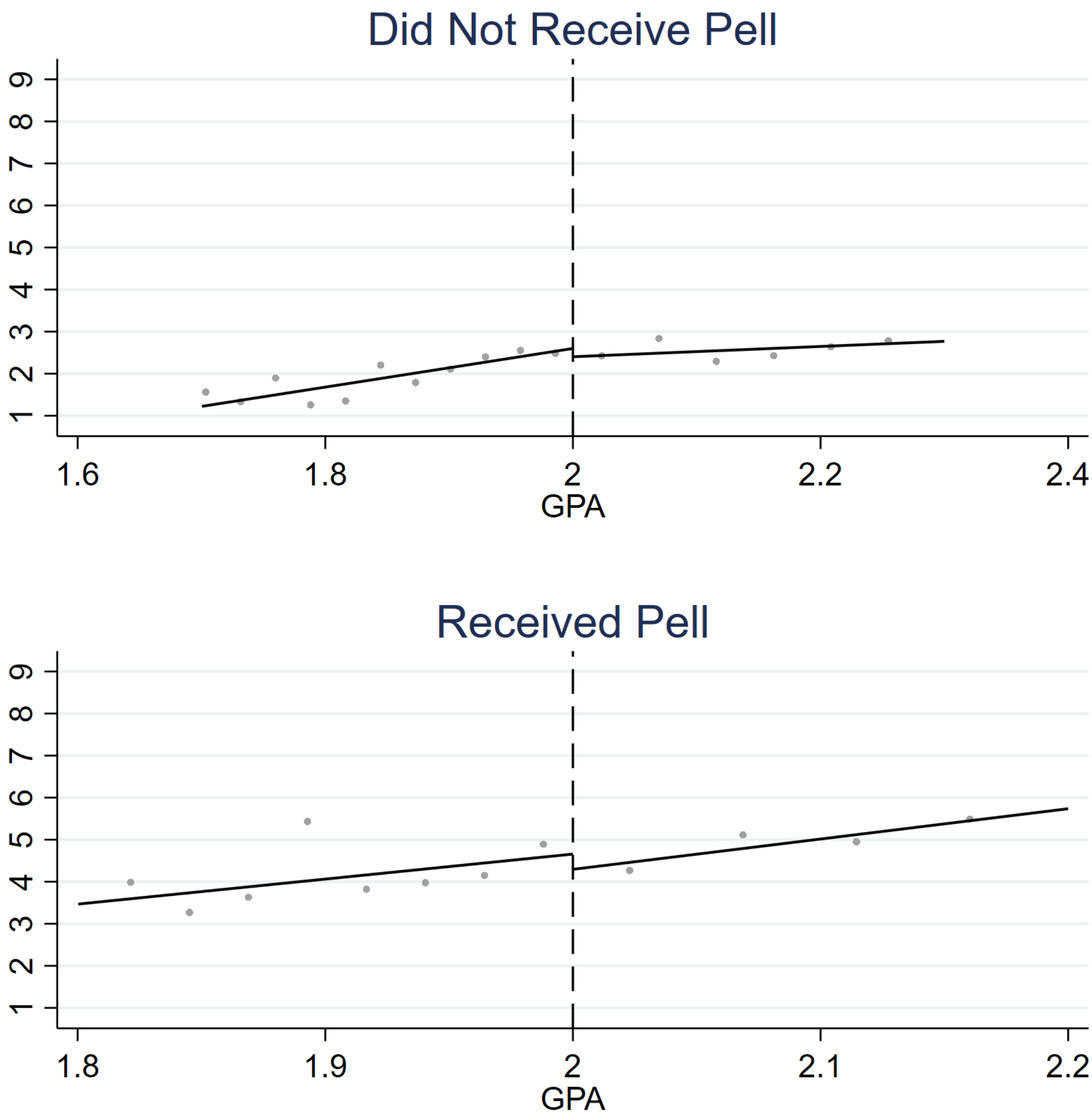
Regression discontinuity results for effect of academic probation after the first semester on six-year graduation. *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of being placed on academic probation and not meeting satisfactory academic progress, which occurs at a 2.0 GPA threshold shown by a vertical dashed line at 2.0. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 0.10 Discontinuities in Outcomes for Upperclassmen – Next Semester Financial Aid



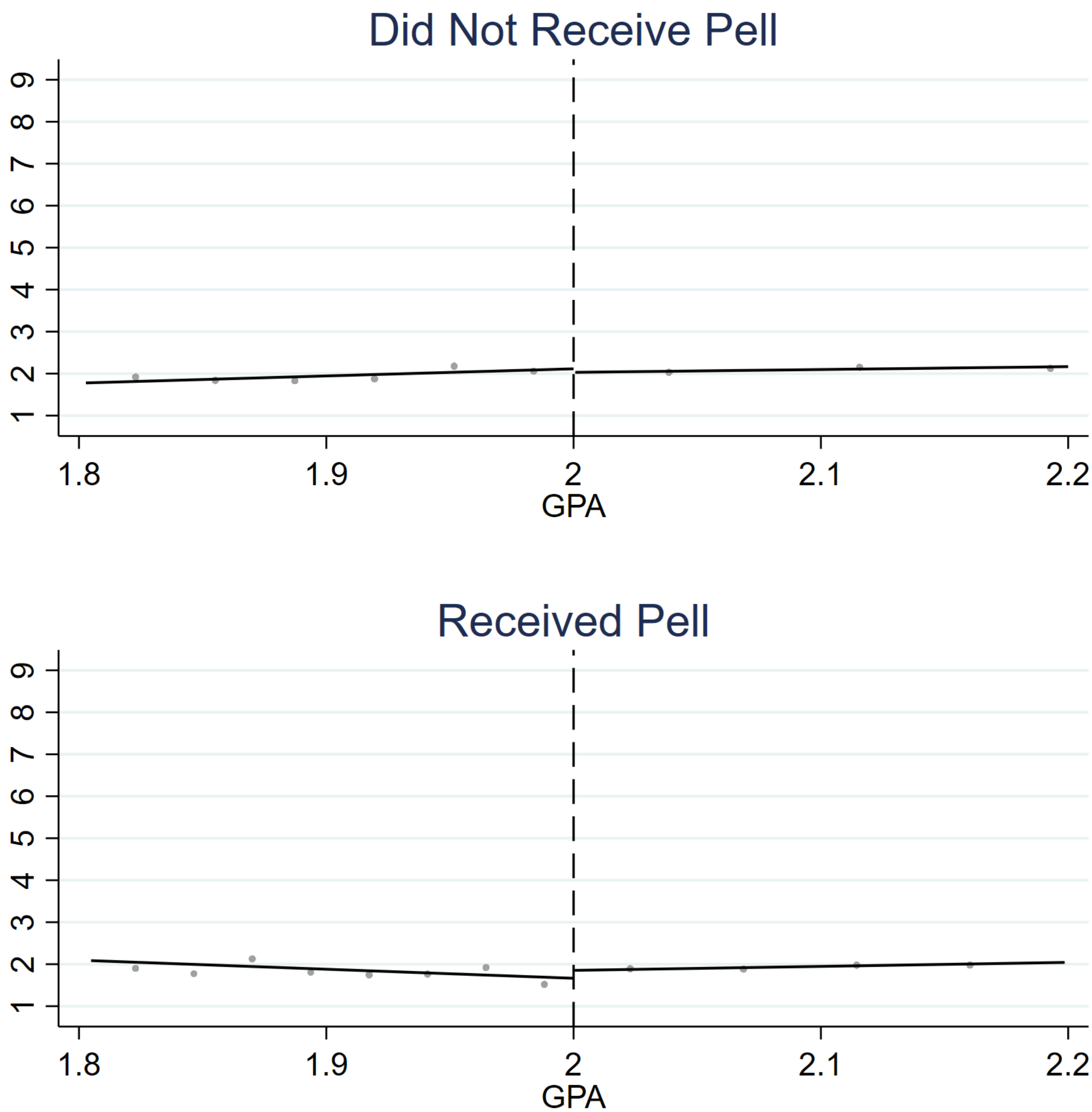
Regression discontinuity results for effect of academic probation after the first semester on six-year graduation. *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of being placed on academic probation and not meeting satisfactory academic progress, which occurs at a 2.0 GPA threshold shown by a vertical dashed line at 2.0. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 0.11 Discontinuities in Outcomes for Upperclassmen – Next Semester Loans



Regression discontinuity results for effect of academic probation after the first semester on six-year graduation. *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of being placed on academic probation and not meeting satisfactory academic progress, which occurs at a 2.0 GPA threshold shown by a vertical dashed line at 2.0. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 0.12 Discontinuities in Outcomes for Upperclassmen – Next Semester GPA



Regression discontinuity results for effect of academic probation after the first semester on six-year graduation. *Notes*: Figures plot conditional average of variable of interest relative to students' cumulative grade point average at time of being placed on academic probation and not meeting satisfactory academic progress, which occurs at a 2.0 GPA threshold shown by a vertical dashed line at 2.0. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Table 0.1 Summary Statistics

	Did Not Receive Pell		Received Pell	
	Mean	SD	Mean	SD
Student Characteristics				
Female	0.545	.498	0.594*	0.491
Native American	0.023	0.149	0.057*	0.232
Asian	0.038	0.191	0.059*	0.236
Black	0.018	0.132	0.029*	0.169
Hispanic	0.434	0.496	0.562*	0.496
Other	0.057	0.232	0.043*	0.204
High School GPA	3.441	0.457	3.377*	0.446
Age At Admission	18.01	0.607	18.02	0.826
Number of Students	8,883		4,632	
Outcomes				
Semester GPA	3.053	0.905	2.909*	0.935
Financial Aid (dollars)	4,439	3,313	7,715*	3,011
Loans Offered (dollars	974.6	1,901	1,770*	1,847
Enroll Next Semester	0.898	0.302	0.889*	0.314
Number of Student Semester Observations	120,547		61,192	

Notes: *Denotes variable is not equal between Pell recipients and those who do not receive the Pell grant at greater than 1% statistical significance in two-way t-tests.

Table 0.2 Covariate Balance – Freshman

	Non-Pell		Pell	
	SAP 1.7 GPA	Academic Probation 2.0 GPA	SAP 1.7 GPA	Academic Probation 2.0 GPA
Female	-0.037	0.050	0.052	0.037
	(0.051)	(0.043)	(0.062)	(0.051)
Native American	0.016	0.014	-0.014	0.042
	(0.024)	(0.019)	(0.043)	(0.032)
Asian	0.016	-0.005	-0.022	0.001
	(0.015)	(0.015)	(0.019)	(0.013)
Black	-0.011	0.007	-0.007	0.008
	(0.019)	(0.016)	(0.028)	(0.024)
Hispanic	-0.052	0.005	0.073	0.035
	(0.051)	(0.043)	(0.062)	(0.050)
Other	0.007	0.039**	0.026	-0.024
	(0.019)	(0.019)	(0.025)	(0.022)
Age at Admission	-0.084	-0.111**	0.082	-0.051
	(0.059)	(0.055)	(0.083)	(0.073)
High School GPA	0.083**	0.059*	-0.015	0.032
	(0.039)	(0.035)	(0.052)	(0.044)
Observations	1571	2405	1,077	1,637

Note: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Bandwidth for each analysis is 0.29.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 0.3 Covariate Balance – Upperclassmen

	Non-Pell Bandwidths			Pell Bandwidths		
	CCT[1]	CERRD[2]	.4	CCT[1]	CERRD[2]	.4
Female	-0.022	-0.036	-0.019	0.000	-0.023	-0.028
	(0.025)	(0.033)	(0.028)	(0.030)	(0.037)	(0.037)
Native American	0.018	0.014	0.021	-0.021	-0.014	-0.021
	(0.016)	(0.021)	(0.016)	(0.023)	(0.029)	(0.025)
Asian	0.005	-0.007	0.006	0.018	0.022	0.010
	(0.010)	(0.014)	(0.010)	(0.013)	(0.017)	(0.015)
Black	-0.002	0.006	-0.005	-0.031*	-0.021	-0.026
	(0.011)	(0.015)	(0.012)	(0.017)	(0.021)	(0.018)
Hispanic	-0.007	0.001	-0.006	0.035	-0.001	0.054
	(0.028)	(0.037)	(0.029)	(0.035)	(0.045)	(0.037)
Other	-0.012	-0.027	-0.009	-0.001	-0.001	-0.002
	(0.014)	(0.020)	(0.011)	(0.014)	(0.019)	(0.014)
Age at Admission	0.024	0.017	0.022	-0.061	-0.010	-0.053
	(0.030)	(0.038)	(0.026)	(0.044)	(0.053)	(0.045)
High School GPA	0.011	0.011	0.012	0.018	0.011	0.014
	(0.026)	(0.035)	(0.023)	(0.028)	(0.035)	(0.029)

Note: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Bandwidth for each analysis are shown in top row (CCT used in primary analysis). [1]Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidth. [2]Coverage error decay rate optimized bandwidth selection from Calonico, Cattaneo, & Farrell (2018).

***Significant at the 1 percent level

**Significant at the 5 percent level

*Significant at the 10 percent level

Table 0.4 Effect of Satisfactory Academic Probation (SAP) and Academic Probation GPA Cutoffs on Next Semester Financial and Academic Outcomes for Freshman

	Non-Pell		Pell	
	SAP 1.7 GPA	Academic Probation 2.0 GPA	SAP 1.7 GPA	Academic Probation 2.0 GPA
A. Next Semester Outcomes for All in Bandwidth				
Next Semester Enrollment	0.010	0.027	-0.0652	0.0353
	(0.051)	(0.038)	(0.059)	(0.044)
Next Semester Financial Aid	-0.420	-0.068	-0.395	-0.175
	(0.403)	(0.344)	(0.445)	(0.326)
Next Semester Loans	-0.044	-0.041	-0.414	0.322
	(0.375)	(0.323)	(0.491)	(0.403)
Observations	1571	2405	1077	1637
B. Next Semester Outcomes for Those Who Enroll Next Semester				
Next Semester GPA	-0.010	0.213**	0.141	0.202*
	(0.144)	(0.100)	(0.162)	(0.122)
Observations	950	1697	690	1199
C. Graduation Outcomes				
Four-Year Graduation	0.034	-0.065*	-0.029	-0.012
	(0.033)	(0.037)	(0.036)	(0.033)
Observations	602	892	388	564
Five-Year Graduation	-0.037	-0.137**	-0.072	-0.074
	(0.061)	(0.057)	(0.056)	(0.067)
Observations	553	819	339	483
Six-Year Graduation	-0.040	-0.130*	-0.065	-0.080
	(0.075)	(0.071)	(0.079)	(0.090)
Observations	504	734	297	420

Notes: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Covariates included in each analysis: female, Native American, Asian, Black, Hispanic, Other, Age at Admission, and High School GPA. Bandwidths are 0.29 GPA units

***Significant at the 1 percent level

**Significant at the 5 percent level

*Significant at the 10 percent level

Table 0.4 Effect of Combined Satisfactory Academic Probation (SAP) and Academic Probation GPA Cutoff on Next Semester Financial and Academic Outcomes for Upperclassmen

	Non-Pell	Pell
	2.0 GPA Cutoff	2.0 GPA Cutoff
A. Next Semester Outcomes for All in Bandwidth		
Next Semester Enrollment	0.012	0.036
	(0.031)	(0.038)
Observations	4,354	2,272
Next Semester Financial Aid	0.250	0.166
	(0.254)	(0.347)
Observations	4,111	1,632
Next Semester Loans	0.180	0.475
	(0.243)	(0.407)
Observations	3,627	1,624
B. Next Semester Outcomes for Those Who Enroll Next Semester		
Next Semester GPA	0.061	-0.227
	(0.010)	(0.141)
Observations	1,829	1,005

Notes: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Covariates included in each analysis: female, Native American, Asian, Black, Hispanic, Other, Age at Admission, and High School GPA. Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths used for each analysis.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Chapter 4 - Charter Schools & School District Finances: How Does Resource Usage Change at the District Level When Charter Schools are Established?

4.1 Introduction

As policymakers in the United States search for ways to improve student educational outcomes and reduce the financial burden of K-12 education on state budgets, school choice policies are regularly championed to address both issues. Support for school choice seems to be growing (Henderson *et al.* 2020) and has had some notable supporters. Famous economist Milton Friedman believed the increased competition would lead to theoretical cost reductions in the educational market as schools competed for students in a more traditional market (Friedman 1997). More recently the Secretary of Education under the Trump Administration, Betsy DeVos, supported school choice initiatives throughout her time as the face of the administration's education policy goals (Strauss, Douglas-Gabriel, & Balingit 2018). Despite the growing support, the body of evidence of school choice's impacts on academic outcome measures can be described as mixed at best. School choice has been shown to improve academic outcomes for both traditional public schools (Hoxby 2003; Dee 1998) and students who choose an alternative education option either through vouchers or charter school enrollment (Egalite & Wolf 2016). These improvements in academic outcomes though appear to come after a "bumpy" transition period of worse performance for students (Hanushek *et al.* 2007; Abdulkadiroglu *et al*. 2018; Gulosino & Liebart 2020). Gains in academic performance are not guaranteed though, with almost just as many compelling studies suggesting no significant differences, or even negative effects, on student performance (Cullen, Jacob, & Levitt 2005; Hanushek *et al.* 2007; Woodworth *et al.* 2015; Gulosino & Liebert 2020).

Aside from student outcomes, public-school district finances are also affected with the movement of students due to school choice. Although most K-12 funding provided by state sources is put through some form of equalization formula (Blagg & Chingos 2017), funding is most often directly tied to student enrollment within the district. Evidence of the effect of school choice on traditional public school district funding has generally shown a decrease in spending per student (Bruno 2019; Ladd & Singleton 2020; Mann & Bruno 2020) overall. The question remains if this decrease in spending is a result of efficiency gains of districts (Buerger & Bifulco 2019), or a reduction of available resources for school districts.

Charter schools date back to the 1990s in the United States. The first charter school law was passed in 1991 in Minnesota, and since then charter schools have been established in most U.S. states and Washington D.C. (Olneck-Brown 2021). Figure 4.1. shows the percentage of all public school students enrolled in charter schools by state for Fall 2016 (U.S. Department of Education 2018). The percentage of public school students enrolled in charter schools varies from state to state, with the United States overall having approximately 6% of public school students attending charters. Delaware reported the highest rate with 44% of all public school students in charters, while six states did not have established charter schools. The growth in charter schools can be explained by the appeal they have to broad groups of students through their flexibility in meeting state standards for curricula and the autonomy to tailor their teaching to specific topics such as STEM, digital media, business, or the arts. In most cases charter schools receive state and federal funding for operation, but they can also be established through private funding in the form of regular donations, foundations, etc..

In this paper I investigate the impact of the most popular school choice policy, the establishment of charter schools on public-school district finances. Charter schools can be established by the public school district, or by external entities such as the state or a non-profit; this analysis considers both district and external charter competition. This is the first analysis to utilize district-level data from across the United States to analyze the impacts of choice on traditional public school district finances. This is impactful for the literature considering the previous lack of generalizability between analyses conducted in a single district (Cullen, Jacob, & Levitt 2005; Hastings & Weinstein 2008; Pathak & Shi 2020) or state (Hanushek *et al.* 2007; Arsen & Ni 2012; Abdulkadiroğlu, Pathak, & Walters 2018; Bruno 2019; Buerger & Bifulco 2019; Ladd & Singleton 2020; Gulosino & Liebert 2020; Mann & Bruno 2020). Estimates from the analysis highlight the considerations policymakers should be making when considering implementing school choice policies, and how funding is allocated for these policies.

I obtained district-level financial data for traditional public-school districts and charter schools from the school year 1998-1999 to 2015-2016 for all fifty states and Washington D.C. from the annual F-33 Financial Surveys required by the Department of Education. This information was then matched to additional district-level data from the Common Core of Data (CCD) provided by the National Center for Education Statistics (NCES). I first present event-study analyses which estimate the effect of the establishment of charter schools for each of the years directly leading to implementation and the years following implementation. I find that enrollment increases in public school districts that establish charter schools. Charter school competition, in which charter schools are established by an entity other than the public-school district, results in a decline in staff and financial assets in the public school district. Overall effects from difference-in-differences style analysis are presented after the event studies for comparability to previous literature. A final analysis is done that estimates the effects of outside charter competition based on the percentage of a county’s students that are enrolled in the charter, i.e. the level of competition. That analysis shows higher levels of charter school competition leads to declines in the number of staff employed by a public school district.

4.2 Methods

4.2.1 Data

I use eight years of financial data from school districts in the United States, from the 2008-09 to the 2015-16 school year. The panel includes 92,847 district-year observations from 13,004 districts in the United States. Districts were excluded from the analysis if they were not a “traditional” school district (school-level codes indicated a district was a vocation or special education system, a nonoperating district that appeared in the dataset for administrative purposes, or an education service agency), did not have a reported Local Education Agency ID (LEAID), could not be matched between the two source datasets, or had charter school competition every year of the panel (thus no variation). The remaining sample covers approximately 80% of all reported districts in the United States The data come from the National Center for Education Statistics (NCES) Common Core of Data (CCD), and the Local Education Agency (School District) Finance Survey (F-33). District level data from the NCES include staffing levels, student-teacher ratios, and location type (urban, suburban, town, or rural). Data from the F-33 includes expenditure, revenue, and whether the district operates at least one charter school. Overall spending per student, instructional spending per student, capital outlay per student, overall revenue per student, and financial assets per student are examined. All financial variables are inflation-adjusted to 2016 dollars.

Descriptive statistics for school districts with an established charter school during the panel and school districts with no charter schools are presented in Table 4.1. Per student spending is significantly higher in districts without a charter school during the eight-year panel, an average of $15,502 versus an average of $13,006 respectively. Standard deviations indicate a wide range of funding levels. This statistically significant higher level of spending extends to instructional spending per student ($7,811 vs. $6,425). Capital expenditures per student are not significantly different between districts with a charter school ($1,246) and districts without one ($1,285). Revenue per student (nearly identical to expenditures per student), and district assets per student ($5,780 vs. $4,703) are also significantly higher for districts without a charter. Enrollment, district teachers, and district staff are significantly higher in districts with at least one charter school, compared with districts that have no charter schools. Lower revenue and expenditure combined with higher levels of enrollment and staffing suggest that districts with charter schools may benefit from economies of scale. Following this pattern, districts with a charter have a higher percentage of English language learners (9.9% vs. 3.9%) and a higher percentage of students who qualify for free or reduced lunch programs (52.5% vs. 44.5%). Districts without charter schools, though have a higher percentage of students with individualized education programs (14.0% vs. 12.2%).

I control for potentially confounding polices and events within states by including state unemployment rates and political characteristics (republican governor, and the share of republicans in state legislatures), available from the University of Kentucky Center for Poverty Research. Lower state house Republican representation is nearly identical for both types of districts (50.7% vs. 50.9%), while districts without a charter have slightly higher upper house Republican representation (53.3% vs. 49.8%). Political unification within the state, measured as a state’s legislature and governor being of the republican party, is similar for each type of school district (39.0% vs. 40.7%). Lastly information on the latest successful school finance litigation within each state was gathered from the National Conference of State Legislatures repositories

(Thatcher 2021). This information is included to control for court-ordered changes in district level spending not captured by other state-level controls mentioned above. Districts without an established charter school were more likely to have had successful school finance litigation during the panel period (53.8% vs. 40.3%).

4.2.2 Empirical Strategy

I use panel data with a difference-in-differences approach to estimate the impact of charter school implementation by comparing changes over time in public school districts that establish a charter school to the contemporaneous changes in districts that do not. Due to differential timing of the establishment of a charter and likely heterogenous treatment effects over time I use the doubly robust (DR) estimand proposed by Sant'Anna and Zhao (2020) and Callaway & Sant'Anna (2021) is used in this analysis. DR flexibly incorporates covariates into a multiple time period difference-in-differences setup with multiple groups, and provides transparent aggregate treatment effects (overall treatment effect and dynamic treatment effects, i.e. event study estimates are both presented in the results). The two-step estimation strategy uses a bootstrap procedure to conduct asymptotically valid inference which can adjust for autocorrelation and clustering. Using the potential outcomes framework, let $Y_t(1)$ and $Y_t(0)$ be the potential outcomes at time t with and without charter school competition. The observed outcome in each period is then expressed as $Y_t = D_t Y_t(1) + (1 - D_t) Y_t(0)$. From this, the average treatment effect on the treated (ATT) is then calculated:

$$ATT_{dr}^{ny}(g,t;\ \delta) = E\left[\left(\frac{G_g}{E[G_g]} - \frac{\dfrac{p_{g,t+\delta}(X)(1 - D_{t+\delta})(1 - G_g)}{1 - p_{g,t+\delta}(X)}}{E\left[\dfrac{p_{g,t+\delta}(X)(1 - D_{t+\delta})(1 - G_g)}{1 - p_{g,t+\delta}(X)}\right]}\right)\left(Y_t - Y_{g-\delta-1} - m_{g,t,\delta}^{ny}(X)\right)\right] \quad (2)$$

where (using the author's notation) the $ATT_{dr}^{ny}(g,t;\ \delta)$ is a simple weighted average of the difference of the outcome variable considered (log expenditures per student, district enrollment, etc.). The first part of the above equation uses inverse probability weighting to place more weight on the most similar school districts with and without charters following Abadie (2005). The second part of the above equation is directly from the outcome regression model proposed by Heckman *et al.* (1997, 1998) where the difference in outcomes over time for school districts with and without charter school competition is calculated (i.e. the difference in differences). Estimates are calculated using the CSDID package in Stata (Rios-Avila, Sant'Anna, & Callaway 2021).

An important assumption for difference-in-differences strategies is the parallel trends assumption, which assumes that treatment and control groups would have the same trends as they did before a policy intervention and thus the difference in the post policy intervention period is truly the response to the policy. In this case, we are assuming that the difference in expenditures per student between public-school districts who establish a charter school and public-school districts without one would continue to be the same had the policy intervention not occurred. To test this assumption, an event study analysis can be done to estimate the effect of the policy change on the outcome variables for the years before and the years after the change. This is shown in the below specification:

$$Y_{dt} = BX_{dt} + \sum_{\tau=1}^{5} B_{-t}\, Charter_{d,t-\tau} + \sum_{\tau=0}^{5} B_{+t}\, Charter_{d,t+\tau} + \delta_d + \omega_t + \varepsilon_{dt} \qquad (3)$$

where B_{-t} is the effect of charter school establishment on public-school district expenditures in the years leading to the charter school (i.e. how public-school district expenditures were in districts who eventually establish a charter school before implementation). If the coefficients on the years leading to implementation are near zero then the parallel trends assumption holds. If not, then other explanations could exist for the continual change of the outcome variable. Due to the importance of the parallel trends assumption, which causal interpretation of the coefficients relies on, event study results are first discussed in the following section. As stated above, the flexibility of the DR estimator allows for multiple calculations of treatment effects, including a dynamic treatment effect (event study). The traditional event study presented above has also come under closer scrutiny lately (Sun & Abraham 2020), and thus dynamic treatment effects from Callaway & Sant'Anna (2021) are presented. Standard errors are clustered at the state level.

4.3 Results

4.3.1 Charter Schools Established by Public-School Districts

The results from the event student analyses are first examined to establish which components of public-school district finances are clearly affected by the establishment of charter schools within a school district, then overall effects measured through difference-in-differences methods are presented. Figure 4.2 shows the results from the dynamic treatment effect analysis of expenditures per student. No evidence is found that expenditures per student is affected after the school district establishes a charter. Figure 4.3 shows dynamic ATT's from an analysis looking at instructional expenditures per student. Instructional expenditures per student appears to decline after a charter school is established within the district. In each of the five years following a charter school being established estimates are all negative, with years 3 through 5 showing a significant decline. This would suggest a shifting of resources away from instructional expenditures to support the establishment of the charter school, yet in the pre-treatment years a decline occurs in the year before treatment. Figure 4.4 shows dynamic ATT's from an analysis looking at capital outlay expenditures per student. In the years following the establishment of a charter school in the school district there does not appear to be a change in capital outlay expenditures per student for the school district. Year 0 and Year 1 are near zero estimates, and though years 2 through 4 have positive estimates they are imprecise.

Thus far, the establishment of a charter school within the school district does not appear to impact district level finances in a consistent manner. Other district characteristics do appear to change though. Figure 4.5 shows a clear increase in the number of students enrolled within the district after the establishment of a charter school. An average 6.5% increase occurs in the first four years following a district establishing a charter school. Examining the pre-treatment period shows nearly zero treatment, the evidence is strongly suggestive of the parallel trends assumption holding. Figure 4.6 has a similar pattern in the number of teachers employed by the district, but post-treatment periods are imprecise. Figure 4.7 shows similar insignificant increases for staff employed in the district in the years following a charter school being established. Two other measures of district finances, revenue per student and financial assets per student, are shown in Figures 4.8 and 4.9. These two components do not appear to be affected by charter school establishment within a district. Revenue per student mirrors closely with expenditures per student, with most districts utilizing all revenue for expenditures each year or only being able to retain small amounts for future budget stabilization. Financial assets also do not appear to be affected, with near zero estimates for most years analyzed.

The analysis of dynamic treatment effects suggest that the establishment of a charter school does not have a clear impact on district level spending. Establishing a charter school within a district though does cause an increase in enrollment within the entire district. The overall effect of each of the event studies is reported in the overall difference-in-differences results in Table 4.2. Panel A shows the corresponding overall estimates from these analyses. We need to turn our attention to only the dependent variables that are supported by the event study analysis. In column 4, enrollment is estimated to increase by 6.5% in the years after a charter school is established.

4.3.2 Charter School Competition within the same County

I now turn to the dynamic treatment effects of charter school competition within the same county as a public school district. In Figure 4.10, estimates of the effect of charter school competition within the county on expenditures per student at the public-school district are presented. Dynamic treatment estimates suggest there is no effect on expenditures per student spending after a competitive charter school is established in the same county. In Figure 4.11 and 4.12 instructional spending per student and capital outplay per student also do not appear to be affected.

Figure 4.13 shows the effect of charter school competition on enrollment within the public-school district. All estimates are imprecise, suggesting a wide range of outcomes after a competing charter enters the same county as a public school district. Teachers employed by the district (Figure 4.14) and staff employed by the district (Figure 4.15) do seem to support this story of students leaving the public school district to attend the outside charter. Staff employed by the district significantly declines in years 1 to 4 following outside competition. Revenue per student (Figure 4.16) does not appear to be affected by outside charter competition. Financial assets per student (Figure 4.17) appear to increase in the years following outside charter competition.

Overall treatment effects are reported in Panel B of Table 4.2. Turning our attention to only the overall estimates that appear to be supported by the dynamic treatment effects analysis, in column 5 I find an overall treatment effect of a 7.8% decline in staff employed by the public school district. In column 8, a significant increase of 30% of financial assets per student is reported.

4.3.3 Alternative Measure of Outside Charter School Competition

A final analysis based on the percentage of a county who enroll in outside charter schools is conducted to determine the extent to which charter school competition is affected by the relative amount of competition from charter schools. In the previous analyses charter school competition was limited to a simple binary indicator if a charter school was established or not, which leads to a difference-in-differences style analysis. This ignores the extent to which students in the area are choosing the charter school compared to traditional public schools in the district (i.e. the binary indicator is the same whether 1% of the county's students enroll in outside charter schools or 10% of the county's students enroll in outside charter schools). Ignoring the level of competition could underestimate the impacts of charter schools with significant enrollments within a county. This issue is explored using a TWFE model in Table 4.3, which examines the same district finance components with the fraction of a county's total enrollment in outside charter schools as the reported independent variable of interest (TWFE in this case has not been documented to lead to biased treatment estimates). Similar to Panel B of Table 4.2, expenditures per student in column 1 of Table 4.3 does not appear to be impacted by the fraction of students enrolled in outside charter schools in the county. Instructional expenditures per student (column 2), capital outlay per student (column 3), enrollment (column 4), and teachers employed by district (column 5) also do not appear to be affected by this measure of charter school competition. The number of staff employed by the public school district (column 6) significantly declines as the fraction of students in the county enroll in outside charter schools. A ten percentage point increase in the fraction of students enrolled in outside charter schools leads to a 8% decline in the number of staff employed by the public school district. This follows in line with the findings in Panel B of Table 4.2. Revenue per student (column 7) and assets per student (column 8) are not affected as outside charter school competition increases.

4.4 Conclusion

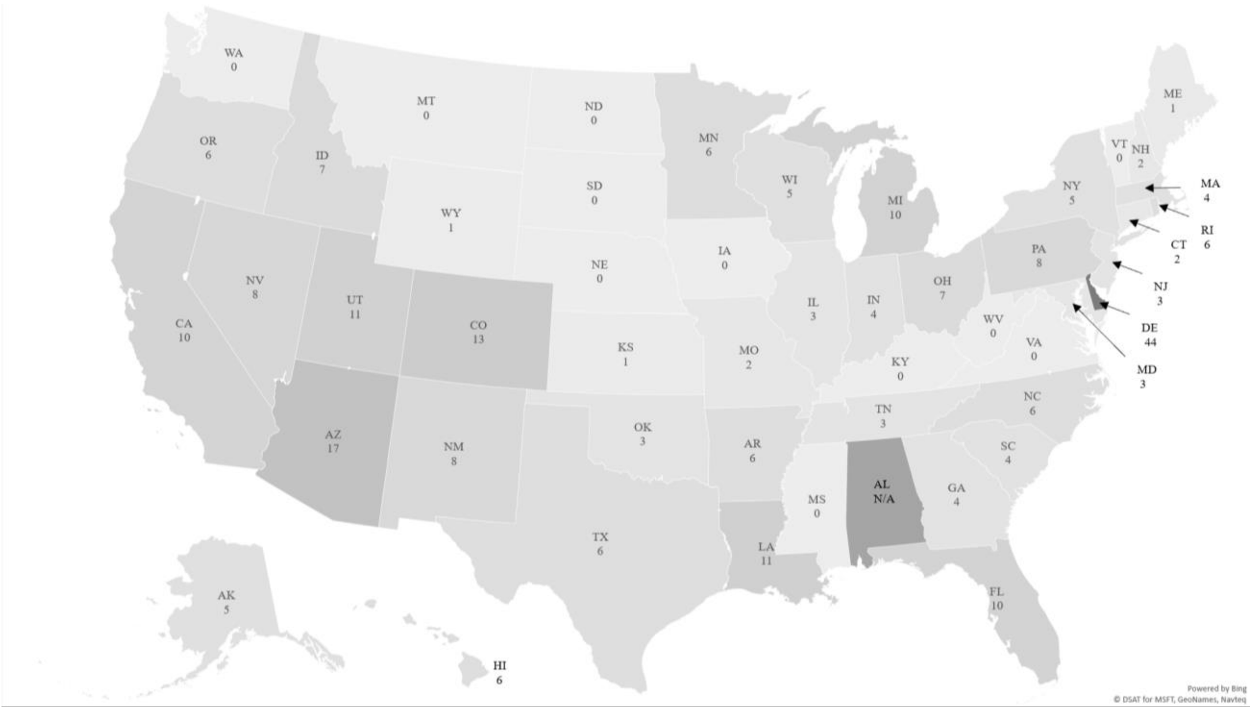
This analysis explores the effects of charter school competition on public school district finances. This analysis is the first to use F-33 panel data from the NCES that covers public-school districts across the entire United States. Both event study designs and difference-in-differences methodology are used to analyze the effects of charter school competition on public school district finances. After the establishment of a charter school within a public school district enrollment within the district increases in the following years.

When examining the effects of a competing charter school being established in the same county as the public-school district, public-school district finances appear to be largely unaffected by the competition. The introduction of a competing charter does lead to a decrease in staff employed by the district and financial assets per student in the following years. An alternative measure of charter school competition, the fraction of students in a county who enroll in outside charter schools, finds significant declines in the number of staff employed by the public school district. This suggests that as the level of charter competition increases, the larger the effects on traditional public schools within the county. These findings call for state policymakers to carefully consider the potential advantages, and disadvantages, to all students when considering charter school policies. While no evidence is presented to impacts of these polices on academic achievement, the increases in enrollment suggest a re-allocation of resources could be needed to the traditional public school districts who are losing these students to charter schools. Alternatively, if the increased enrollment is a result of students who previously attended private schools enrolling in the newly formed charter school, then this represents increased usage of public resources. Investigating student-level enrollment data would answer this question.

Community leaders also need to understand where the additional students are coming from when a public school district establishes a charter school. If the increased school choices are attracting outside families through Tiebout sorting, where families choose the communities they live in based on the public amenities they prefer, other community resources could also be experiencing increases in use. Increased population density over time or significant changes of net migration into these communities are two future research opportunities that do not need student-level data to be explored.

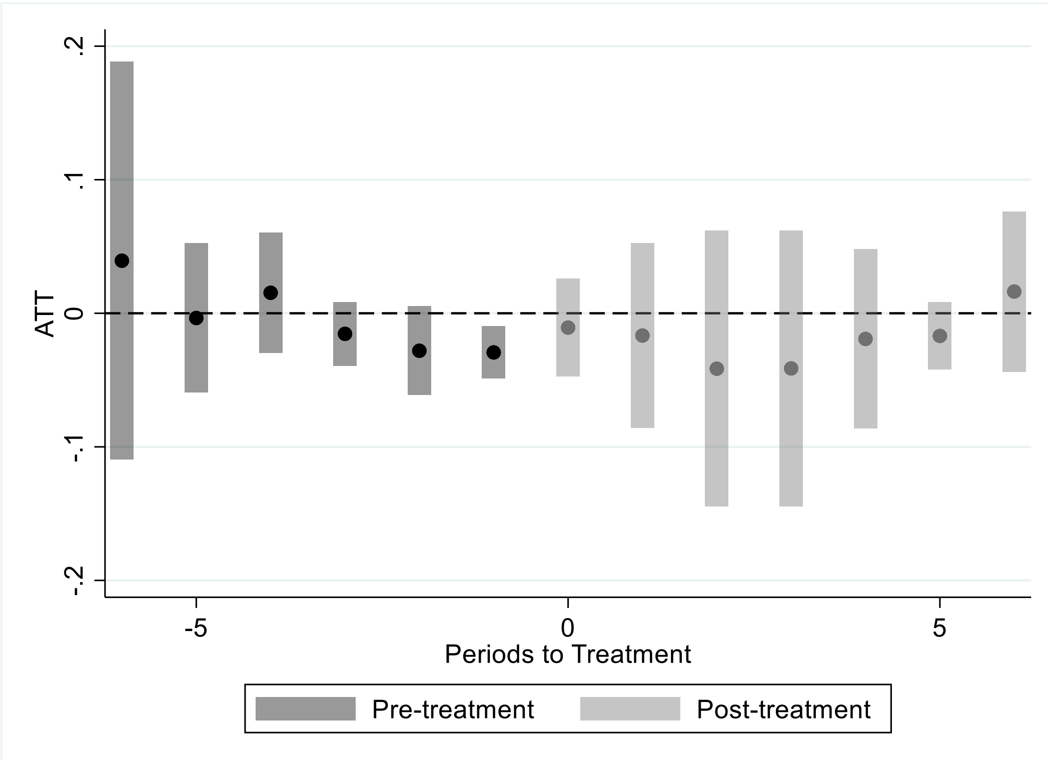
One of the primary goals of this analysis was to increase the generalizability of its findings to more school districts than previous analyses of single districts or states. While this goal was achieved, the unique characteristics of each state policy likely explain some of the inconsistencies between the results. This is a tradeoff that most researchers are familiar with. This study also had other limitations and tradeoffs. The primary of these is the lack of incorporating student achievement outcomes into the research question. The overall answer to the effects of these policies needs to consider academic achievement of students who leave because of school choice and those who stay, and that question would be best answered at the student level. An additional limitation of the analysis is the lack of information on state inter-district transfers. Some states allow for students to transfer from their assigned public school district to another school district without moving, which is fundamentally a school choice policy. While this information is tracked by districts, it is not reported at the national level and thus its effect is unknown. Two characteristics though should alleviate concerns that inter-district transfers could be confounding results found here: (1) inter-district transfers are not allowed in all states, and (2) inter-district transfers in most cases must be approved by the receiving district. The degree to which an inter-district transfer affects the receiving district's finances also varies by state, with some states having a higher degree of funding tied to the student (Arsen & Ni 2012) and others having a higher weight placed on local property tax revenues. A future opportunity for research is to examine differential charter effects by rural-urban classification. Little has been said about rural students who realistically do not have choices even after state policy gives them the right to choose. Even in urban school districts, location is still an important factor when a student or family is selecting a school for attendance (Hastings & Weinstein 2008).

Figure 4.1 Percentage of all public school students enrolled in public charter schools, by state: Fall 2016



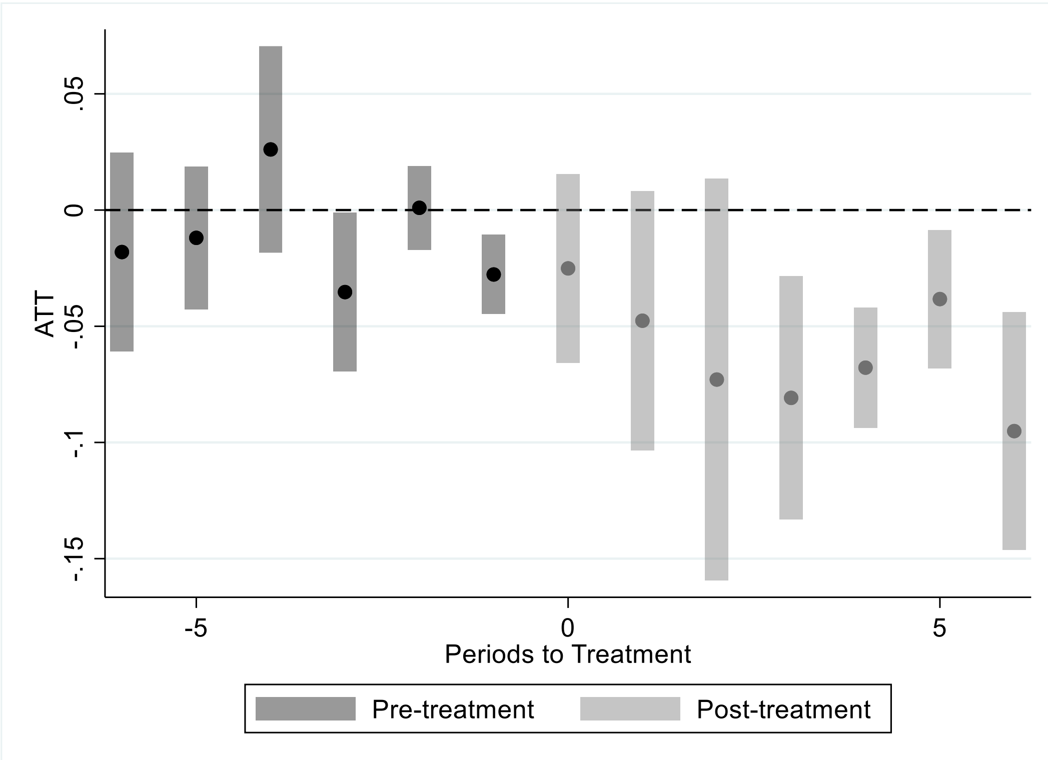
Source: U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 2000-01 through 2016-17. Figure generated in Excel by author

Figure 4.2 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Expenditures per Student



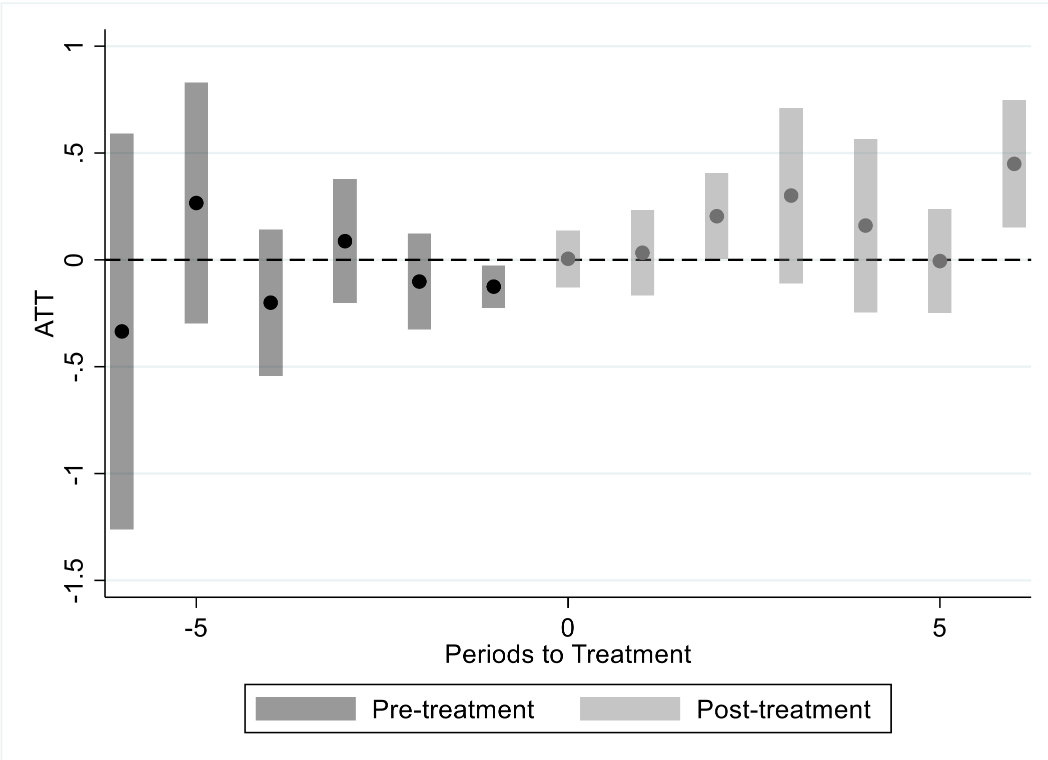
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, expenditures per student, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.3 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Instructional Expenditures per Student



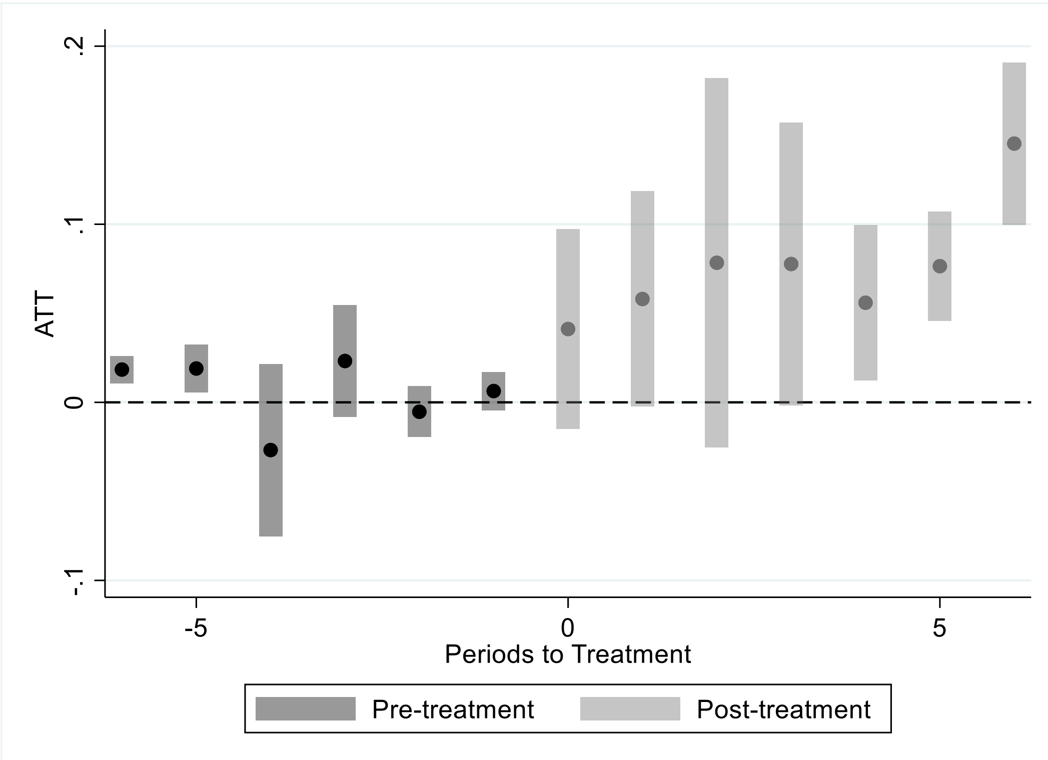
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, instructional expenditures per student, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.4 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Capital Outlay Expenditures per Student



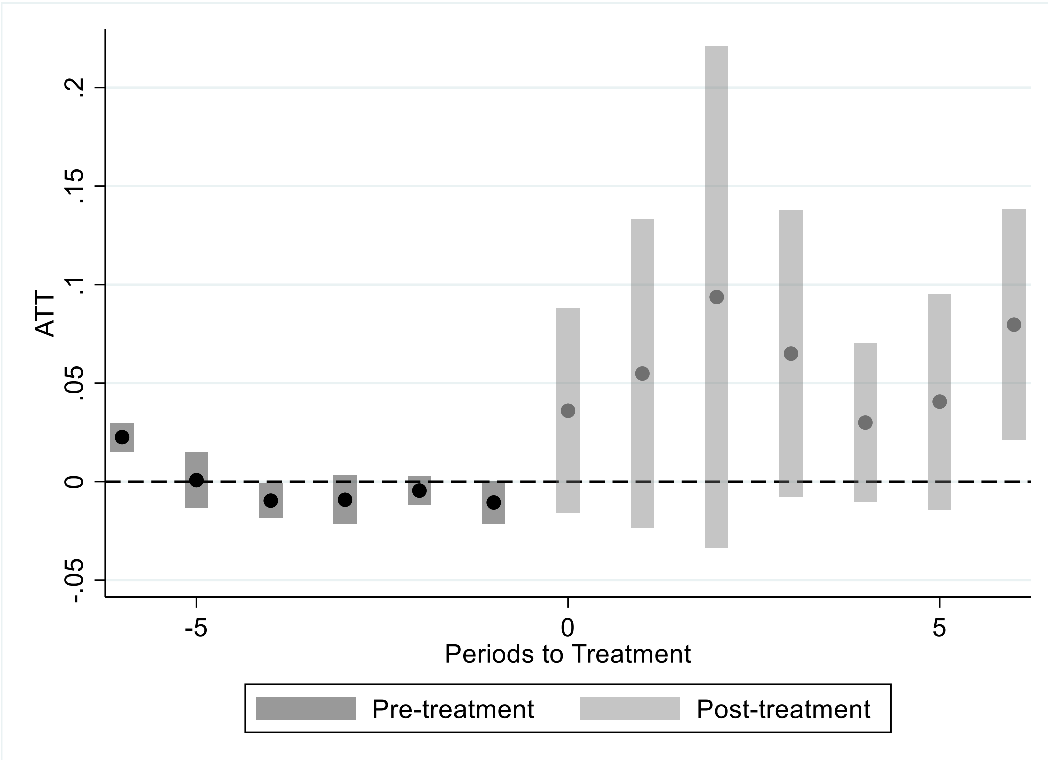
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, capital outlay expenditures per student, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.5 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Enrollment



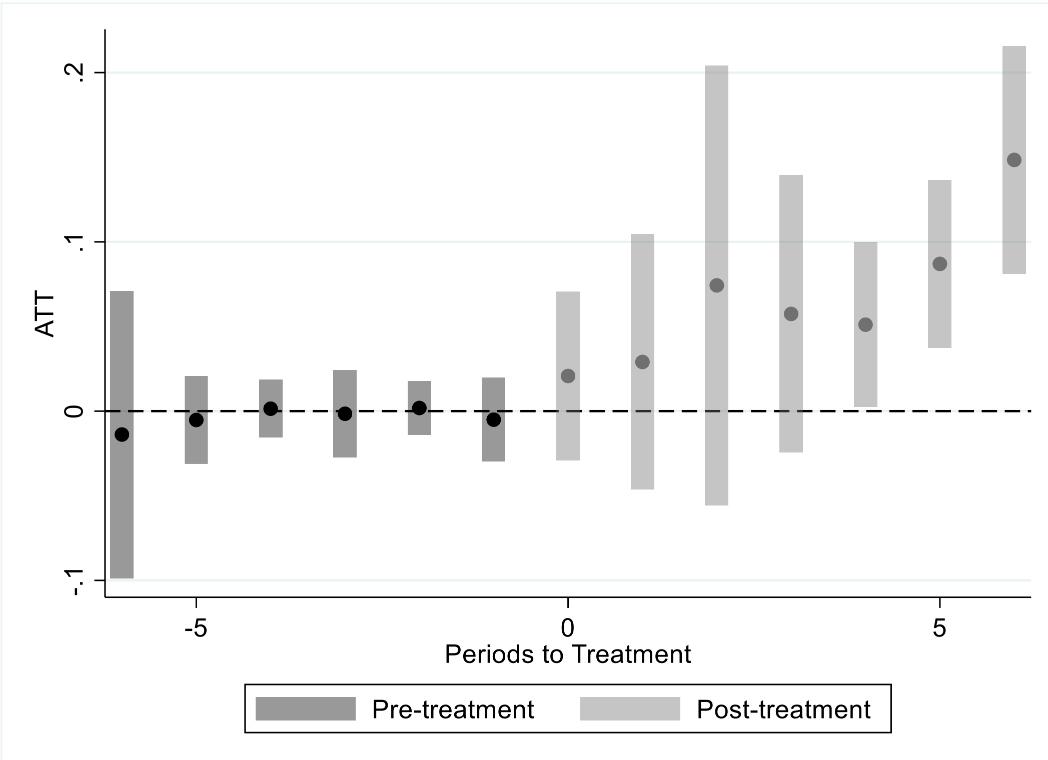
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, enrollment, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.6 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Teachers in District



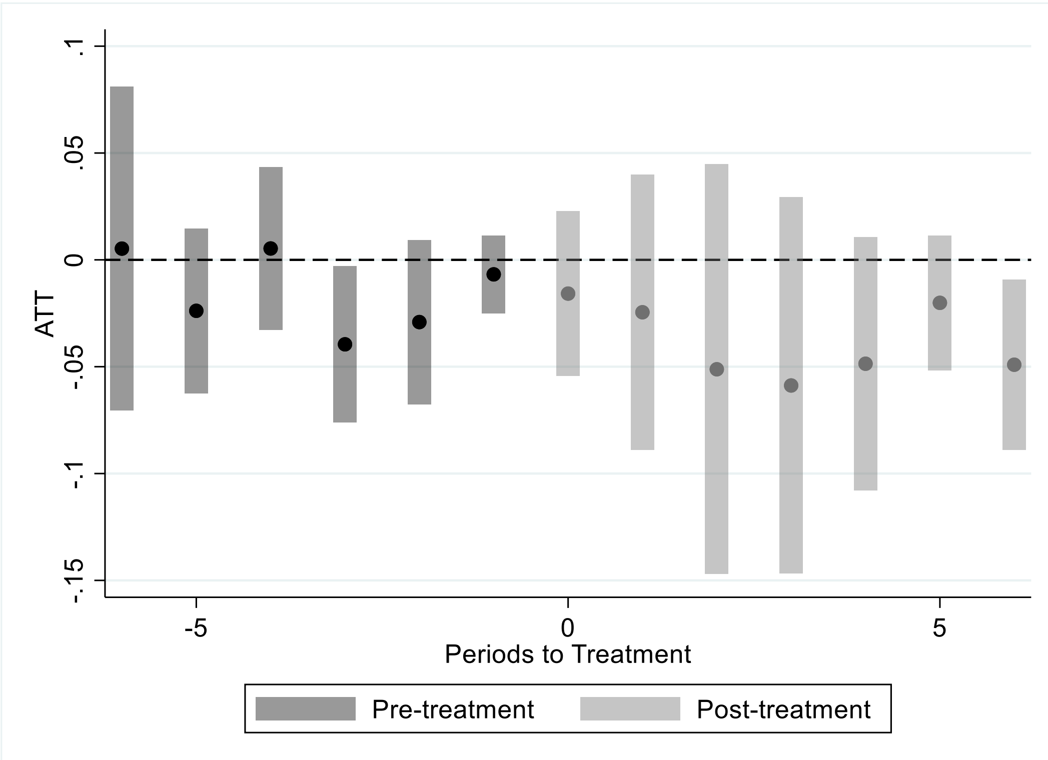
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, teachers, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.7 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Staff in District



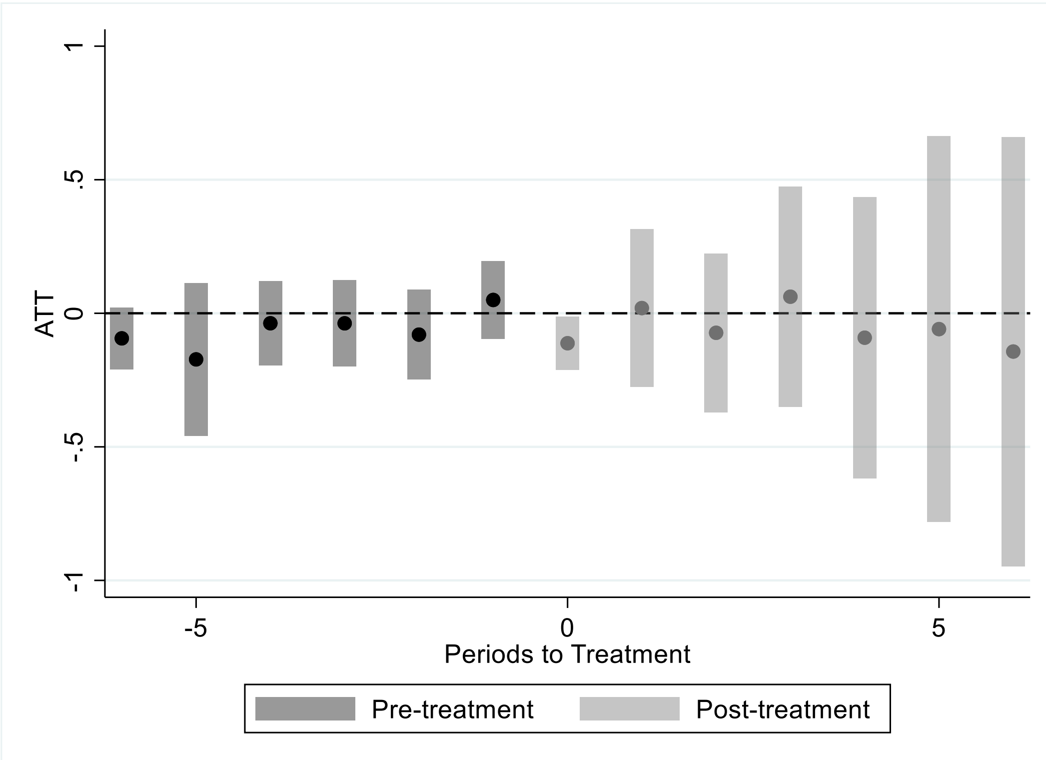
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, staff, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.8 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Revenue Per Student



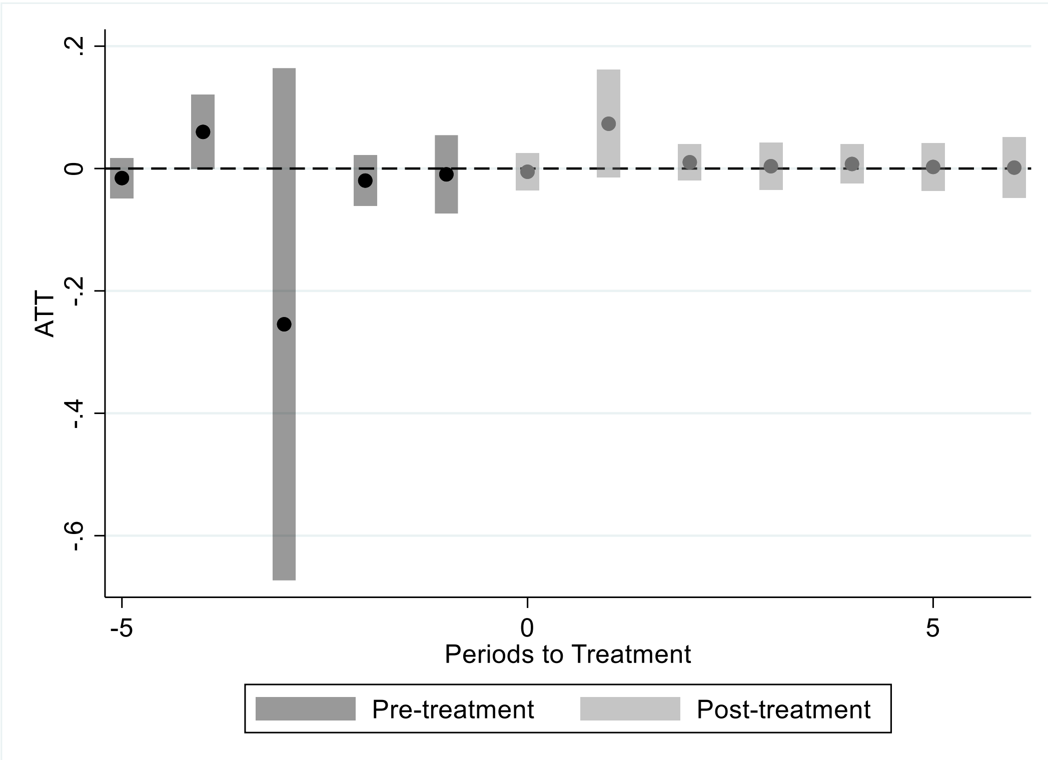
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, revenue per student, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.9 Event Study – Dynamic Effects of Average Treatment Effect (Charter School Established by District) on the Treated – Assets Per Student



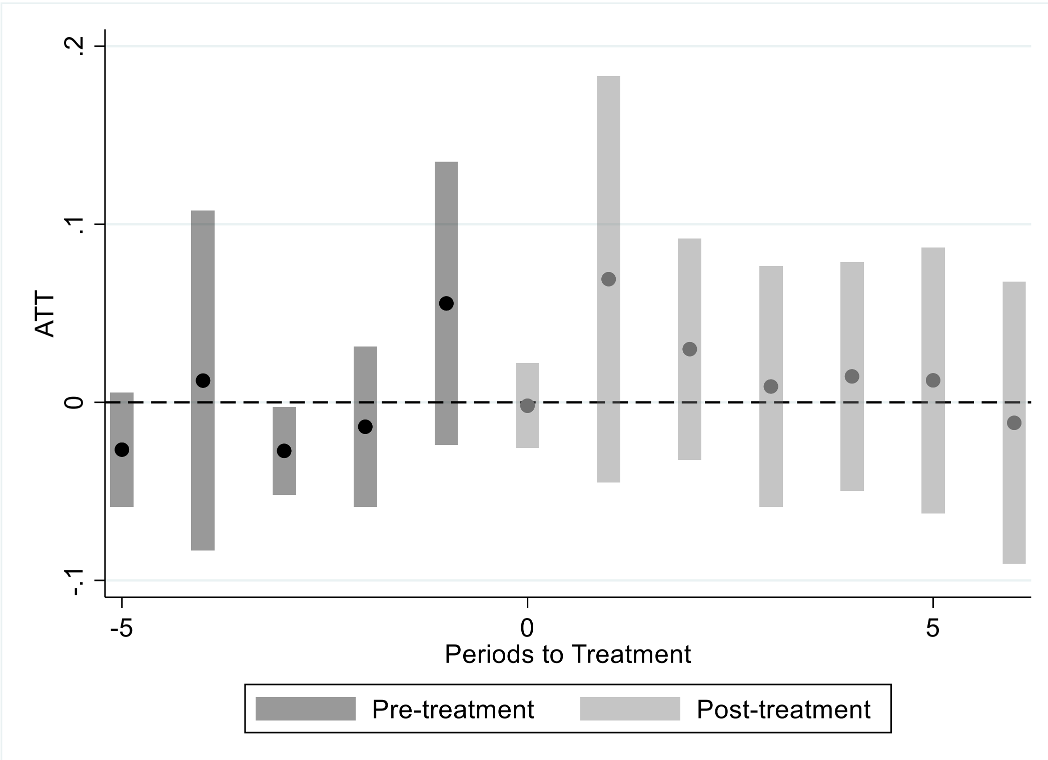
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, assets per student, separately for each period and group. Treated group consists of districts that establish a charter school, the control group of districts who do not and not yet treated charter school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.10 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Expenditures Per Student



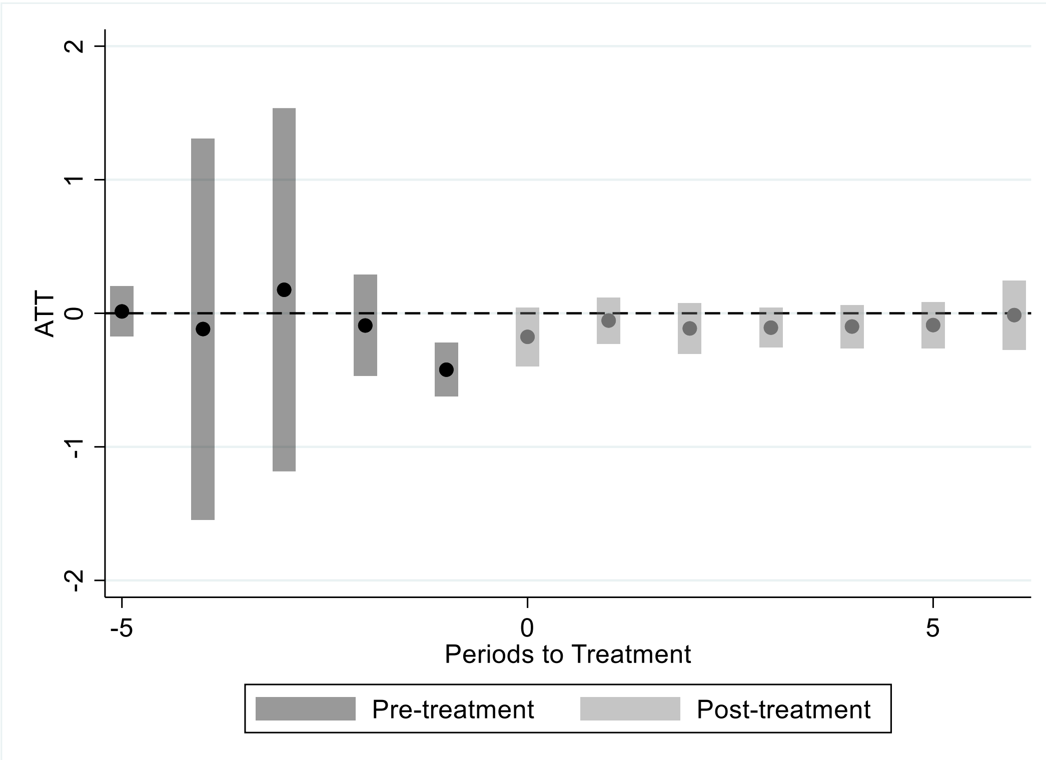
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, expenditures per student, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.11 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Instructional Expenditures Per Student



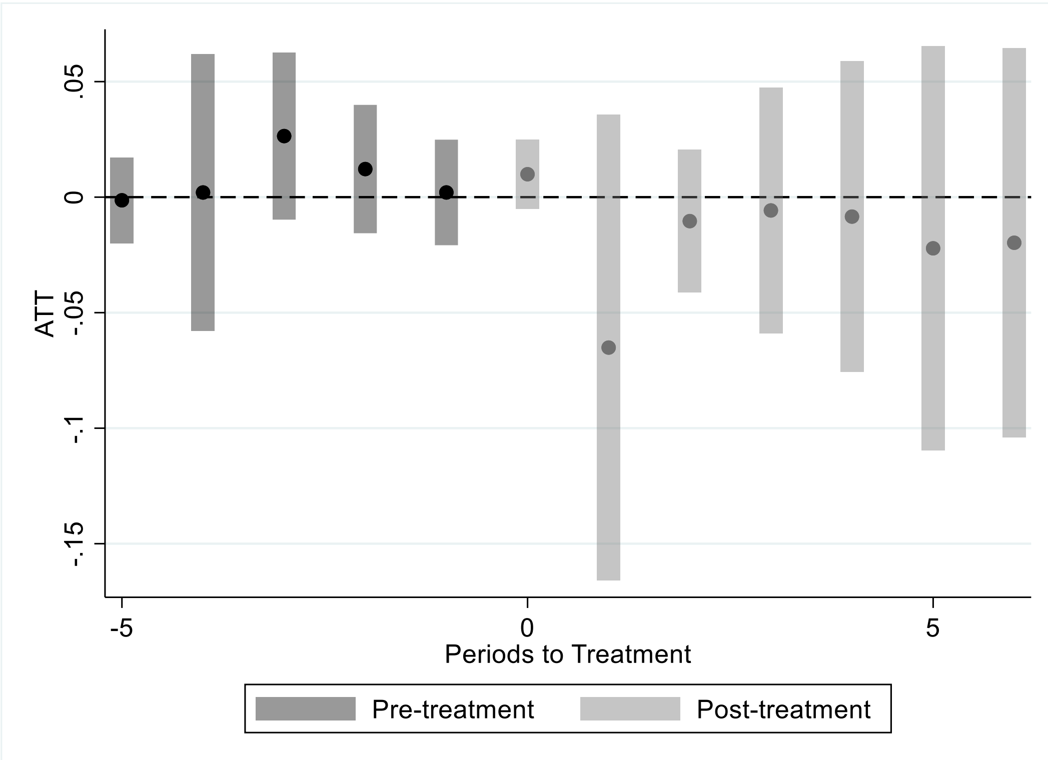
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, instructional expenditures per student, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.12 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Capital Outlay Per Student



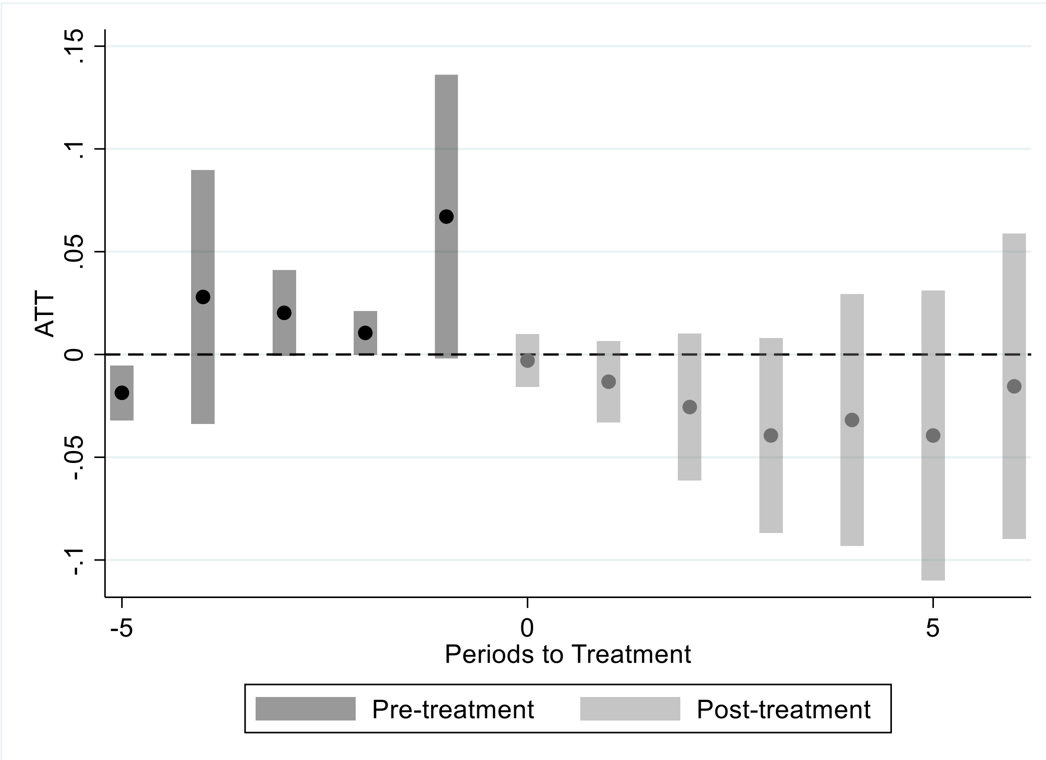
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, capital outlay expenditures per student, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.13 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Enrollment



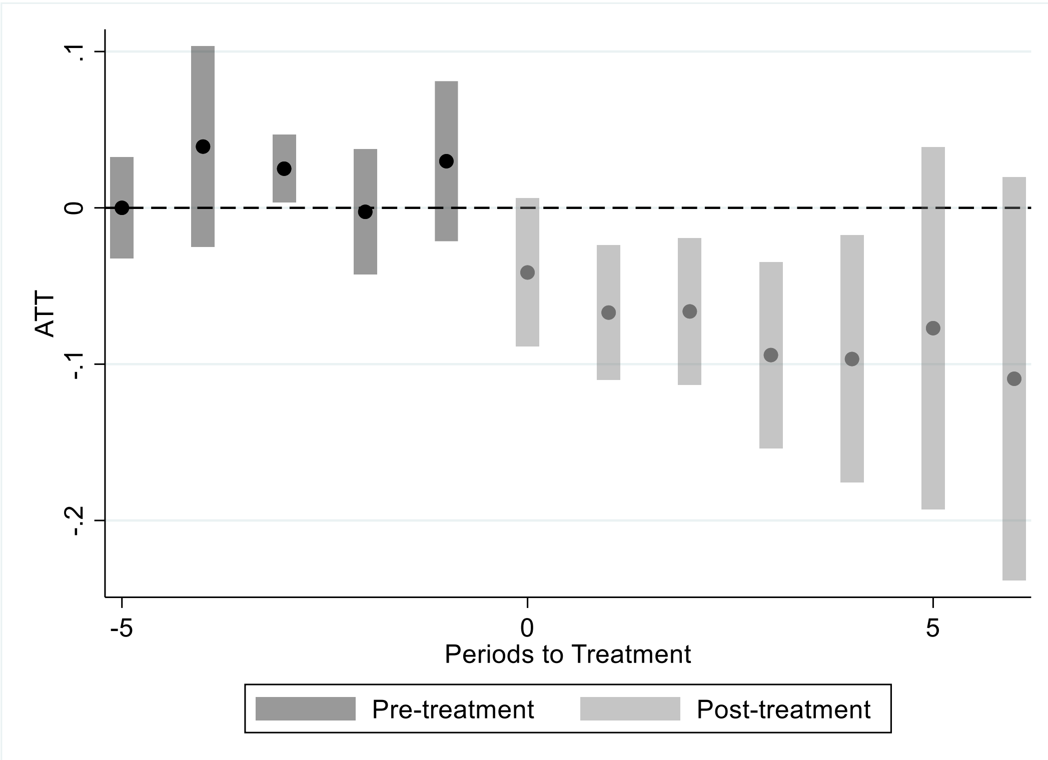
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, outlay, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.14 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Teachers in District



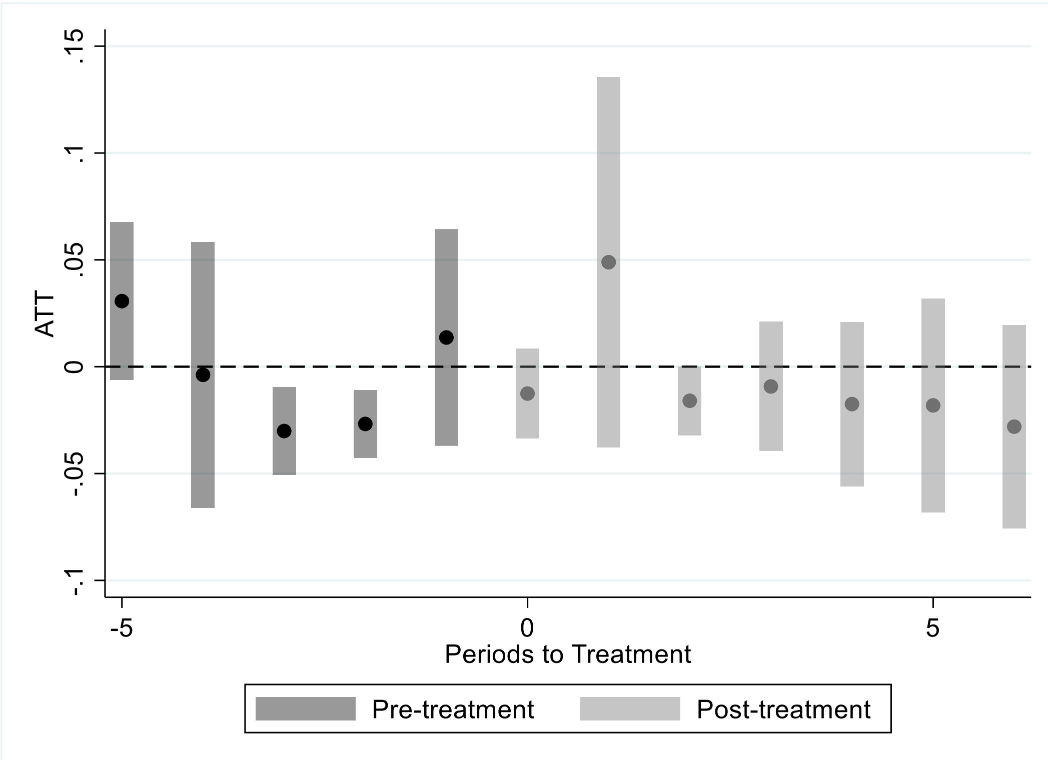
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, teachers, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.15 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Staff in District



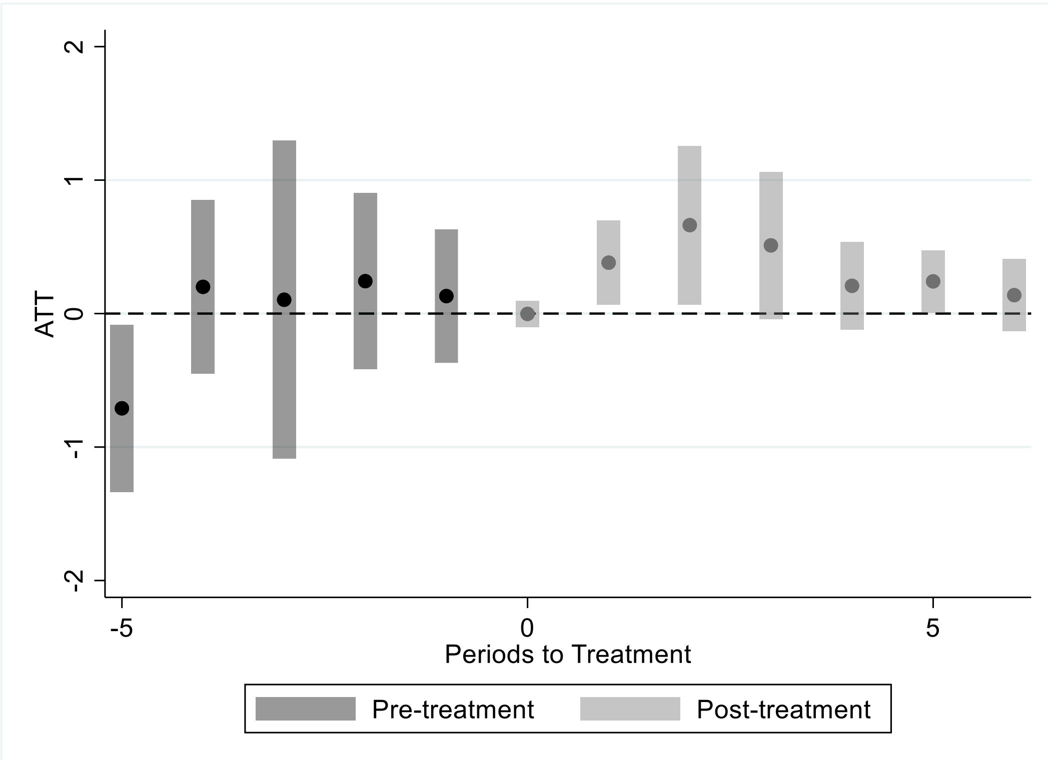
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, staff, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.16 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Revenue Per Student



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, revenue per student, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Figure 4.17 Event Study – Dynamic Effects of Average Treatment Effect (Competing Charter School Established in Same County) on the Treated – Assets Per Student



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, assets per student, separately for each period and group. Treated group consists of districts that have a competing charter in county, the control group of districts who do not and not yet treated charter competition school districts. Covariates include: log enrollment, state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP

Table 4.1 District Level Summary Statistics by Districts without a Charter School & those with a Charter School

	No charter within District			Charter within District		
	N	Mean	(SD)	N	Mean	(SD)
Expenditures Per Student	86,784	15,502	9,591	6,063	13,006*	5,133
Instructional Per Student	86,784	7,811	4,454	6,063	6,425*	1,835
Capital Expenditures Per Student	86,784	1,285	3,131	6,063	1,246	2,668
Revenue Per Student	86,784	15,605	9,511	6,063	12,976*	4,816
District Assets Per Student	86,784	5,780	13,091	6,063	4,703*	5,344
Enrollment	86,784	2,722	5,724	6,063	11,294*	25,808
District Teachers	86,784	176.8	358.8	6,063	650.1*	1,330
District Staff	86,784	350.0	743.7	6,063	1,244*	2,474
% of District English Language Learners	86,784	0.039	0.082	6,063	0.099*	0.132
% of District Free/Reduced Lunch	86,784	0.445	0.225	6,063	0.525*	0.313
% of District with IEP	86,784	0.140	0.051	6,063	0.122*	0.064
State Unemployment Rate	86,784	6.971	2.117	6,063	7.552*	2.203
% of Lower State House Republican	86,784	0.507	0.162	6,063	0.509	0.145
% of Upper State House Republican	86,784	0.533	0.178	6,063	0.498*	0.156
Political Change	86,784	0.390	0.488	6,063	0.407*	0.491
Successful School Finance Litigation	86,784	0.538	0.499	6,063	0.403*	0.491
Number of Districts	12,103			901		

Notes: Financial and district characteristic information for the school years 2009-2010 to 2017-2018 is gathered from district-reported F-33 Financial Survey information. State unemployment rates and political information from the University of Kentucky Center for Poverty Research. Finance litigation information reported from the National Conference of State Legislatures.
*Significantly different at 5% level in t-test analysis

Table 4.2 Overall Average Treatment Effect on the Treated of Charter Schools on Public School District Finances

	(1) Expenditures	(2) Instructional	(3) Capital	(4) Enrollment	(5) Teachers	(6) Staff	(7) Revenue	(8) Assets
A. Charter is established within the district								
ATT	-0.023	-0.056**	0.127*	0.065**	0.056	0.052	0.036	-0.046
	(0.034)	(0.024)	(0.067)	(0.033)	(0.037)	(0.037)	(0.032)	(0.153)
B. Charter is competing within the same county								
ATT	0.140	0.018	-0.096	-0.017	-0.024	-0.078***	-0.007	0.308***
	(0.018)	(0.034)	(0.078)	(0.030)	(0.022)	(0.032)	(0.018)	(0.082)
Observations	94,135	94,135	94,135	94,135	94,135	94,135	94,135	94,135

Note: Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Controls include: log enrollment (except column 4), state’s unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP. District and year fixed effects are included in each regression, with state-linear time trends. Standard errors robust to heteroskedasticity are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4.3 Effect of Charter School Competition by Fraction of Charter Enrollment in County on Public School District Finances

	(1) Expenditures	(2) Instructional	(3) Capital	(4) Enrollment	(5) Teachers	(6) Staff	(7) Revenue	(8) Assets
Charter Fraction	0.03	-0.04	0.54	-0.02	-0.02	-0.08**	-0.01	0.44
	(0.043)	(0.032)	(0.471)	(0.049)	(0.031)	(0.034)	(0.030)	(0.339)
Observations	92,254	92,254	92,254	92,254	92,254	92,254	92,253	92,254
R-squared	0.213	0.459	0.021	0.119	0.228	0.245	0.347	0.097
Number of Districts	12,931	12,931	12,931	12,931	12,931	12,931	12,931	12,931
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Dependent variable in each OLS regression is reported in the second row (each is in the log form). Controls include: log enrollment (except column 4), state's unemployment rate, state political characteristics (political change and the % of each state legislative body being republican), an indicator for the years since the last successful school finance litigation, the % of a school district who are English language learners, % who qualify for national free or reduced lunch programs, and the % who have an IEP. District and year fixed effects are included in each regression, with state-linear time trends. Robust standard errors, clustered at the district level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter 5 - Conclusions: The Importance of Considering Unintended Consequences When Crafting Policy

This dissertation explores the unintended consequences from education policy. The analyses are conducted in secondary and post-secondary settings, using multiple research designs, and examining individual and district-level outcomes. The multiple methods, and levels of analysis, were chosen not only to emphasize the complications associated with education policy research, but also to demonstrate the opportunity within the field. Each of these policies examined in the analysis intends to improve student outcomes but have unintended consequences on students. Following I summarize the main findings, limitations, and policy implications for each chapter.

In chapter two, “Is Thursday the New Friday? The Four-Day School Week and Teen Traffic Safety”, I investigate whether district-level schedule changes impact the traffic safety of driving age teens within the district. This is the largest analysis to date within the four-day school week literature, incorporating district-level outcomes in nine different states across the United States. With concerns of additional weekend night driving opportunities motivating the analysis, we surprisingly find no evidence of increased fatal accident involvement in school districts that adopt the four-day school week schedule. In analyses further examining how the schedule change has impacted fatal accident involvement no evidence is found that raises concerns over the additional weekend night, and evidence of a decline for female driving age teens is found. While this single analysis looking at driving age teen traffic safety cannot answer the question of what is happening to students in four-day school districts after adopting the schedule, it is one piece of evidence that can serve to re-assure policymakers and local education leaders seeking information on the potential downsides of the policy. Future opportunities for research lie in obtaining more general traffic accident data, such as done in the secondary analysis of Colorado and Idaho but applied to additional states and years. Incorporating more years into each analysis, i.e. before school year 2010-2011, could allow for more power since many rural school districts first adopted the four-day school week schedule in the previous decade. The research frontier is broad for four-day school week policy though, with potentially even more interesting research still to be done utilizing YRBSS data in a longitudinal setting.

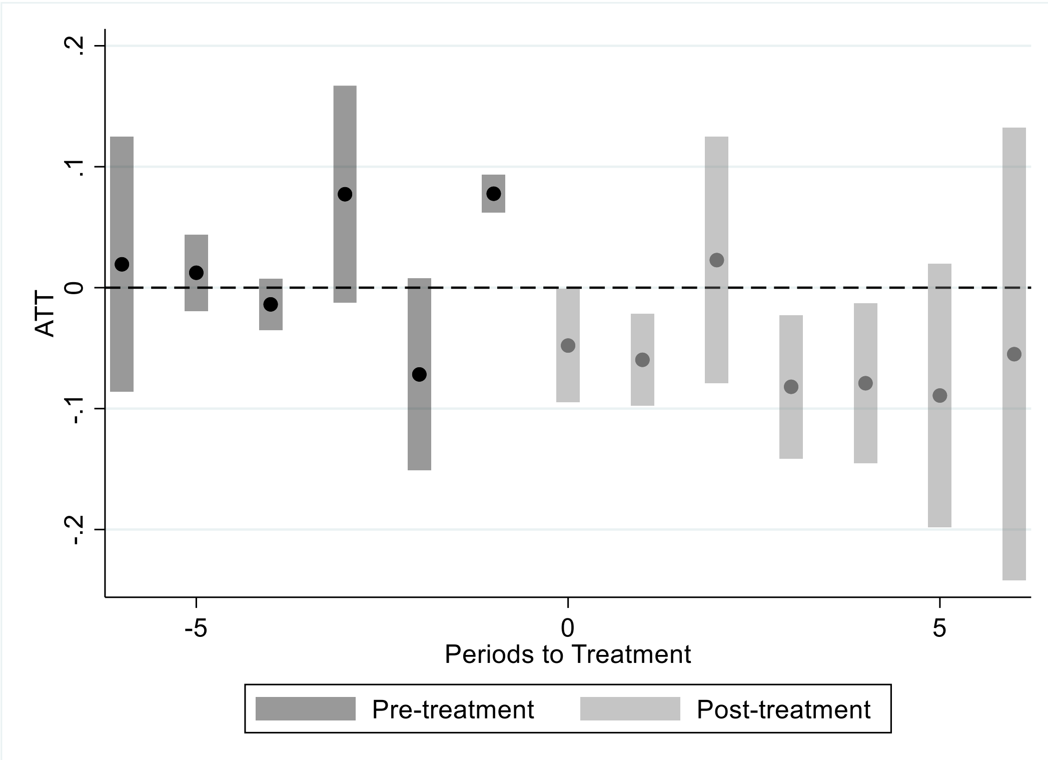
In chapter three, “Academic Probation & Financial Aid: Financial Aid Implications of Probation”, we examine for the first time in the literature how the underlying financial aid status of undergraduate students impacts their responses to being placed on academic probation. This analysis is unique it its setting within a large, essentially open access public university in a minority-majority U.S. state. Using individual student data in a regression discontinuity design we study otherwise similar students who fall just above and below the academic probation GPA cutoff. We find clear decreases in the financial aid packages awarded to Pell grant students who face satisfactory academic progress requirements at the same GPA cutoff, while non-Pell students on academic probation have no measurable financial aid implications from academic probation. Even though a clear difference exists in the financial implications of academic probation for both groups of students, we find similar responses academically the following semester. Non-Pell students significantly increase their next semester GPA, while Pell students have a similar in magnitude response. Neither group appears to respond to being placed on academic probation by disenrolling from the university in the following semester, and no evidence is found that suggests graduation rates are affected. These results would suggest that student responses to satisfactory academic progress requirements, academic probation, and potentially merit aid progress requirements are not directly tied to the financial implications of the policies. While more evidence is clearly needed, the results of this particular analysis apply to undergraduate students near the 2.0 academic probation GPA cut-off, administrators should consider the merits of disentangling progress requirements from the loss of financial aid. At the very least in the semester following when a student has below desired performance. Additional research on merit aid programs, and their GPA requirements, at the same university provides a great opportunity to further understand the relationship between student responses to GPA requirements and financial aid.

In chapter four, “Charter Schools & School District Finances: How Does Resource Usage Change at the District Level When Charter Schools are Established?, we examine whether charter school competition affects public school district finances. This is the first analysis expanding to multiple states and utilizing a rich financial dataset in the choice literature. Charter school competition has often only considered the presence of a charter school within the community, and ignored if the charter is founded by the public school district or an outside organization. This small difference in who founds the charter leads to different implications for the public school district affected by charter school competition. Using a national dataset of school district level finances, including: expenditures, district enrollment, and district level staffing information we use difference-in-differences methods to find the causal effect of charter schools on public school district finances. After a school district establishes a charter school district student enrollment and teachers employed by the district increase in the subsequent years. No measurable change is found in per student expenditures, revenue, or assets for school districts. When examining if a public school district faces an outside charter competing within the same county for students, an opposite effects is found for enrollment.. These findings are in line with expected responses of families and students when additional education options are presented, they use them. State and local policymakers must consider the implications of this change in usage from each type of charter school. Additional research examining student-level enrollment choices is the next step in understanding the full consequences of the policy. If enrollment increases attributed to charters are students returning to public sector from private schools, this could result in a strain of existing public resources if funding does not respond. Alternatively if enrollment changes are the result of public school students choosing charter schools, resources can just be re-allocated appropriately.

These chapters hopefully demonstrate the importance and value of fully considering the potential unintended consequences when enacting policy, particularly within education. In some instances, the unintended effects could be of more social importance than the goal of the policy. This often happens when the unintended effects are negative for individuals, but this also can occur when unintended effects give insights to other areas. As in the case of compulsory school attendance laws mentioned in chapter 1, the unintended consequence of that policy was different education levels for otherwise similar individuals born just before and after the school start cutoff date. Using that natural experiment, unintentionally created by policy, Angrist & Krueger (1991) were able to estimate a measure of returns on education that did not suffer from the bias present in other analyses. It is well understood and often stated that many experiments economists would like to conduct, or questions we would like to analyze, cannot be studied directly due to ethical reasons. You cannot randomize education levels for individuals, or in the case of my research being placed on academic probation; but you can use econometric techniques on natural experiments to come close to understanding their effects.

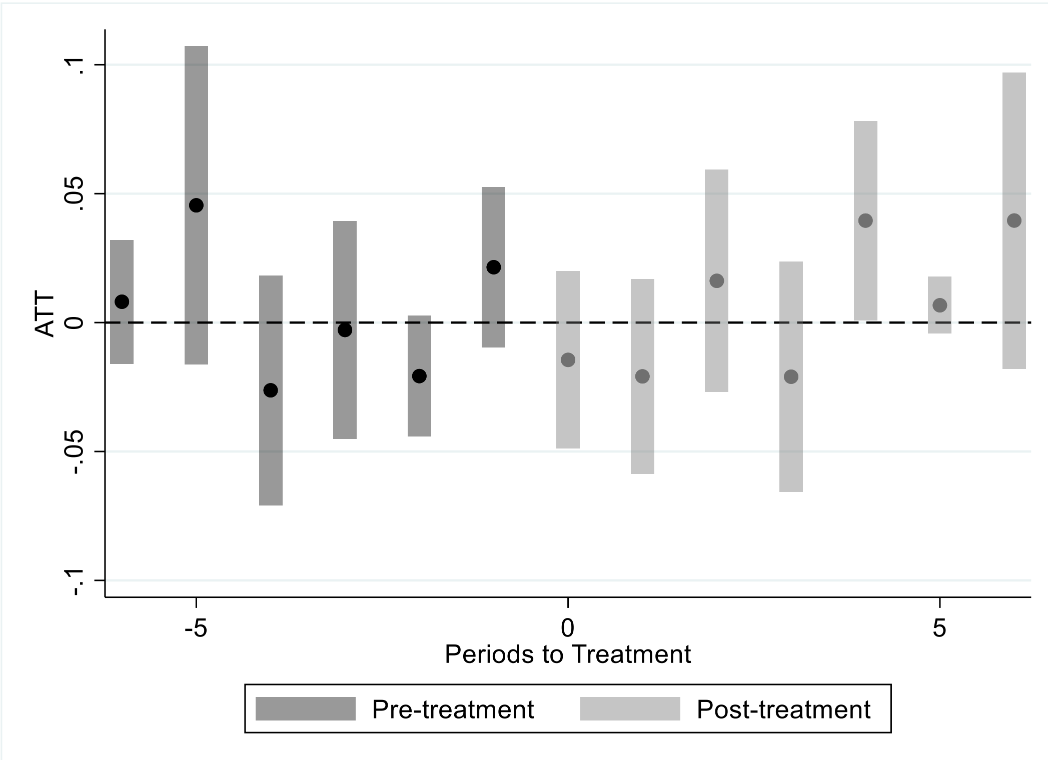
Appendix To Chapter 2

Figure 2A.1 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Sunday



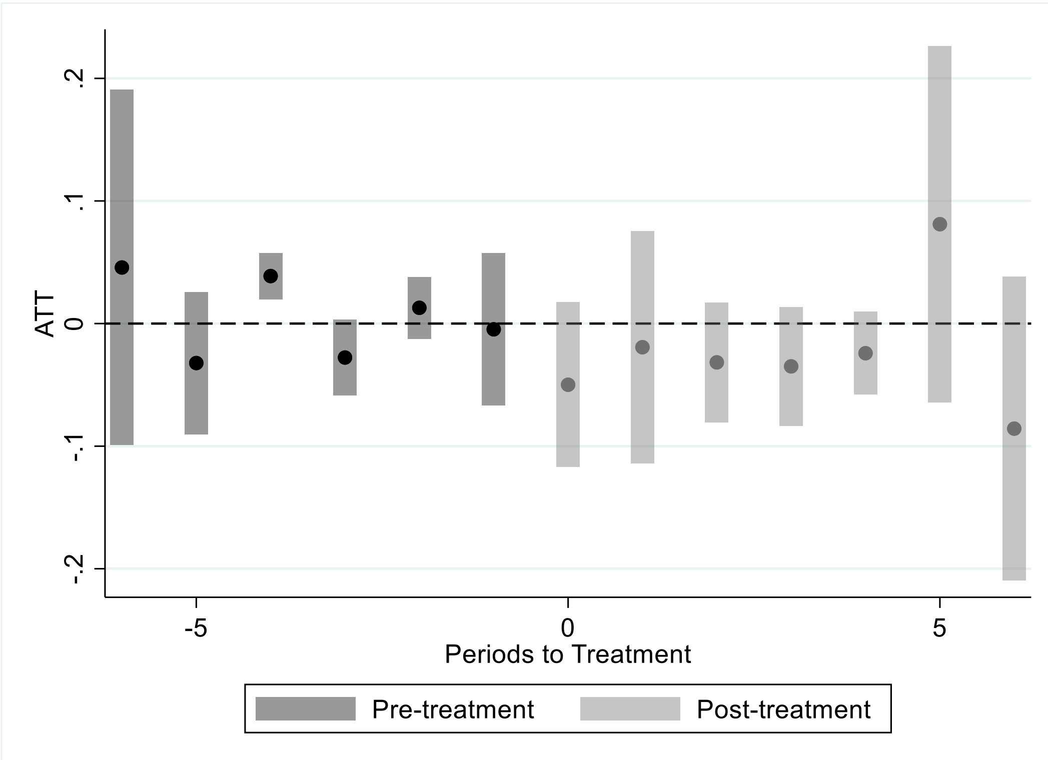
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.2 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Monday



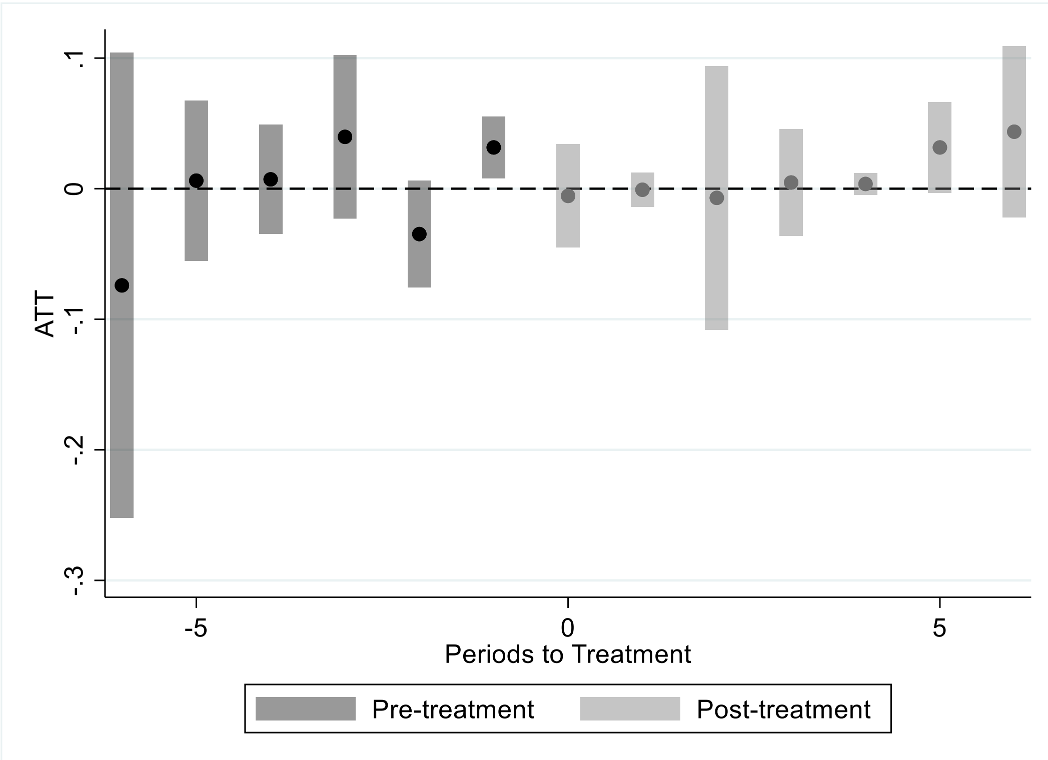
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.3 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Tuesday



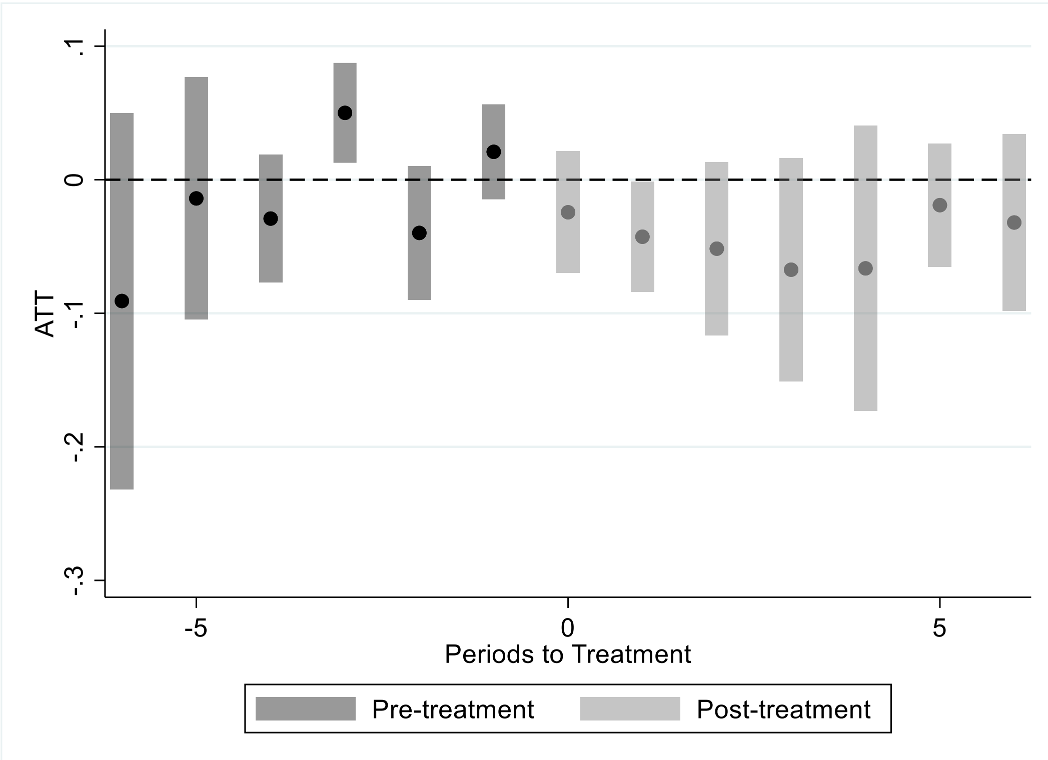
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.4 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Wednesday



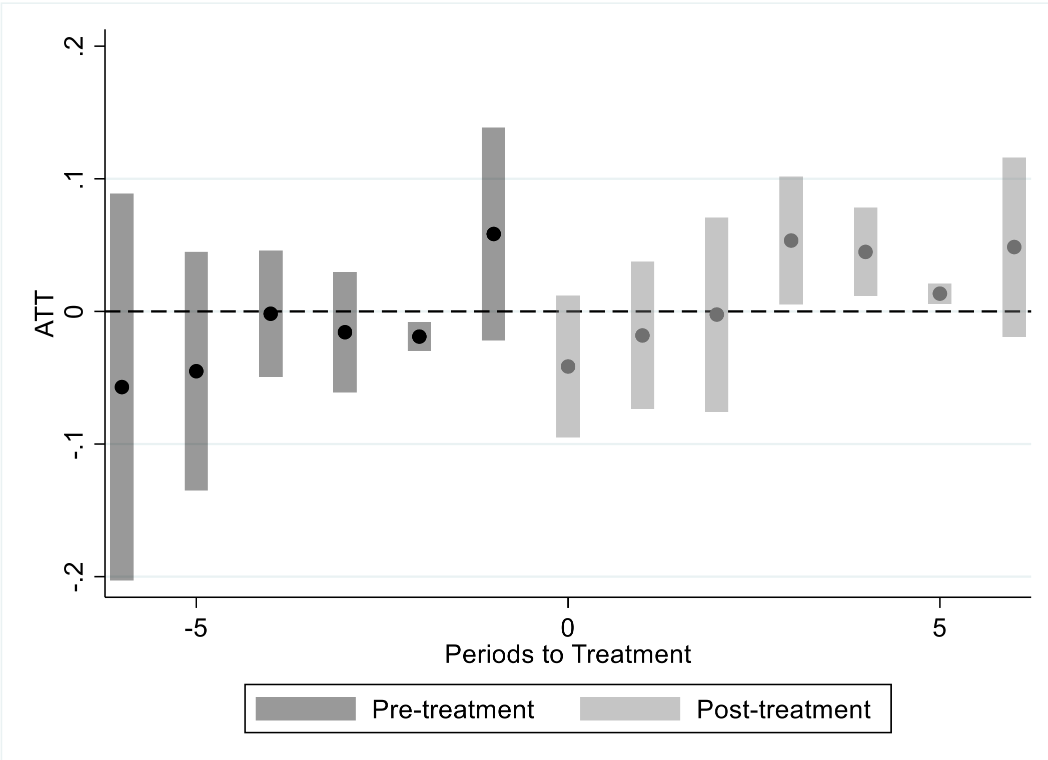
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.5 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Thursday



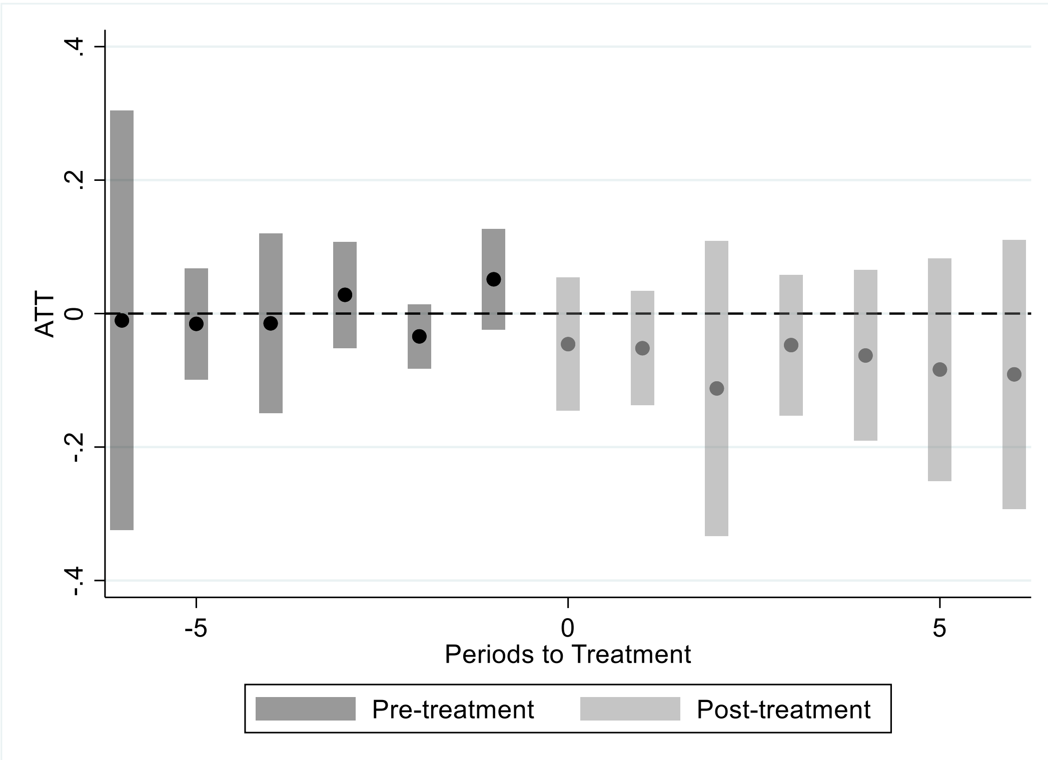
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.6 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Friday



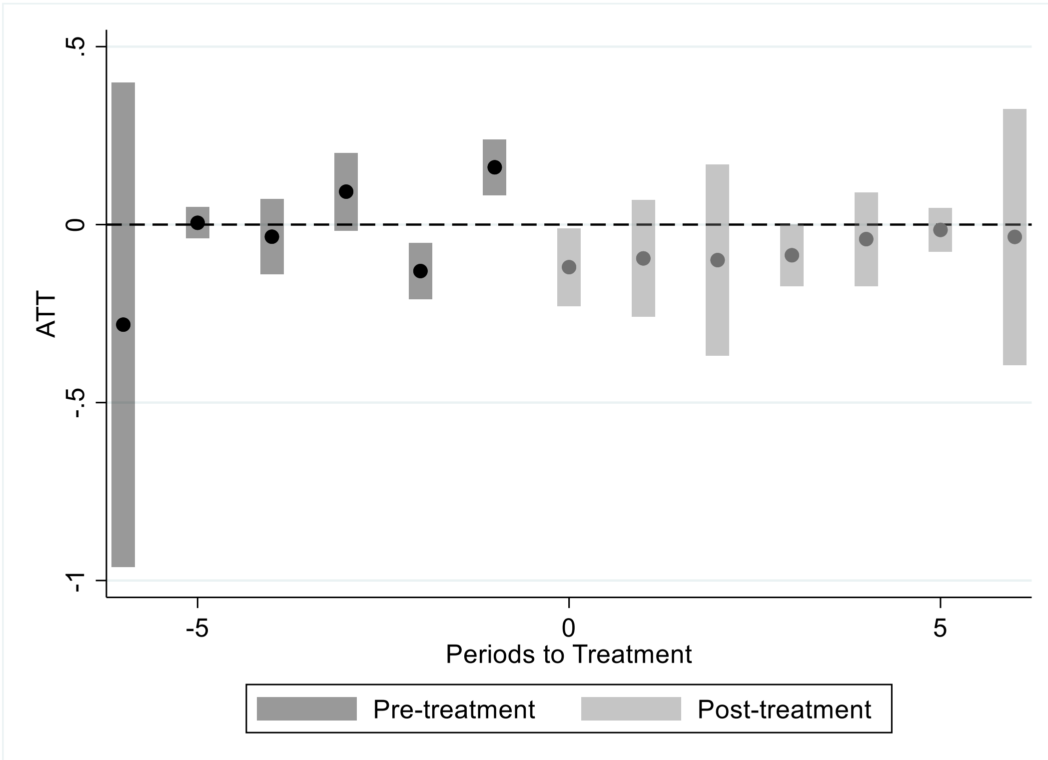
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.7 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District on Saturday



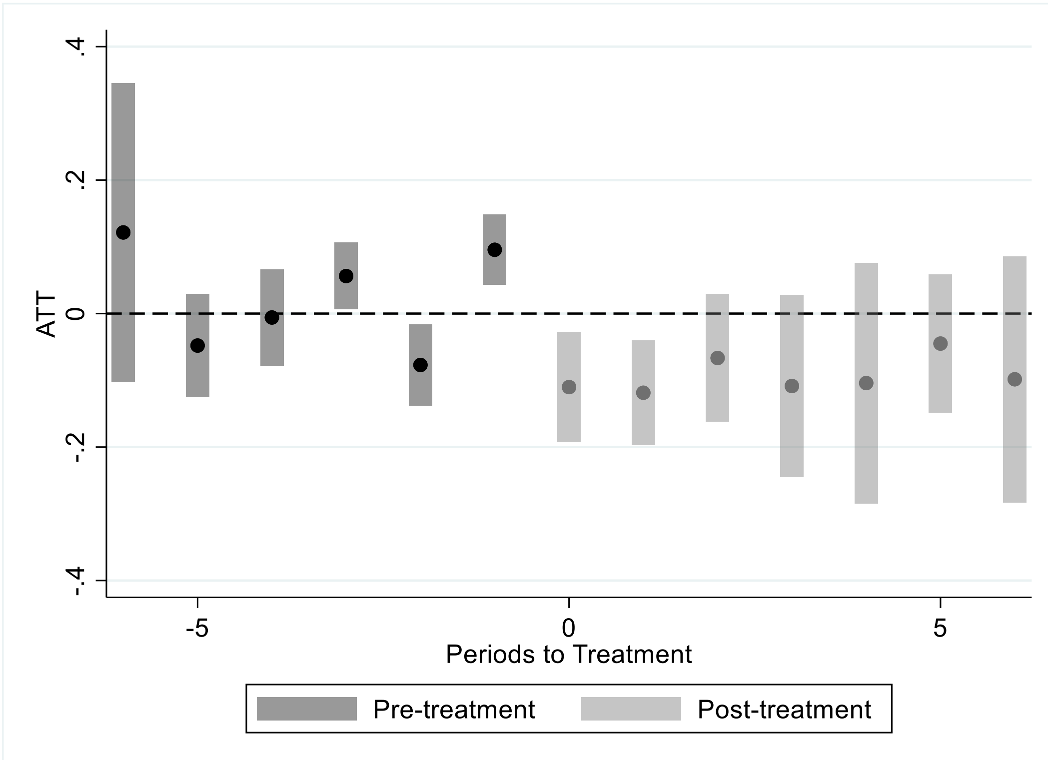
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.8 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated Male 15-18 Year Olds Involved in Fatal Accident in District



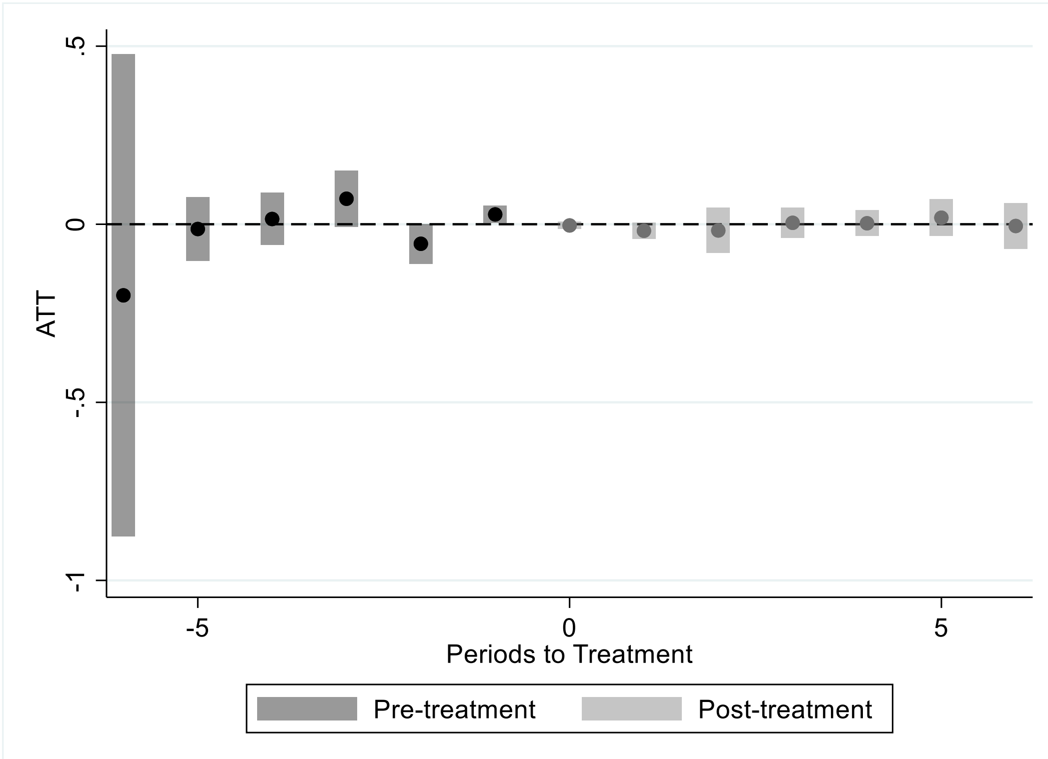
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.9 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated Female 15-18 Year Olds Involved in Fatal Accident in District



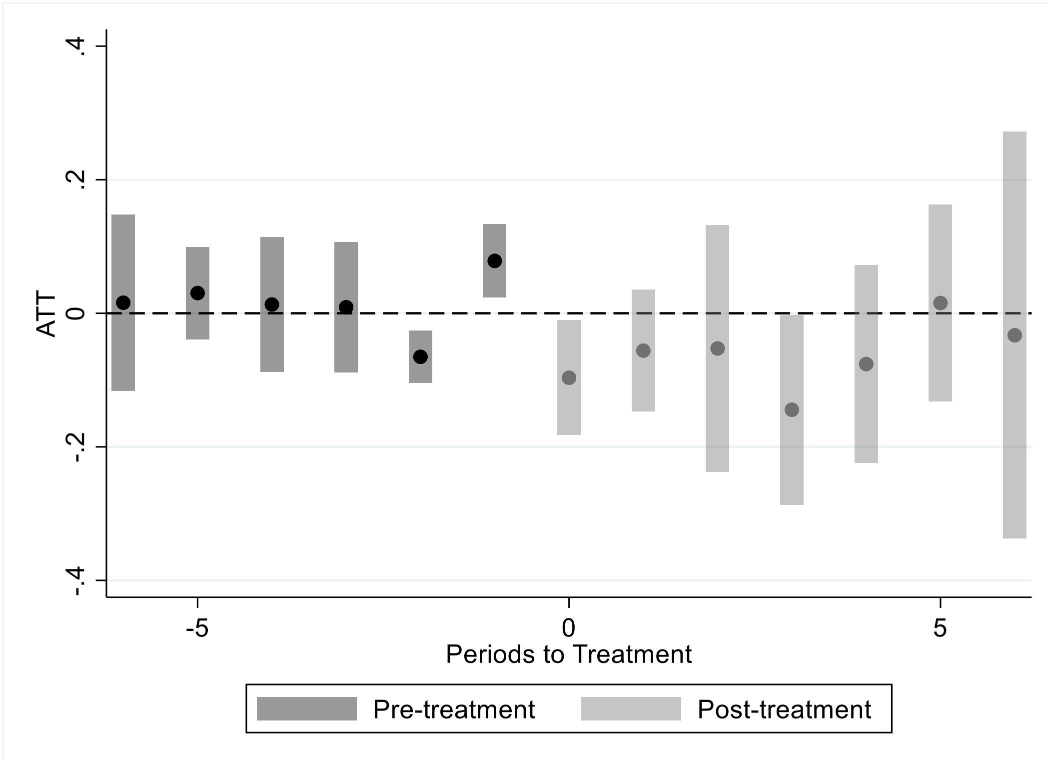
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.10 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District during Night Hours



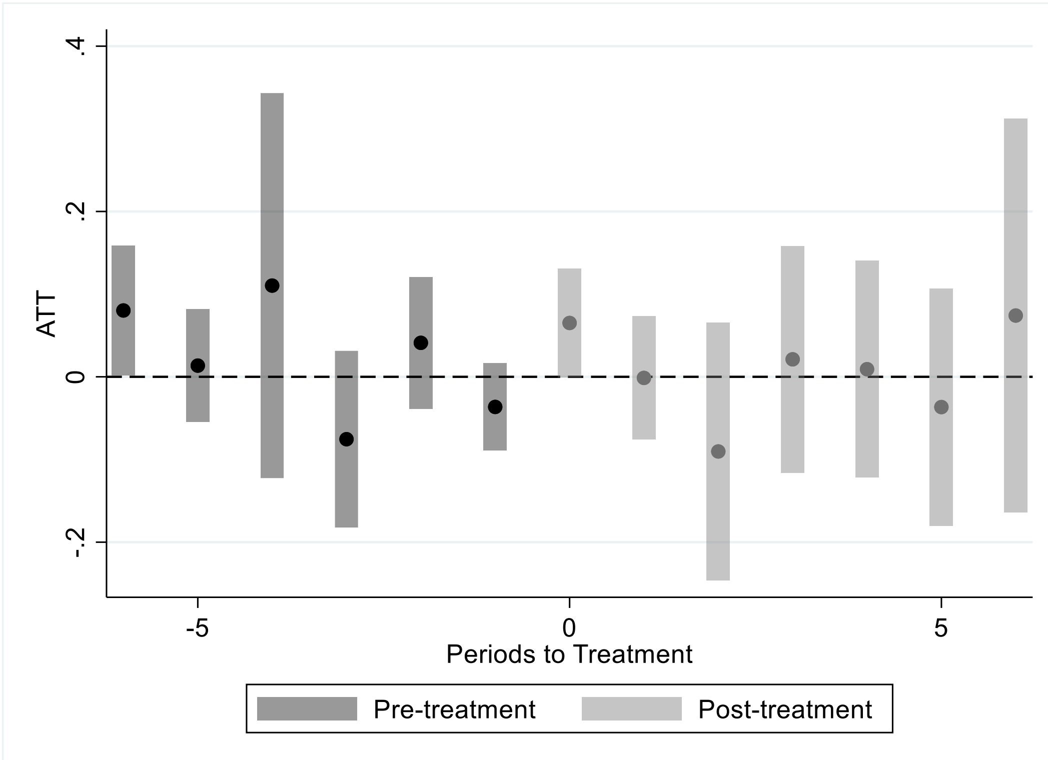
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.11 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in District during Day Hours



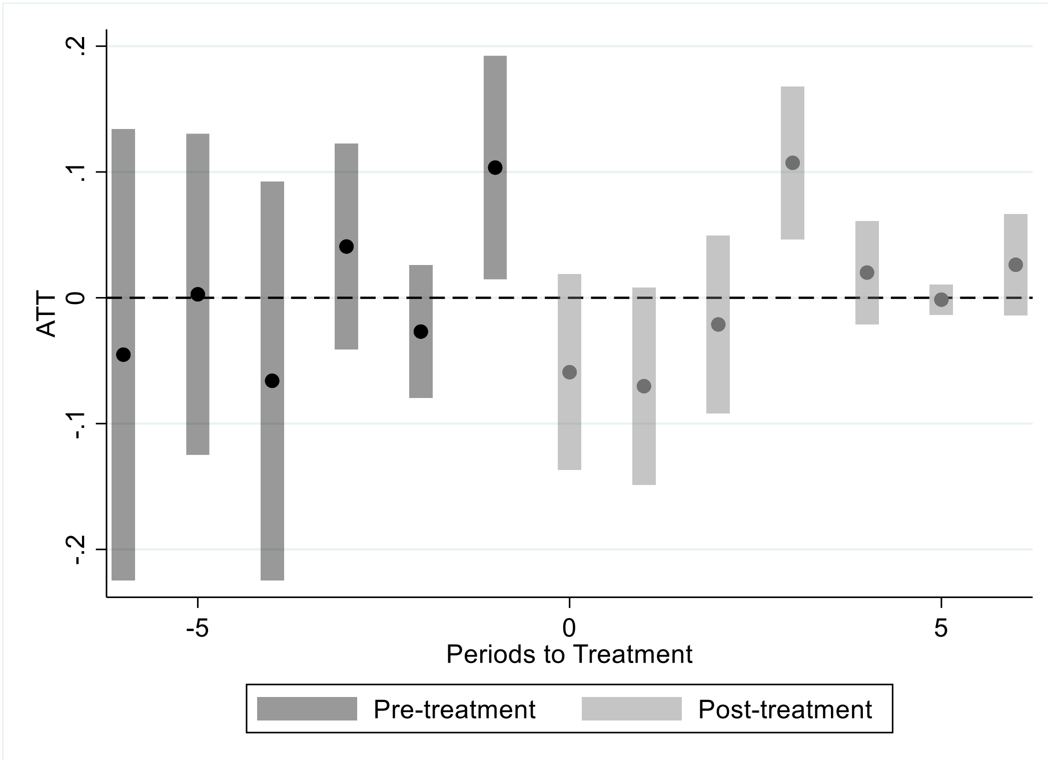
Note: Time of day is split into day and night based on previous literature (Dee, 1999). Night is from the hours of midnight to 4:59 am and day is from the hours 7:00 am to 3:59 pm.
Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.122 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 15-18 Year Olds Involved in Fatal Accident in Summer Months



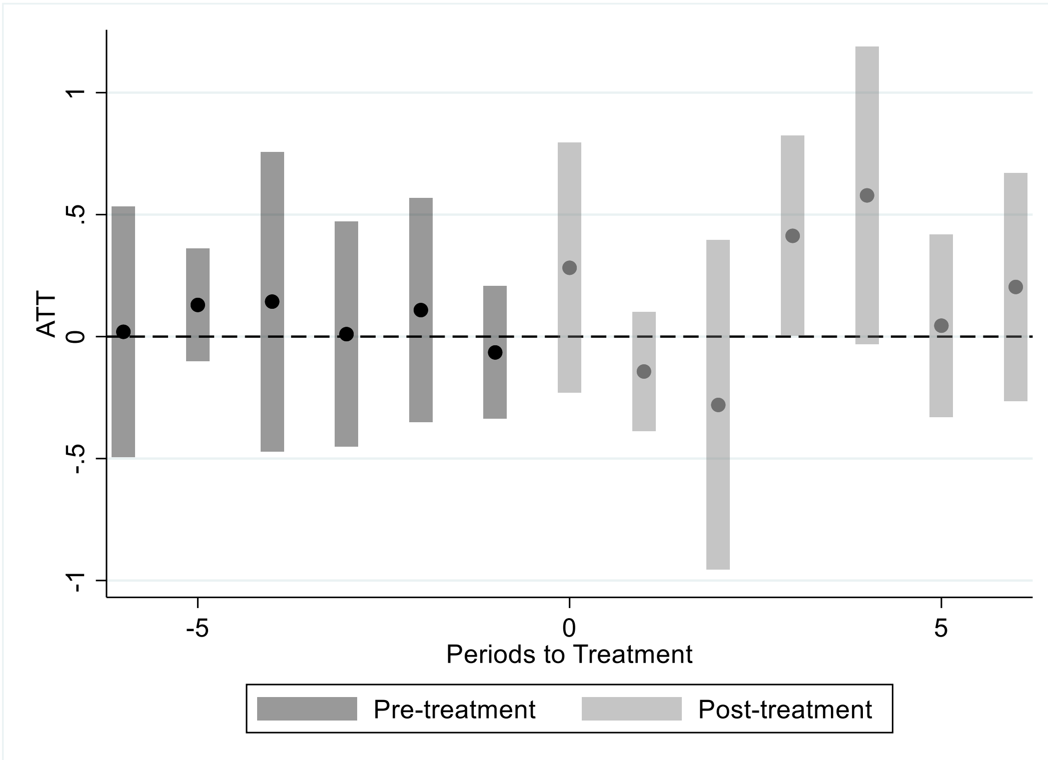
Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.133 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated 26-31 Year Olds Involved in Fatal Accident in Summer Months



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Figure 2A.144 Event Study – Dynamic Effects of Average Treatment Effect (Four-Day School Week) on the Treated All Individuals Involved in Fatal Accident in Summer Months



Note: Figure plots coefficient estimates from Callaway & Santana (2021) Doubly Robust estimator with 95% confidence intervals from a regression of outcome variable, involvement in a fatal accident, separately for each period and group. Treated group consists of districts that adopt a four-day school week, the control group five-day school week districts and not yet treated four-day school week districts. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios.

Table 2A.1 Year Districts Adopt a Four-Day School Week

State	Never Four-Day	Year (Fall) School District Switches								Total Districts (by State)
		2010	2011	2012	2013	2014	2015	2016	2017	
Colorado	70	0	3	5	2	2	2	3	10	97
Idaho	71	0	12	2	3	3	2	0	0	93
Kansas	265	0	1	0	0	0	1	3	0	270
Missouri	490	0	3	2	3	0	6	3	7	514
New Mexico	50	0	4	0	0	1	0	1	6	62
Oklahoma	407	0	2	5	4	7	1	44	1	471
South Dakota	106	0	5	0	0	0	0	1	1	113
Wyoming	31	0	0	0	0	0	0	1	2	34
Total Switches (by Year)	1,490	0	30	14	12	13	12	56	27	1,654

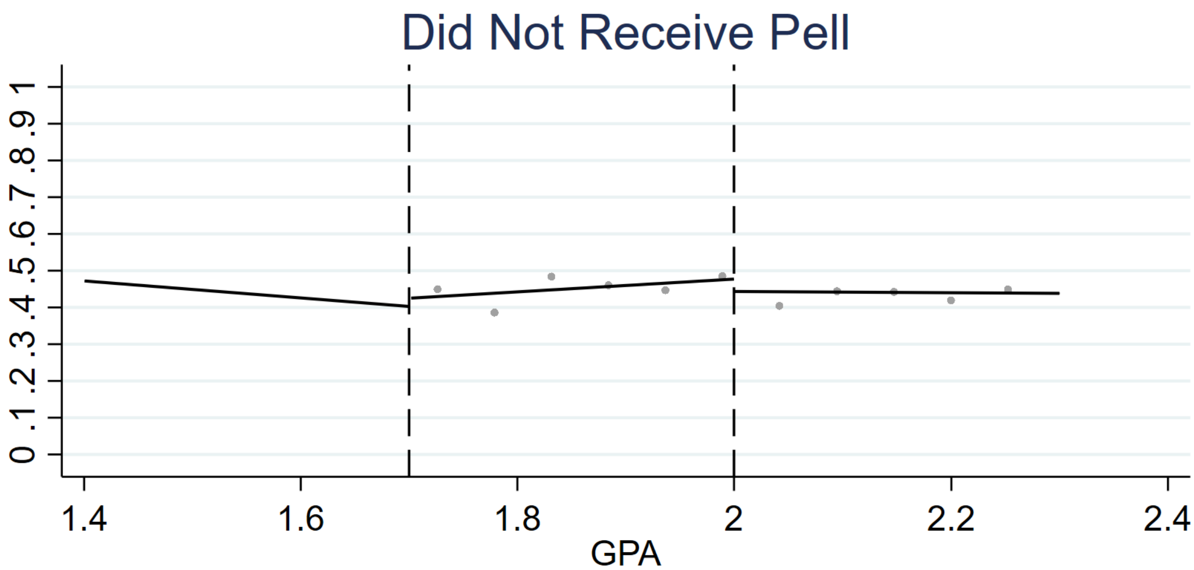
Table 2A.2 Average Treatment Effect on Treated of Four-Day School Week Schedule on Number of Individuals Involved in Fatal Accident for Summer Months

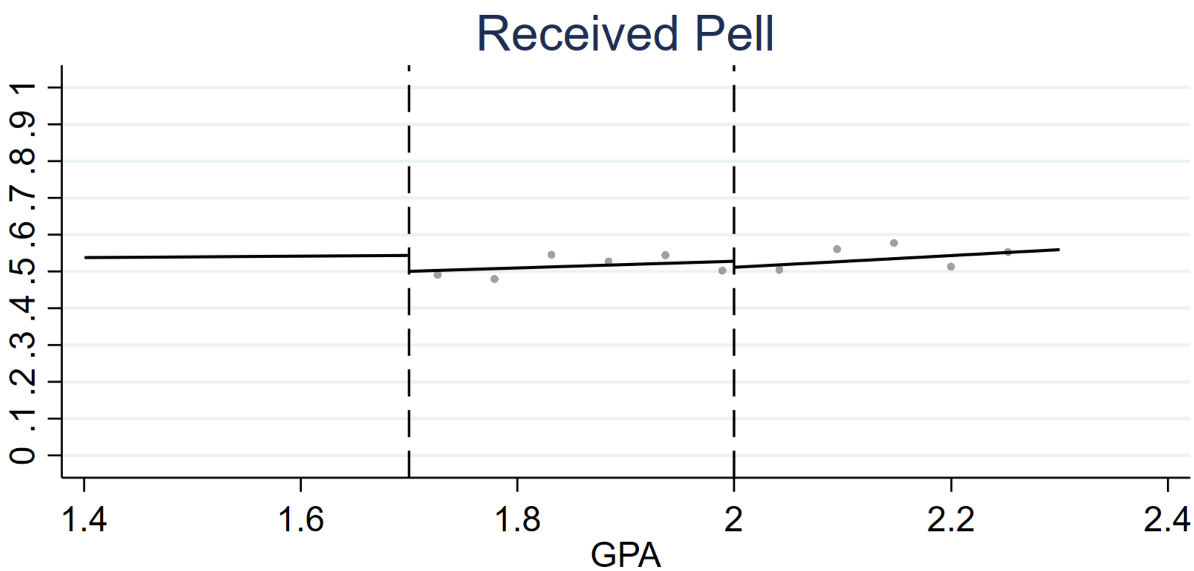
	15-18 Year Olds	26-31 Year Olds	All in District
ATT – Four Day School Week	0.001	-0.020	0.126
	(0.043)	(0.029)	(0.156)
Mean Outcome	0.060	0.060	0.633
for Four-Day District	(0.369)	(0.354)	(2.149)
Observations	13,232	13,232	13,232

Doubly Robust difference-in-differences estimator with not yet treated observations included with control group of never treated observations. Covariates include: district enrolment (thousands), free-reduced lunch percentage, and student teacher ratios. Standard errors robust to heteroskedasticity in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Appendix to Chapter 3

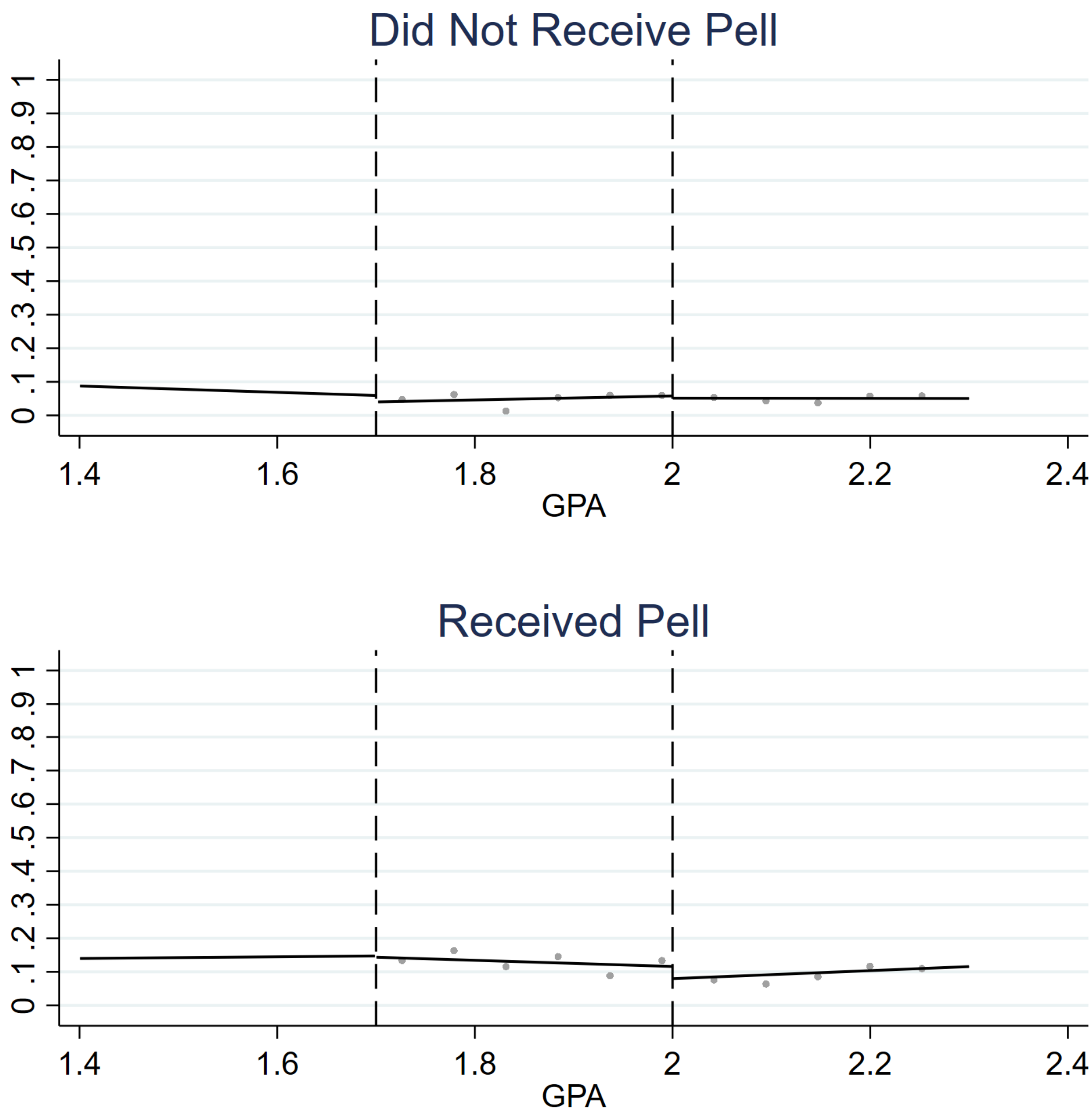
Figure 3A.1 Discontinuities in Observables – Female – Freshmen





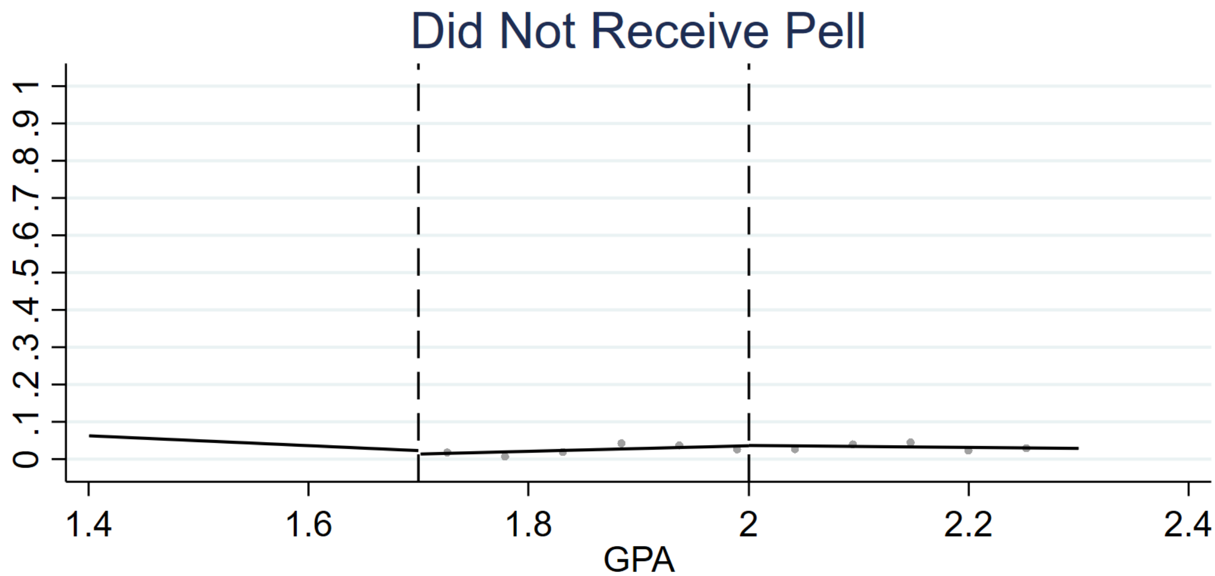
Regression discontinuity results for female (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29.

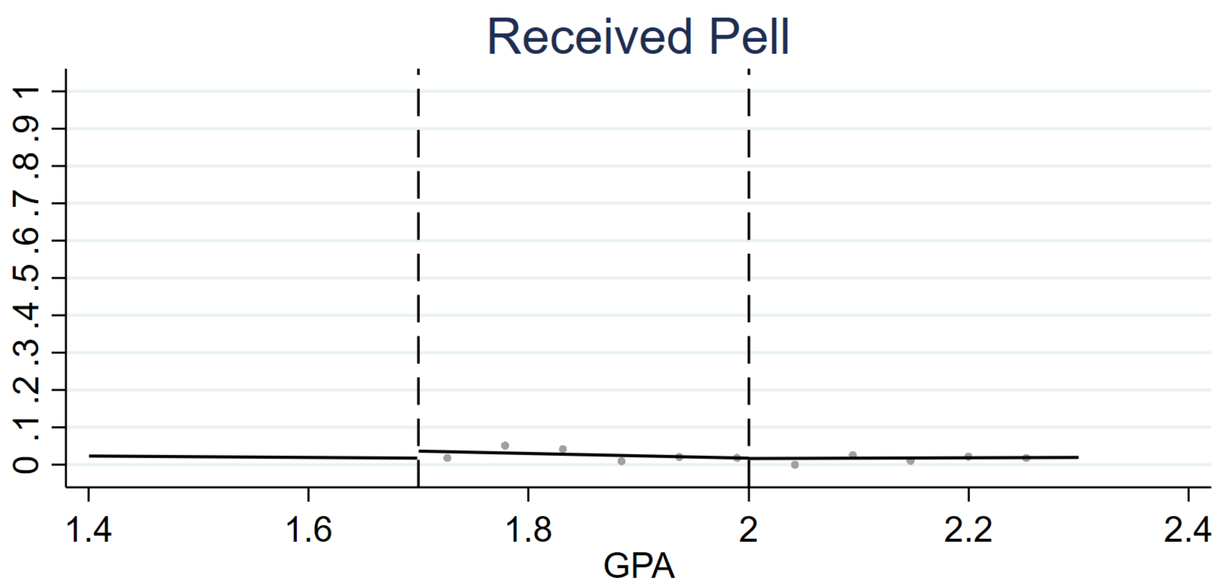
Figure 3A.2 Discontinuities in Observables – Native American – Freshmen



Regression discontinuity results for Native American (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

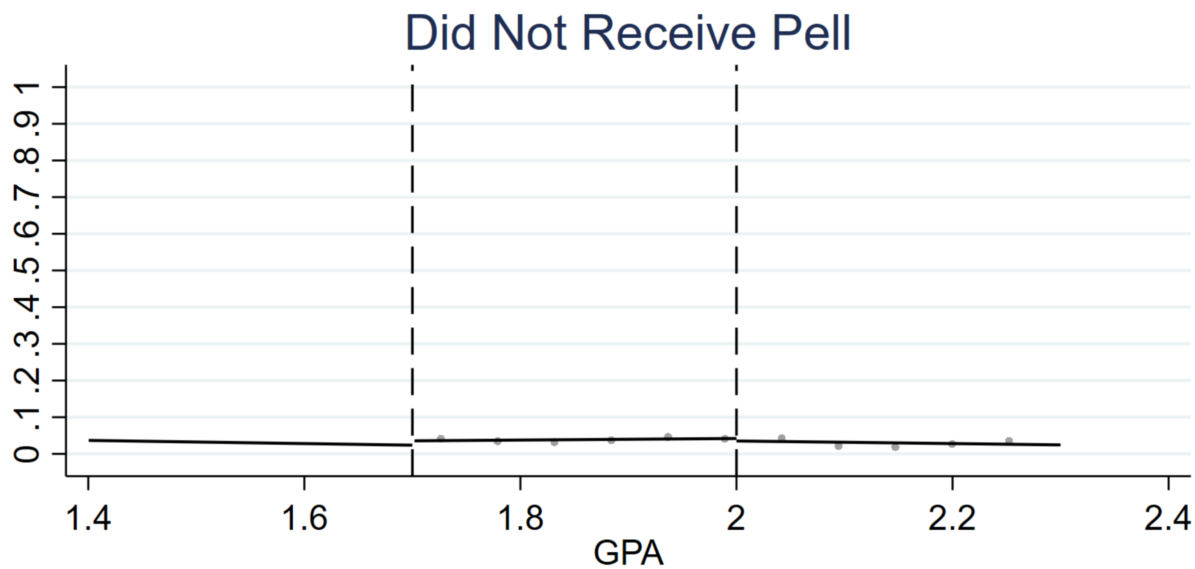
Figure 3A.3 Discontinuities in Observables – Asian – Freshmen

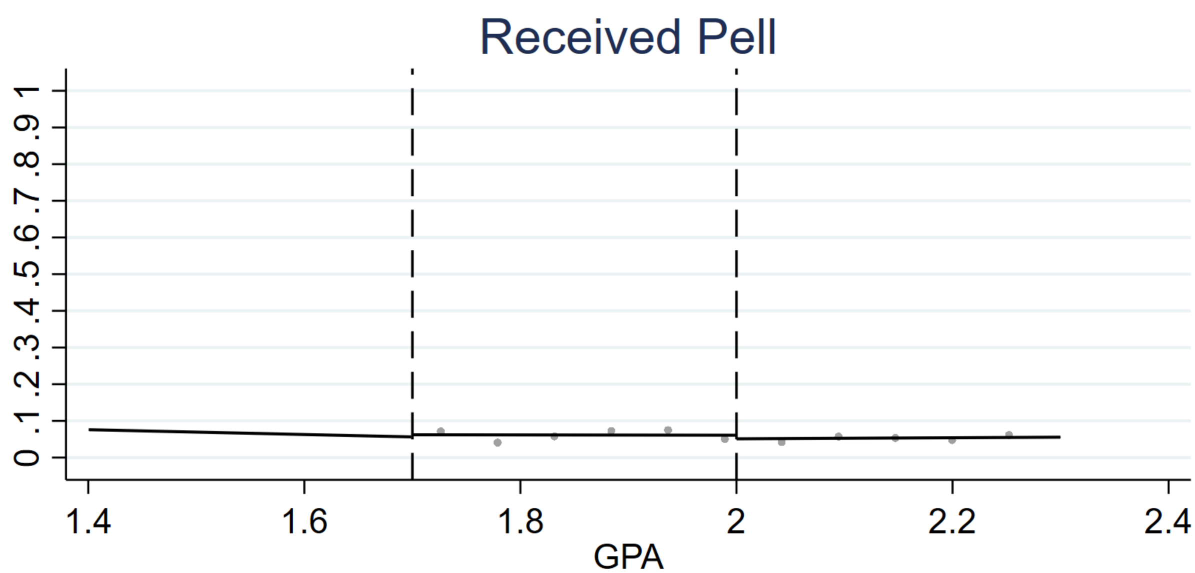




Regression discontinuity results for Asian (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

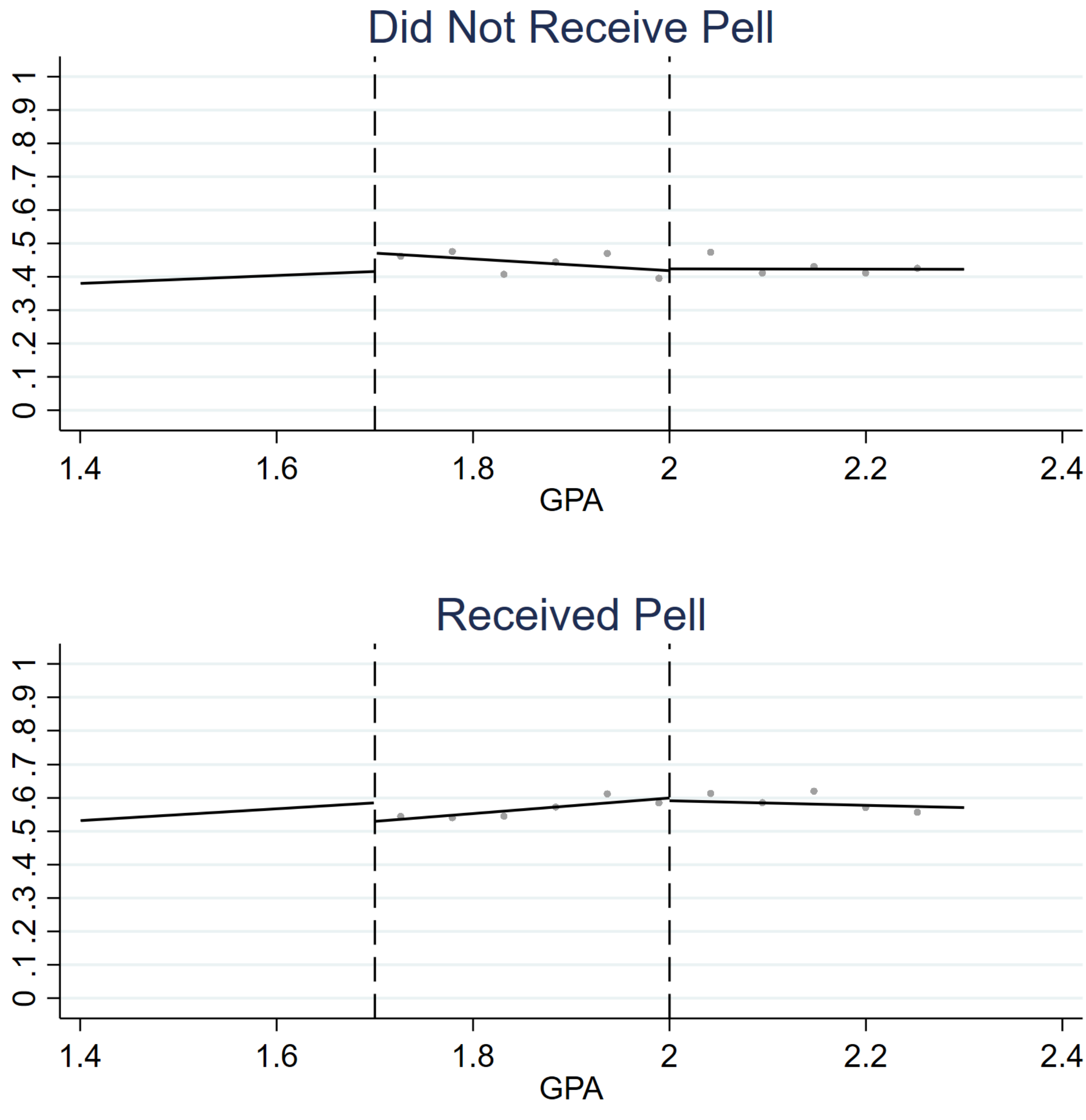
Figure 3A.4 Discontinuities in Observables – Black – Freshmen





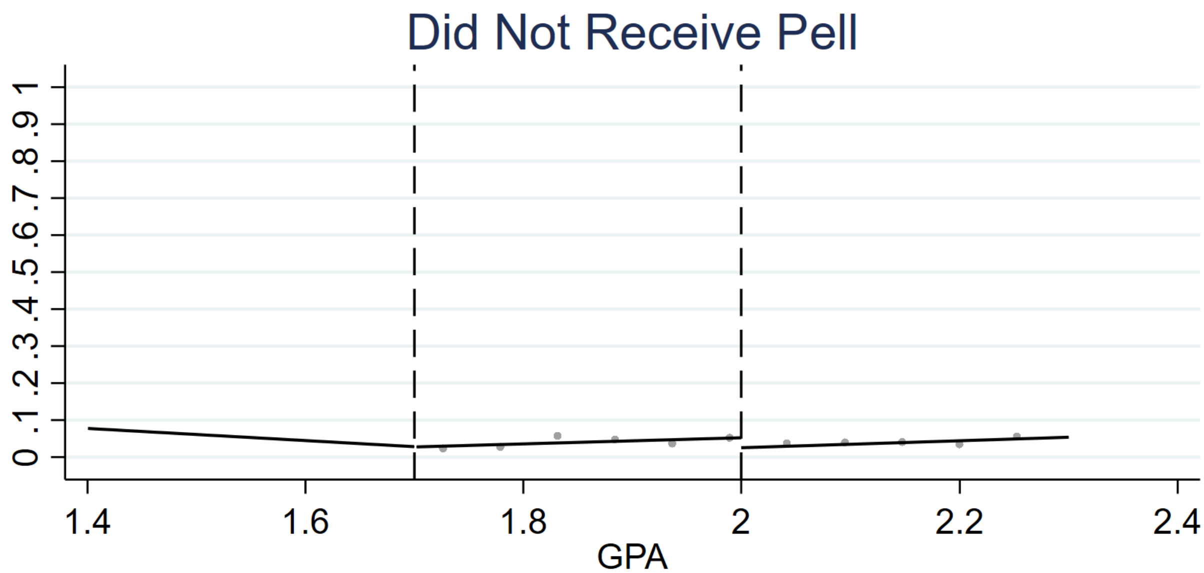
Regression discontinuity results for Black (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

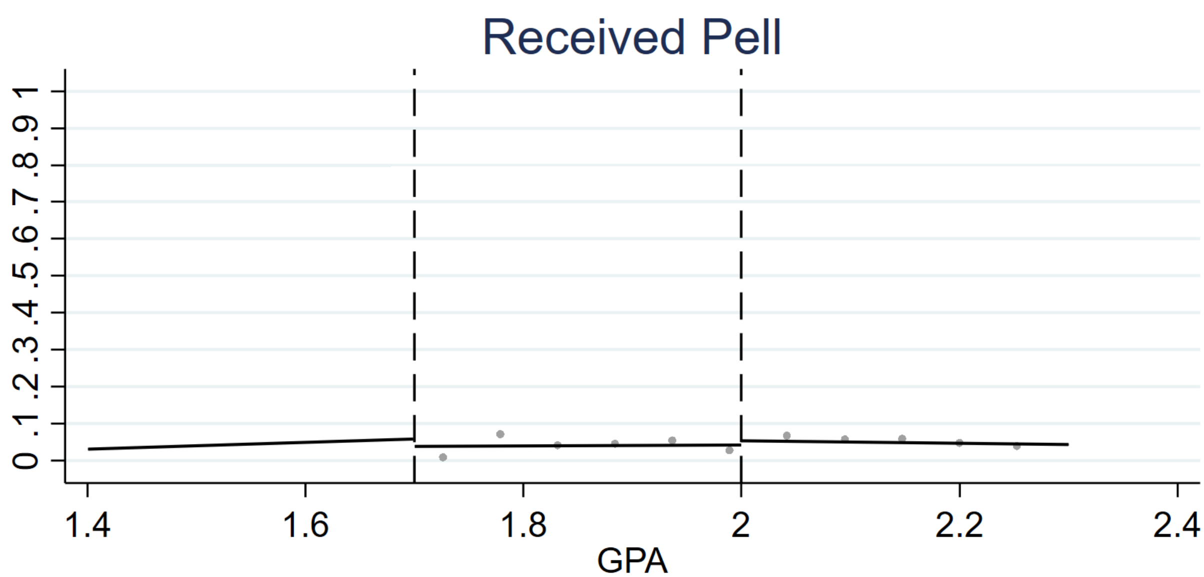
Figure 3A.5 Discontinuities in Observables – Hispanic – Freshmen



Regression discontinuity results for Hispanic (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

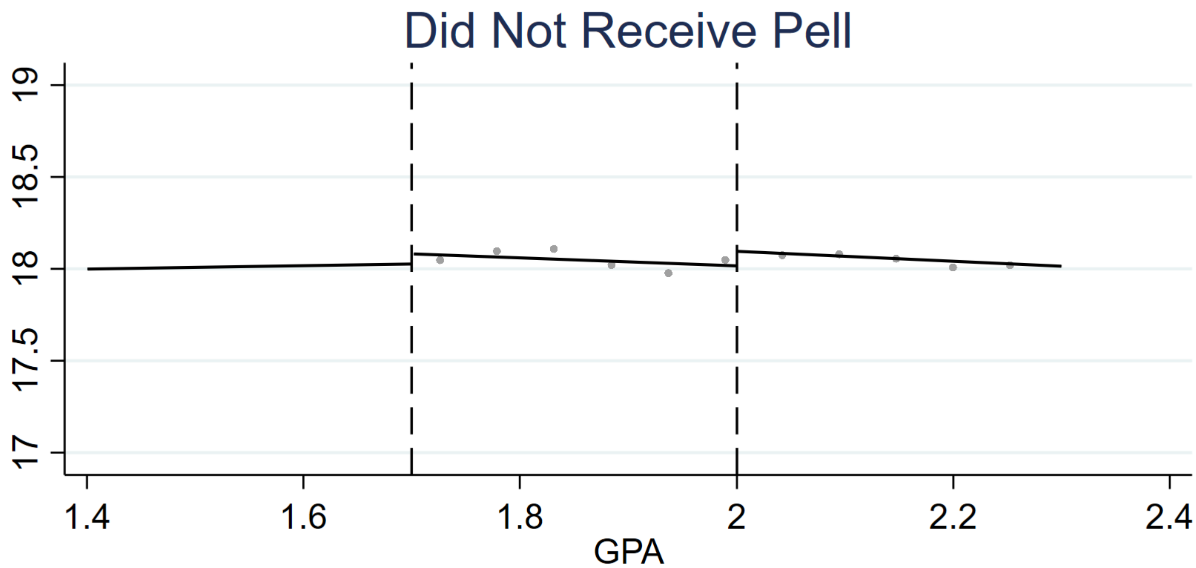
Figure 3A.6 Discontinuities in Observables – Other Race – Freshmen

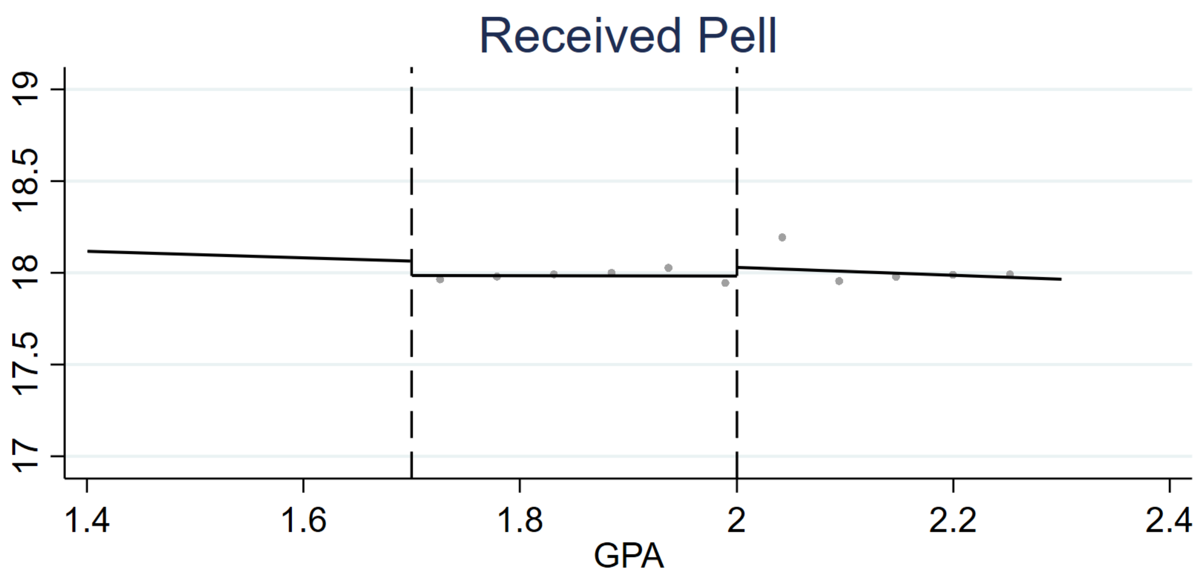




Regression discontinuity results for Other Race (indicator variable) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

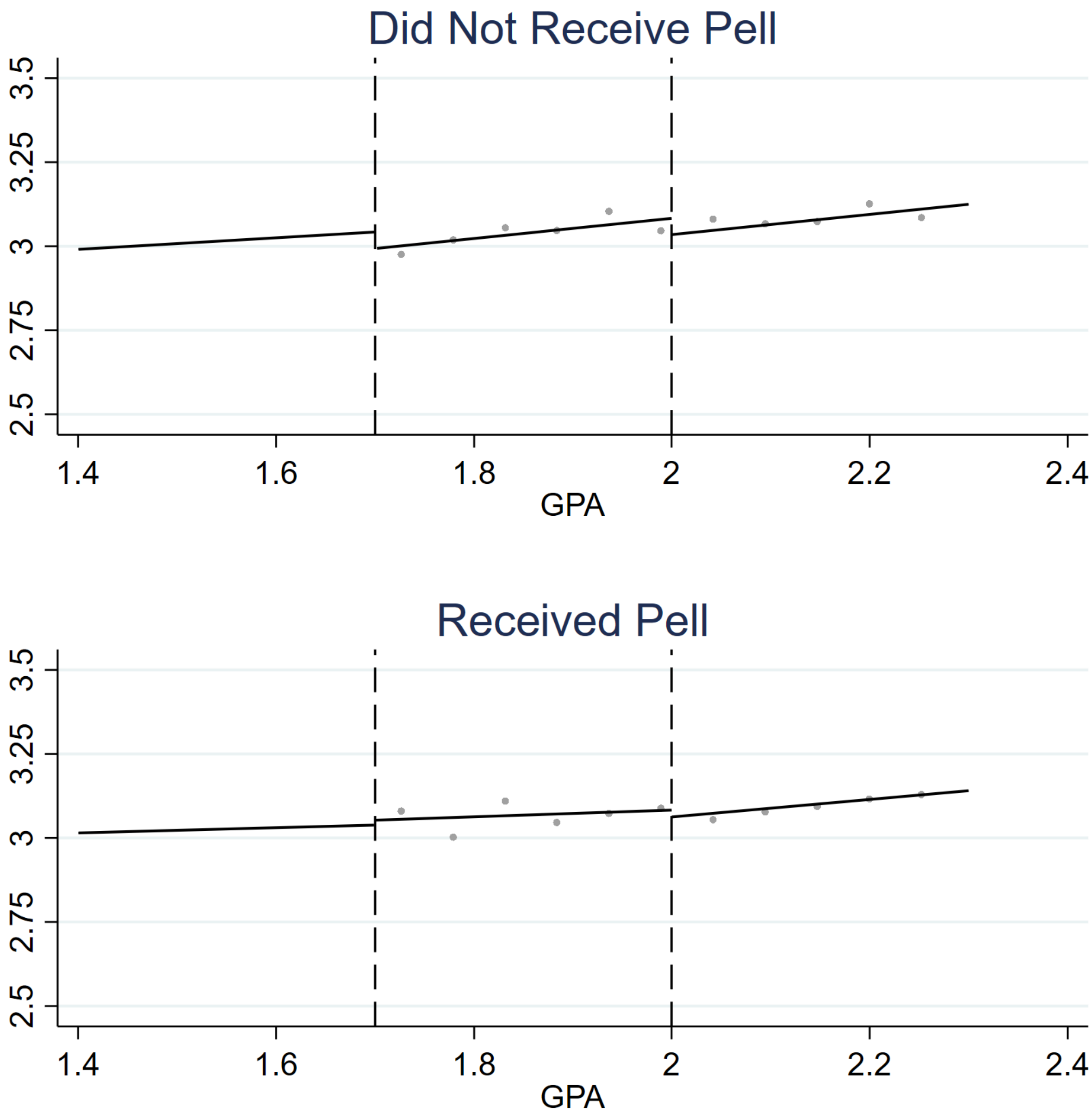
Figure 3A.7 Discontinuities in Observables – Age Admitted – Freshmen





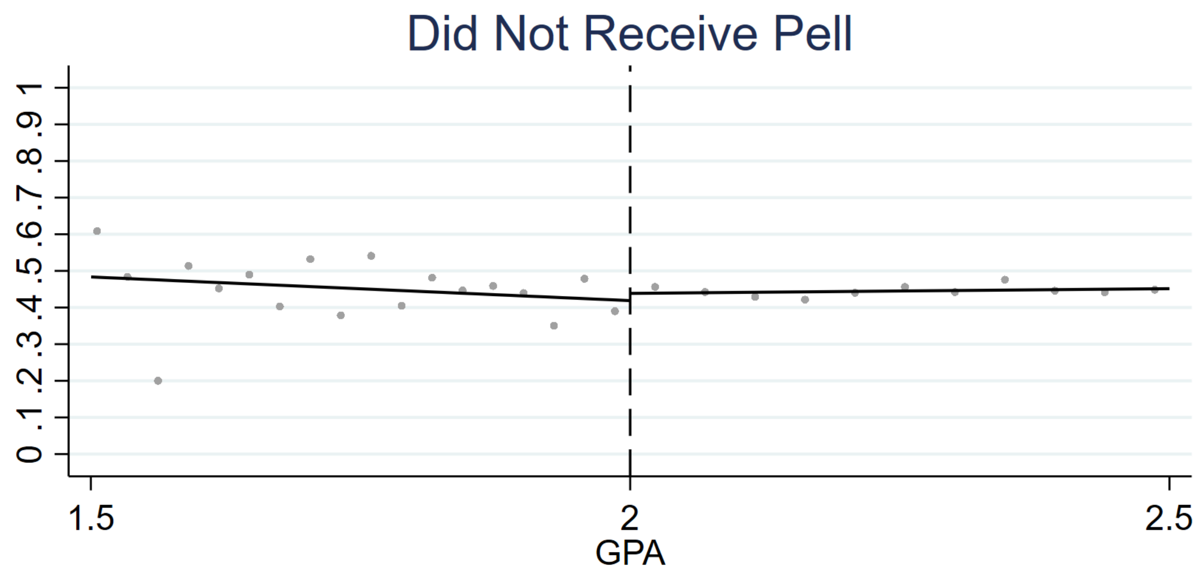
Regression discontinuity results for Age Admitted (in years) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

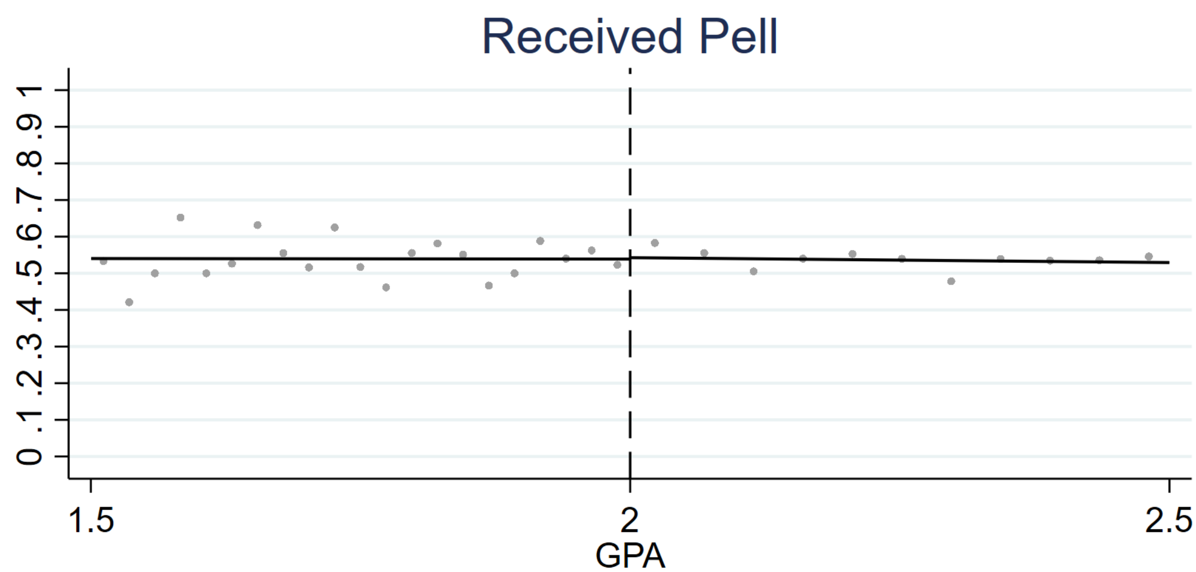
Figure 3A. 8 Discontinuities in Observables – High School GPA – Freshmen



Regression discontinuity results for High School GPA (in GPA units) following the semester either did not meet satisfactory academic progress or were placed on academic probation. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of not meeting satisfactory academic progress GPA requirements, GPA of 1.7, and being placed on academic probation, which occurs at a 2.0 GPA. Each cutoff is shown by a dashed line. Estimates are based on linear regression as described in empirical strategy section using bandwidths of 0.29

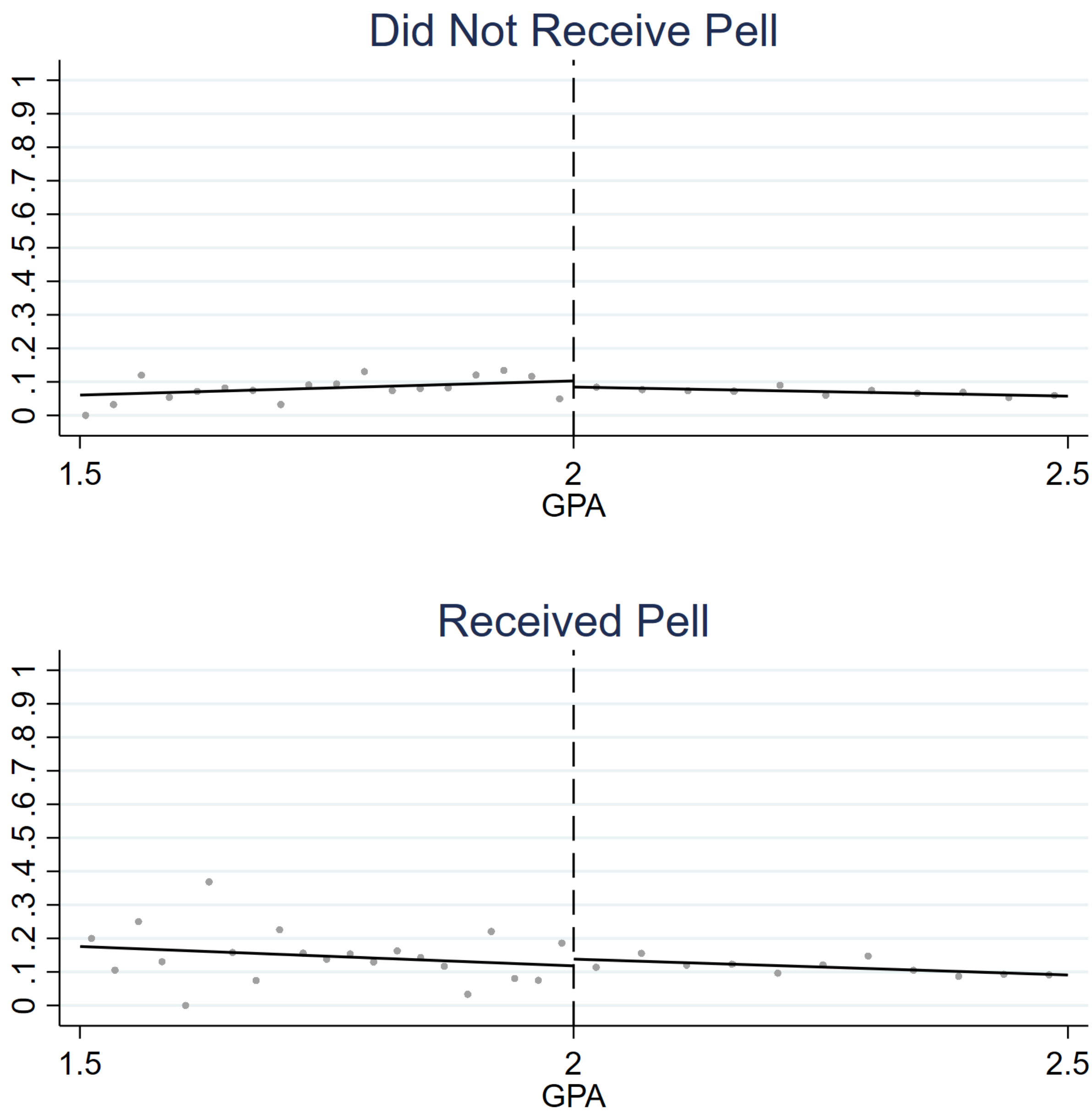
Figure 3A.9 Discontinuities in Observables – Female – Upperclassmen





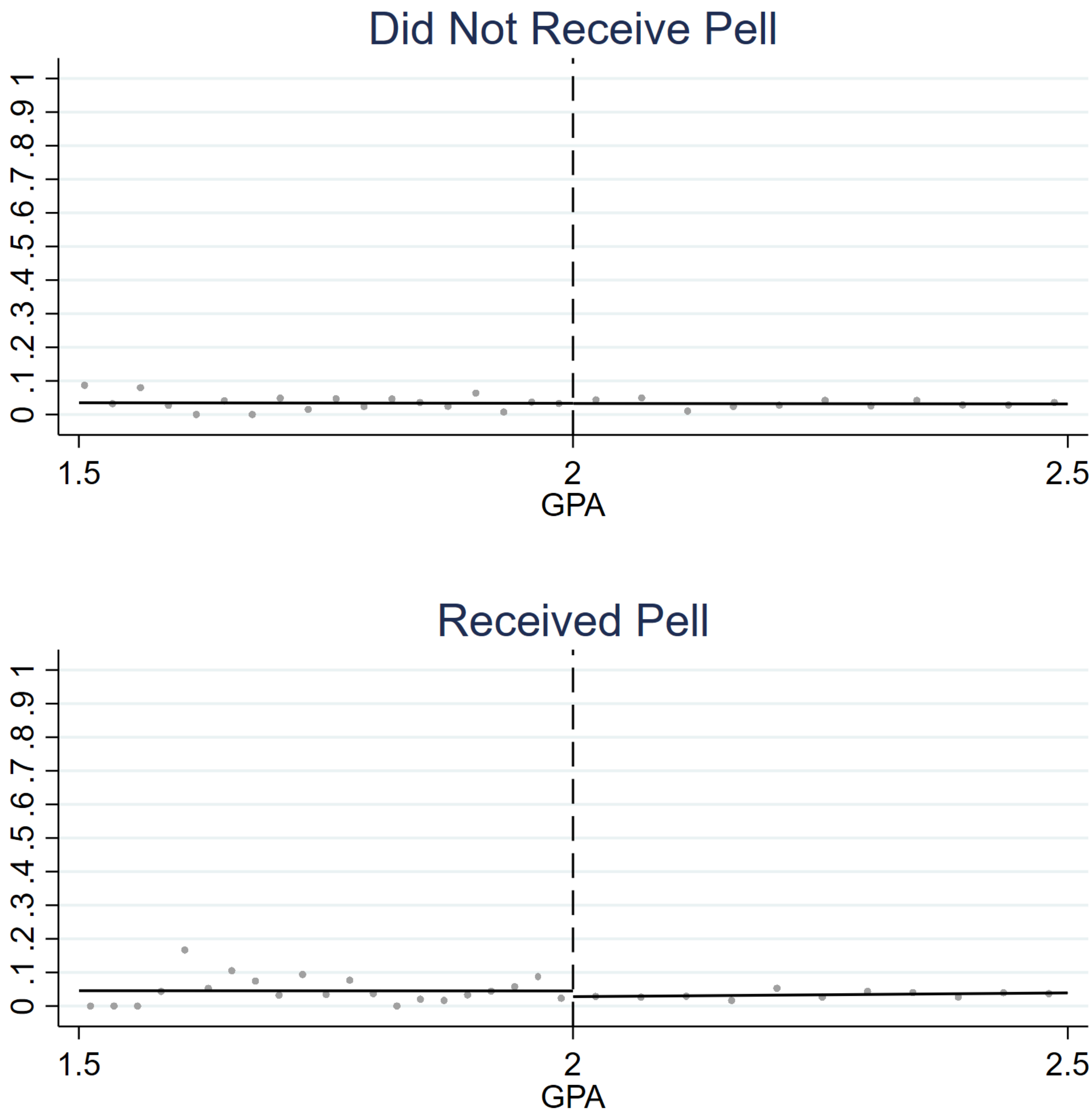
Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 3A.10 Discontinuities in Observables – Native American – Upperclassmen



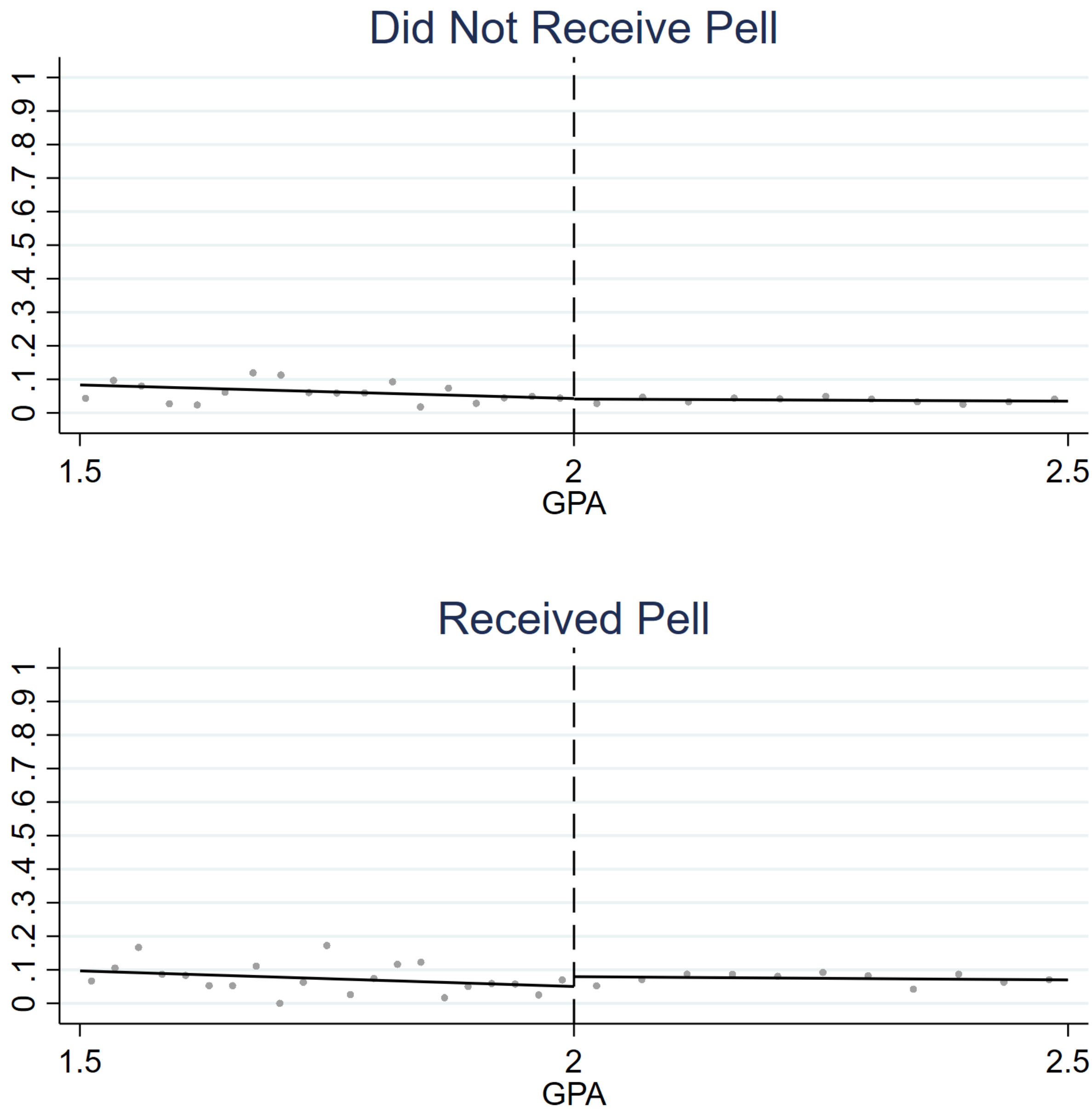
Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 3A.11 Discontinuities in Observables – Asian – Upperclassmen



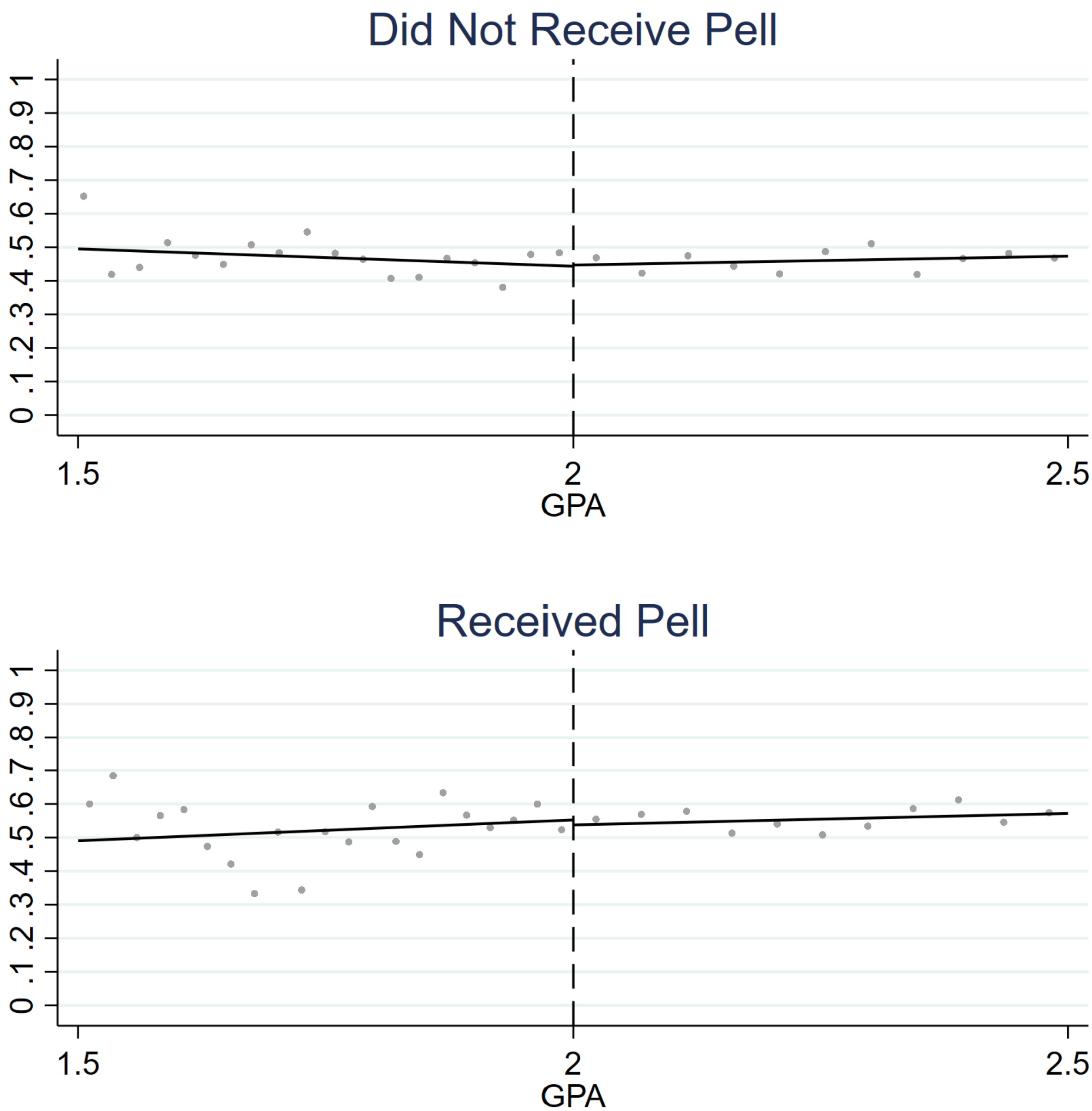
Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 3A.12 Discontinuities in Observables – Black – Upperclassmen



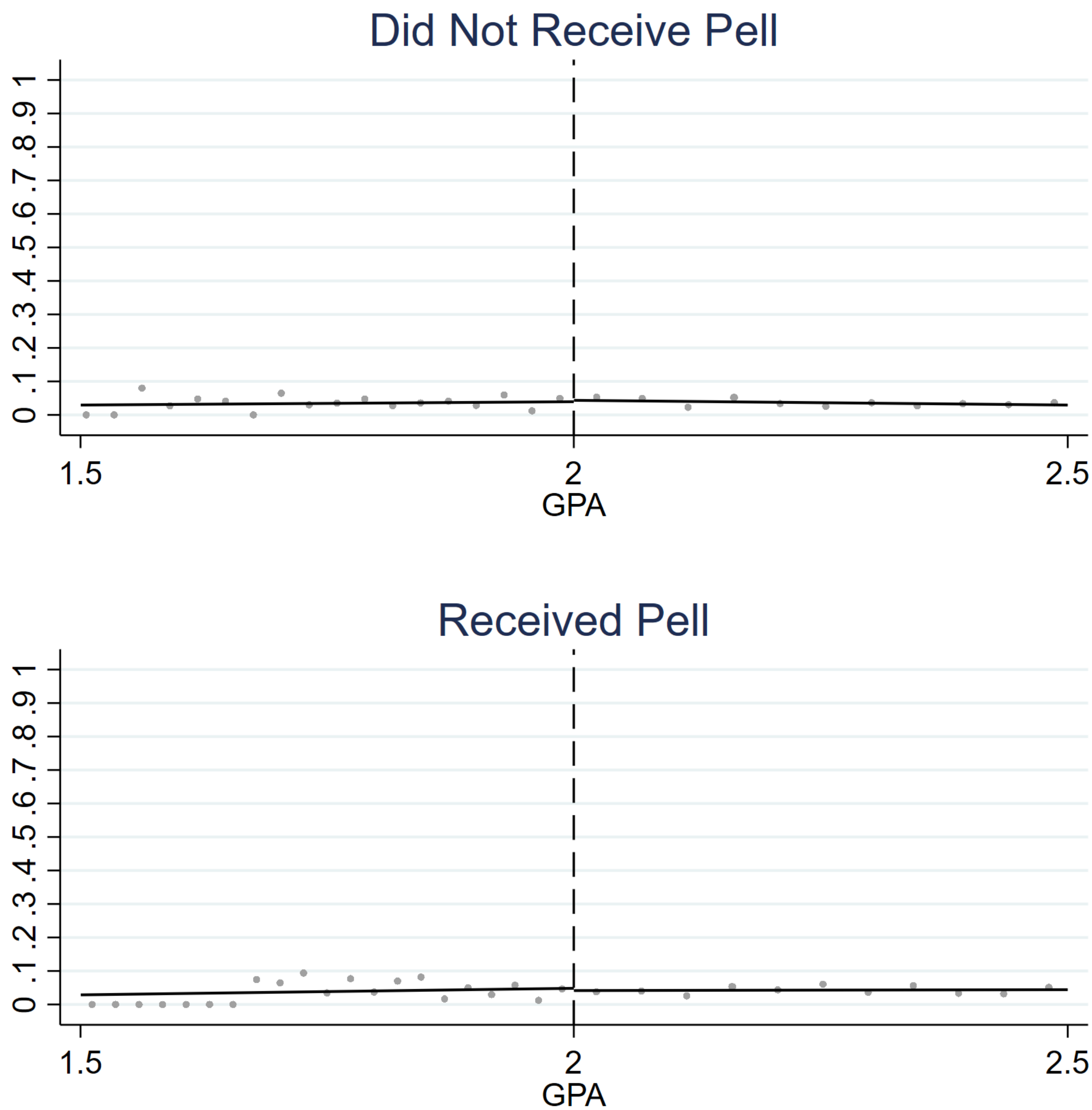
Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 3A.13 Discontinuities in Observables – Hispanic – Upperclassmen



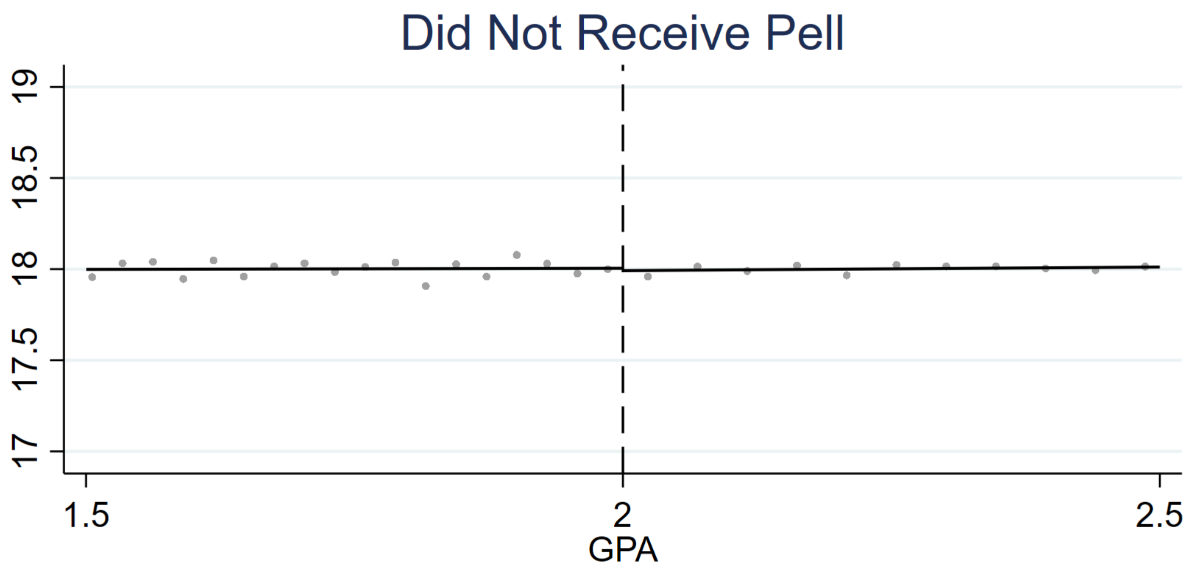
Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

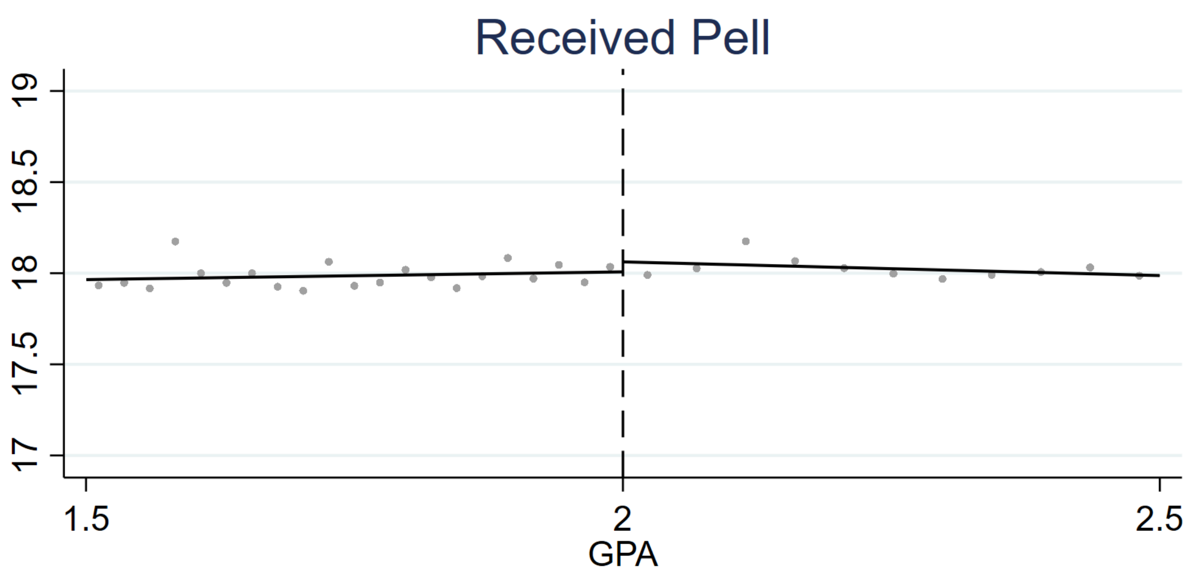
Figure 3A. 14 Discontinuities in Observables – Other Race – Did not receive Pell



Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

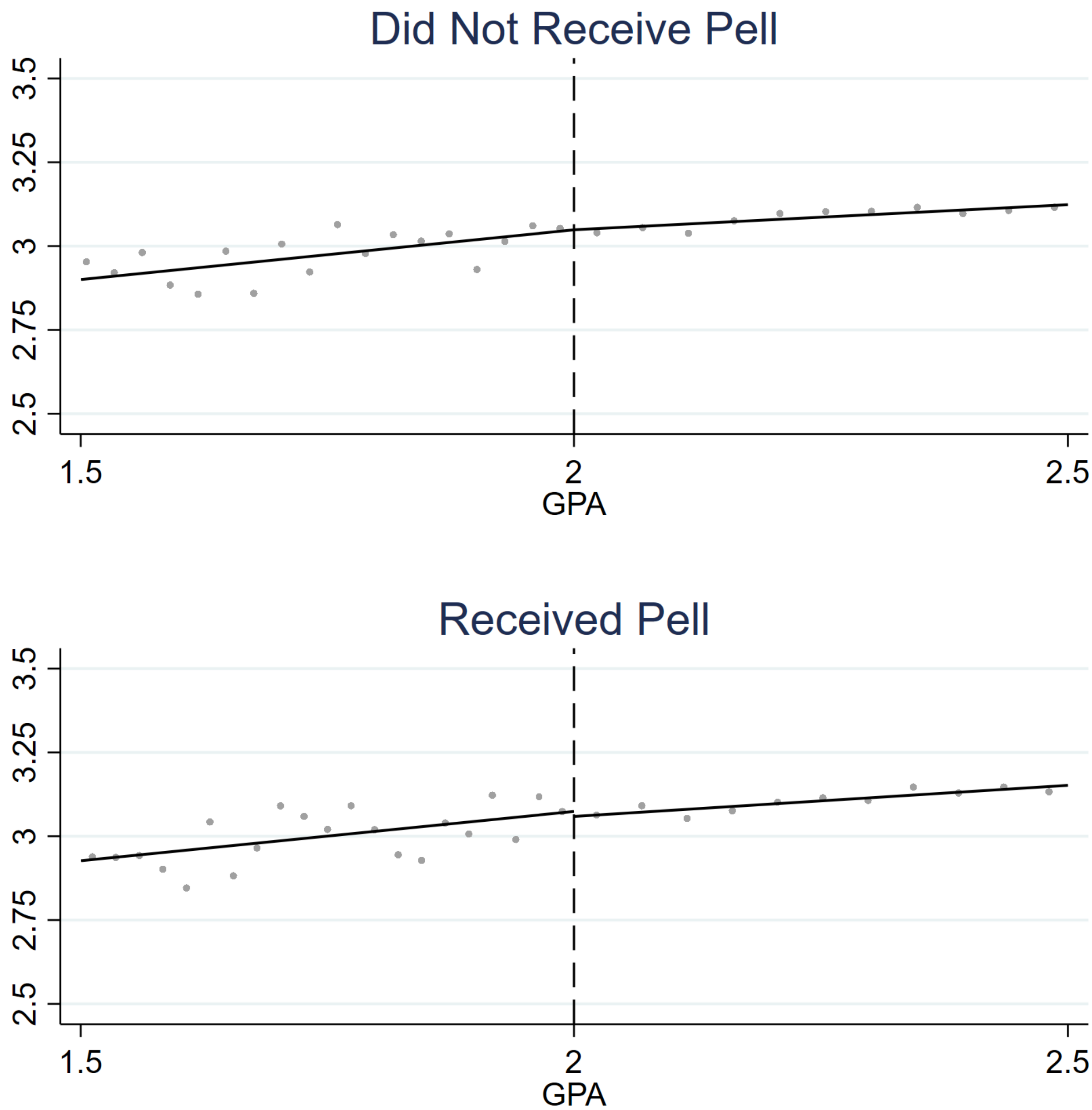
Figure 3A.15 Discontinuities in Observables – Age Admitted – Did not receive Pell





Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Figure 3A.16 Discontinuities in Observables – High School GPA – Upperclassmen



Regression discontinuity results for Female (indicator variable) following the semester they were placed on academic probation and did not meet satisfactory academic progress. *Notes*: Figures plot averages of variable interest relative to students' cumulative grade point average at time of being placed on academic probation, which occurs at a 2.0 GPA threshold shown by a dashed vertical line. Estimates are based on linear regression as described in empirical strategy section using Calonico, Cattaneo, and Titiunik (2014) data-driven optimal bandwidths.

Table 0A.1 Covariate Balance – Four-Year Graduation Study

	Non-Pell		Pell	
	SAP 1.7 GPA	Academic Probation 2.0 GPA	SAP 1.7 GPA	Academic Probation 2.0 GPA
Female	-0.136	0.006	0.039	-0.020
	(0.084)	(0.075)	(0.104)	(0.093)
Native American	0.028	0.027	-0.088	0.018
	(0.040)	(0.039)	(0.066)	(0.062)
Asian	-0.002	-0.001	-0.023	0.006
	(0.014)	(0.027)	(0.027)	(0.026)
Black	-0.008	0.046	-0.014	0.004
	(0.028)	(0.031)	(0.044)	(0.045)
Hispanic	-0.007	-0.026	0.143	0.172*
	(0.084)	(0.073)	(0.103)	(0.091)
Other	0.008	0.056*	0.025	-0.034
	(0.028)	(0.031)	(0.044)	(0.040)
Age at Admission	0.005	-0.072	0.037	-0.023
	(0.066)	(0.077)	(0.113)	(0.083)
High School GPA	0.032	0.070	-0.100	-0.007
	(0.064)	(0.062)	(0.086)	(0.081)
Observations	602	892	388	564

Note: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Bandwidth for each analysis is 0.29.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 0A.2 Covariate Balance – Five-Year Graduation Study

	Non-Pell		Pell	
	SAP 1.7 GPA	Academic Probation 2.0 GPA	SAP 1.7 GPA	Academic Probation 2.0 GPA
Female	-0.128	0.002	0.077	-0.003
	(0.088)	(0.078)	(0.110)	(0.101)
Native American	0.015	0.022	-0.067	0.009
	(0.042)	(0.042)	(0.072)	(0.067)
Asian	0.002	0.004	-0.038	0.014
	(0.015)	(0.028)	(0.028)	(0.030)
Black	-0.013	0.022	-0.011	0.010
	(0.030)	(0.030)	(0.050)	(0.053)
Hispanic	0.013	-0.038	0.148	0.176*
	(0.087)	(0.075)	(0.110)	(0.099)
Other	0.004	0.058*	0.009	-0.034
	(0.029)	(0.033)	(0.047)	(0.040)
Age at Admission	0.001	-0.042	0.082	-0.024
	(0.067)	(0.080)	(0.128)	(0.096)
High School GPA	0.048	0.101	-0.065	0.056
	(0.066)	(0.065)	(0.092)	(0.090)
Observations	553	819	339	483

Note: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Bandwidth for each analysis is 0.29.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Table 0A.3 Covariate Balance – Six-Year Graduation Study

	Non-Pell		Pell	
	SAP 1.7 GPA	Academic Probation 2.0 GPA	SAP 1.7 GPA	Academic Probation 2.0 GPA
Female	-0.120	0.028	0.035	-0.118
	(0.092)	(0.082)	(0.120)	(0.112)
Native American	0.012	0.023	-0.033	0.068
	(0.045)	(0.046)	(0.080)	(0.077)
Asian	0.001	0.004	-0.025	0.008
	(0.016)	(0.031)	(0.026)	(0.019)
Black	-0.013	0.024	-0.011	0.014
	(0.033)	(0.033)	(0.059)	(0.064)
Hispanic	-0.006	-0.050	0.056	0.096
	(0.091)	(0.079)	(0.120)	(0.111)
Other	0.001	0.073**	0.002	-0.018
	(0.028)	(0.036)	(0.053)	(0.046)
Age at Admission	0.024	-0.032	0.047	-0.061
	(0.068)	(0.088)	(0.140)	(0.111)
High School GPA	0.052	0.107	-0.078	-0.011
	(0.070)	(0.069)	(0.103)	(0.101)
Observations	504	734	297	420

Note: Estimated robust standard errors are displayed in parentheses. Estimates are based on linear regression as described in empirical strategy section. Bandwidth for each analysis is 0.29.
***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

References

Abdulkadiroğlu, A., Pathak, P. A., & Walters, C. R. (2018). Free to choose: Can school choice reduce student achievement?. American Economic Journal: Applied Economics, 10(1), 175-206.

Adams, S., Blackburn, M. L., & Cotti, C. D. (2012). Minimum wages and alcohol-related traffic fatalities among teens. Review of Economics and Statistics, 94(3), 828-840.

Amys, R. D. (2016). The four-day school week: Research on extended weekends (Doctoral dissertation, Montana State University).

Anderson, D. M., & Walker, M. B. (2015). Does shortening the school week impact student performance? Evidence from the four-day school week. Education Finance and Policy, 10(3), 314-349.

Angrist, J. D., & Keueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings?. The Quarterly Journal of Economics, 106(4), 979-1014.

Arsen, D., & Ni, Y. (2012). The effects of charter school competition on school district resource allocation. Educational Administration Quarterly, 48(1), 3-38.

Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2016). Heaping-induced bias in regression-discontinuity designs. Economic inquiry, 54(1), 268-293.

Barrow, L., & Rouse, C. E. (2018). Financial incentives and educational investment: The impact of performance-based scholarships on student time use. Education Finance and Policy, 13(4), 419-448.

Becker, G. S. (2009). Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago press.

Bergman, P., & McFarlin Jr, I. (2018). Education for all? A nationwide audit study of school choice (No. w25396). National Bureau of Economic Research.

Blagg, K., & Chingos, M. M. (2017). How Do School Funding Formulas Work? [Data Visualization]. Washington, D.C.: Urban Institute.

Bowman, N. A., Jang, N., Kivlighan, D. M., Schneider, N., & Ye, X. (2019). The Impact of a Goal-Setting Intervention for Engineering Students on Academic Probation. Research in Higher Education, 1-25.

Brunner, E., Hyman, J., & Ju, A. (2020). School finance reforms, teachers' unions, and the allocation of school resources. Review of Economics and Statistics, 102(3), 473-489.

Bruno, P. (2019). Charter competition and district finances: Evidence from California. Journal of Education Finance, 44(4), 361-384.

Buerger, C., & Bifulco, R. (2019). The effect of charter schools on districts' student composition, costs, and efficiency: The case of New York state. Economics of Education Review, 69, 61-72.

Burgess, S., & Briggs, A. (2010). School assignment, school choice and social mobility. Economics of Education Review, 29(4), 639-649.

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. Journal of Econometrics.

Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. Journal of the American Statistical Association, 113(522), 767-779.

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

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

Carpenter, C., & Dobkin, C. (2017). The minimum legal drinking age and morbidity in the United States. Review of Economics and Statistics, 99(1), 95-104.

Carruthers, C. K., & Özek, U. (2016). Losing HOPE: Financial aid and the line between college and work. Economics of education review, 53, 1-15.

Casey, M. D., Cline, J., Ost, B., & Qureshi, J. A. (2018). Academic Probation, Student Performance, and Strategic Course-Taking. Economic Inquiry, 56(3), 1646-1677.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G., & Keele, L. (2016). Interpreting regression discontinuity designs with multiple cutoffs. The Journal of Politics, 78(4), 1229-1248.

Cornelisz, I., van der Velden, R., De Wolf, I., & van Klaveren, C. (2020). The consequences of academic dismissal for academic success. Studies in Higher Education, 45(11), 2175-2189.

Cullen, J. B., Jacob, B. A., & Levitt, S. D. (2005). The impact of school choice on student outcomes: an analysis of the Chicago Public Schools. Journal of Public Economics, 89(5-6), 729-760.

Dee, T. S. (1998). Competition and the quality of public schools. Economics of Education review, 17(4), 419-427.

Dee, T. S. (1999). State alcohol policies, teen drinking and traffic fatalities. Journal of public Economics, 72(2), 289-315.

Dee, T. S., & Evans, W. N. (2001). Behavior policies and teen traffic safety. American Economic Review, 91(2), 91-96.

Dee, T. S., Grabowski, D. C., & Morrisey, M. A. (2005). Graduated driver licensing and teen traffic fatalities. Journal of health economics, 24(3), 571-589.

Doherty, S. T., Andrey, J. C., & MacGregor, C. (1998). The situational risks of young drivers: The influence of passengers, time of day and day of week on accident rates. Accident Analysis & Prevention, 30(1), 45-52.

Donis-Keller, C., & Silvernail, D. L. (2009). Research brief: A review of the evidence on the four-day school week. Center for Education Policy, Applied Research and Evaluation, University of Southern Maine.

EdChioce, 2019. The ABCs of School Choice 2019. [online] EdChoice. Available at: <https://www.edchoice.org/research/the-abcs-of-school-choice/> [Accessed 8 March 2021].

Bush, M. (2010). Compulsory School Age Requirements. [online]. Education Commission of the States. Available at: https://www.ncsl.org/documents/educ/ECSCompulsoryAge.pdf [Accessed 30 September 2021].

Egalite, A. J., & Wolf, P. J. (2016). A review of the empirical research on private school choice. Peabody Journal of Education, 91(4), 441-454.

Erwin, C., & Binder, M. (2020). Does broad-based merit aid improve college completion? Evidence from New Mexico's lottery scholarship. Education Finance and Policy, 15(1), 164-190.

Fischer, S., & Argyle, D. (2018). Juvenile crime and the four-day school week. Economics of education Review, 64, 31-39.

Friedman, M. (1997). Public schools: Make them private. Education Economics, 5(3), 341-344.

Gilpin, G. (2019). Teen driver licensure provisions, licensing, and vehicular fatalities. Journal of health economics, 66, 54-70.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing (No. w25018). National Bureau of Economic Research.

Gower, M. L. (2017). Interpreting the Impact of the Four-Day School Week: An Examination of Performance Before and After Switching to the Four-Day School Week (Doctoral dissertation, Lindenwood University).

Grau, E., & Shaughnessy, M. F. (1987). The Four Day School Week: An Investigation and Analysis.

Gulosino, C., & Liebert, J. (2020). Examining Variation Within the Charter School Sector: Academic Achievement in Suburban, Urban, and Rural Charter Schools. Peabody Journal of Education, 95(3), 300-329.

Gunderson, M., & Oreopolous, P. (2020). Returns to education in developed countries. In The economics of education (pp. 39-51). Academic Press.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. Journal of public Economics, 88(9-10), 1721-1746.

Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. Journal of public economics, 91(5-6), 823-848.

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. The Quarterly journal of economics, 123(4), 1373-1414.

Henderson, M. B., Houston, D. M., Peterson, P. E., & West, M. R. (2020). PUBLIC SUPPORT GROWS FOR HIGHER TEACHER PAY AND EXPANDED SCHOOL CHOICE RESULTS FROM THE 2019 EDUCATION NEXT SURVEY. Education Next, 20(1), 8-28.

Hoxby, C. M. (2003). School choice and school productivity. Could school choice be a tide that lifts all boats?. In The economics of school choice (pp. 287-342). University of Chicago Press.

Hill, P. T., & Heyward, G. (2015). The four-day school week in rural Idaho schools.

Hewitt, P. M., & Denny, G. S. (2011). The four-day school week: Impact on student academic performance. The Rural Educator, 32(2).

Imazeki, J., & Reschovsky, A. (2003). Financing adequate education in rural settings. Journal of Education Finance, 29(2), 137-156.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. The Review of economic studies, 79(3), 933-959.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. Journal of econometrics, 142(2), 615-635.

Jones, T. R., Kreisman, D., Rubenstein, R., Searcy, C., & Bhatt, R. (2020). The Effects of Financial Aid Loss on Persistence and Graduation: A Multi-Dimensional Regression Discontinuity Approach. Education Finance and Policy, 1-60.

Karaca-Mandic, P., & Ridgeway, G. (2010). Behavioral impact of graduated driver licensing on teenage driving risk and exposure. Journal of health economics, 29(1), 48-61.

Keall, M. D., Frith, W. J., & Patterson, T. L. (2005). The contribution of alcohol to night time crash risk and other risks of night driving. Accident Analysis & Prevention, 37(5), 816-824.

Ladd, H. F., & Singleton, J. D. (2020). The fiscal externalities of charter schools: Evidence from North Carolina. Education Finance and Policy, 15(1), 191-208.

Lindo, J. M., Sanders, N. J., & Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. American Economic Journal: Applied Economics, 2(2), 95-117.

Lovenheim, M. F., & Walsh, P. (2018). Does choice increase information? Evidence from online school search behavior. Economics of Education Review, 62, 91-103.

Mann, B. A., & Bruno, P. (2020). The effects of charter school enrollment losses and tuition reimbursements on school districts: Lifting boats or sinking them?. Educational Policy, 0895904820951124.

Mann, B., Kotok, S., Frankenberg, E., Fuller, E., & Schafft, K. (2016). Choice, cyber charter schools, and the educational marketplace for rural school districts. The rural educator, 37(3).

Martorell, P., McFarlin Jr, I., & Xue, Y. (2015). Does failing a placement exam discourage underprepared students from going to college?. Education Finance and Policy, 10(1), 46-80.

Morton, E. (2021). Effects of four-day school weeks on school finance and achievement: Evidence from Oklahoma. Educational Researcher, 50(1), 30-40.

Moss, B. G., & Yeaton, W. H. (2015). Failed warnings: Evaluating the impact of academic probation warning letters on student achievement. Evaluation Review, 39(5), 501-524.

Olneck-Brown, B., 2021. Charter Schools: Overview. [online] Ncsl.org. Available at: <https://www.ncsl.org/research/education/charter-schools-overview.aspx> [Accessed 18 October 2021].

Ost, B., Pan, W., & Webber, D. (2018). The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies. Journal of Labor Economics, 36(3), 779-805.

Pathak, P. A., & Shi, P. (2020). How well do structural demand models work? Counterfactual predictions in school choice. Journal of Econometrics.

Rice, T. M., Peek-Asa, C., & Kraus, J. F. (2003). Nighttime driving, passenger transport, and injury crash rates of young drivers. Injury Prevention, 9(3), 245-250.

Richards, R. L. (1990). A comparative study using selected characteristics of four-day week schools and five-day week schools in rural New Mexico (Doctoral dissertation, Texas Tech University).

Rios-Avila, F., Sant'Anna, P., & Callaway, B. (2021). CSDID: Stata module for the estimation of Difference-in-Differences models with multiple time periods.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Sabia, J. J., Pitts, M. M., & Argys, L. M. (2019). Are minimum wages a silent killer? New evidence on drunk driving fatalities. Review of Economics and Statistics, 101(1), 192-199.

Sagness, R. L., & Salzman, S. A. (1993). Evaluation of the Four-Day School Week in Idaho Suburban Schools.

Sant’Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. Journal of Econometrics, 219(1), 101-122.

Schudde, L., & Scott-Clayton, J. (2016). Pell grants as performance-based scholarships? An examination of satisfactory academic progress requirements in the nation’s largest need-based aid program. Research in Higher Education, 57(8), 943-967.

Scott-Clayton, J., & Schudde, L. (2020). The Consequences of Performance Standards in Need-Based Aid Evidence from Community Colleges. Journal of Human Resources, 55(4), 1105-1136.

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.

Shope, J. T., & Bingham, C. R. (2008). Teen driving: motor-vehicle crashes and factors that contribute. American Journal of Preventive Medicine, 35(3), S261-S271.

Sneyers, E., & De Witte, K. (2018). Interventions in higher education and their effect on student success: a meta-analysis. Educational Review, 70(2), 208-228.

Strauss, V., Douglas-Gabriel, D., & Balingit, M. (2018). DeVos seeks cuts from education department to support school choice. Washington Post, 13.

Tharp, T. W., Matt, J., & O'Reilly, F. L. (2016). Is the Four-Day School Week Detrimental to and Student Success?. Journal of Education and Training Studies, 4(3), 126-132.

Thatcher, D., 2021. School Finance Litigation Citations (NCSL). [online] ncsl.org. Available at: <https://www.ncsl.org/research/education/state-role-in-education-finance.aspx> [Accessed 8 March 2021].

Thattai, D. (2001). A history of public education in the United States.

Thompson, P. N. (2021). Is four less than five? Effects of four-day school weeks on student achievement in Oregon. Journal of Public Economics, 193, 104308.

University of Kentucky Center for Poverty Research. 2018. "UKCPR National Welfare Data, 1980-2017." Lexington, KY. http://ukcpr.org/resources/national-welfare-data (accessed Dec 12, 2018)

U.S. Department of Education (2018). National Center for Education Statistics, Common Core of Data (CCD), "Public Elementary/Secondary School Universe Survey," 2000-01 through 2016-17.

U.S. Department of Education, National Center for Education Statistics (2019). Integrated Postsecondary Education Data System (IPEDS): Institutional Characteristics component and Winter 2019-20, Student Financial Aid component.

U.S. Department of Education (2020). National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS): Fall Enrollment component.

Vorona, R. D., Szklo-Coxe, M., Lamichhane, R., Ware, J. C., McNallen, A., & Leszczyszyn, D. (2014). Adolescent crash rates and school start times in two central Virginia counties, 2009-2011: a follow-up study to a southeastern Virginia study, 2007-2008. Journal of clinical sleep medicine, 10(11), 1169-1177.

Ward, J. (2019). The Four-day School Week and Parental Labor Supply. Available at SSRN 3301406.

Welch, J. G. (2014). HOPE for community college students: The impact of merit aid on persistence, graduation, and earnings. Economics of Education Review, 43, 1-20.

Wright, N. A. (2020). Perform better, or else: Academic probation, public praise, and students decision-making. Labour Economics, 62, 101773.

Yeaton, W. H., & Moss, B. G. (2018). A Multiple-Design, Experimental Strategy: Academic Probation Warning Letter's Impact on Student Achievement. The Journal of Experimental Education, 1-21.