ESSAYS IN THE ECONOMICS OF EDUCATION POLICY

By

Kaitlyn E. Elgart

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 10, 2024

Nashville, Tennessee

Approved:

Lesley Turner, Ph.D.

Analisa Packham, Ph.D.

Michelle Marcus, Ph.D.

Jason Grissom, Ph.D.

ACKNOWLEDGMENTS

Completing this dissertation was far from a solo endeavor, and I have many people to thank for helping me achieve this goal. I am very grateful to my adviser, Lesley Turner, for being an excellent mentor and teacher throughout my graduate school experience and for always holding me to high standards. For their mentorship and encouragement, I thank members of my dissertation committee, Analisa Packham, Michelle Marcus, and Jason Grissom. I am grateful for the comments and feedback received from the faculty in the Vanderbilt Economics Department, my colleagues at the Tennessee Education Research Alliance, and the collaboration of my coauthor Susan Kemper Patrick. This dissertation would not have been possible without the advice, encouragement, and friendship of my classmates, and I am especially grateful to Jessica Kiser, Sam Crowell, Kayleigh McCrary, and Tucker Smith.

I am eternally grateful to my amazing friends and family who have supported me on this journey over the past five years. To my friends, Natasha Levitin, Lizzy Hedrick, Cara Miller, and many others, thank you for visiting me, believing in me, and always picking up the phone when I needed you. To the friends I have gained during my years in Nashville, thank you for making this city a home. I must also thank my fluffy four-legged research assistant, Penny, for bringing me lots of joy, snuggles, and long walks at the park on study breaks. Finally, I am most grateful to all of my family for their endless love and support in my graduate school pursuits and beyond. In particular I thank my mom, Kim Elgart, for answering many, many, many phone calls and for being my biggest cheerleader in all things. Thank you to my dad, Russ Elgart, for being my favorite proofreader and for the many morale-boosting dinners together in Nashville over the years. Thank you to my brother, Evan Elgart, for always knowing how to make me laugh and for all the weekends spent together in Tennessee and Alabama at concerts and football games.

This dissertation is dedicated to the women who raised me: my mother Kim, and in memory of my grandmothers Hilda and Linda. Thank you for teaching me to dream big.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1 Motivations and Consequences of School Policing: Evidence from the COPS Hiring Program

Kaitlyn Elgart[1]

1.1 Introduction

Law enforcement officers have been present in U.S. schools since the 1950s, when the first school police officers were assigned in Flint, Michigan as a way to combat crime on the school grounds (Counts et al. 2018). However, over the last 50 years, the proportion of students in the U.S. attending a school with an armed police officer present has dramatically increased. In 1970, less than 5% of schools had a regular law enforcement presence. In 2019, more than half of all public schools and 70% of public high schools reported having an armed School Resource Officer (SRO) present on their campus (NCES 2021).

The policy debate surrounding SROs and safety in schools has intensified in recent years. Some states have passed laws requiring SROs to be stationed in all public schools in the aftermath of school shooting incidents (Florida State Legislature, 2018; Maryland Association of Boards of Education, 2018), given that prior research has shown that increased police presence overall can be an effective crime deterrent (Levitt 2002; Evans and Owens 2007). Conversely, several major school districts such as Minneapolis and Denver chose to cut ties with their local police departments and remove police from schools in response to the Black Lives Matter movement and concerns about the School-to-Prison pipeline[2]. Additionally, advocates' concerns about the unintended consequences of placing police in schools, particularly for students from disadvantaged backgrounds or minority students, have influenced policy discussions surrounding these staffing decisions.

In this paper, I study the effect of school-based policing on student discipline outcomes using evidence from a federal grant program which provided funding for local law enforcement agencies (LEAs) to place police officers in schools. The Community Oriented Policing Services (COPS) hiring program, funded by the U.S. Department of Justice, provides three-year grant awards to local police departments for law enforcement hiring purposes. I exploit variation in timing of grant receipt in a difference-in-differences framework to examine the causal impact of receipt of a COPS grant for school-based policing (SBP) on student discipline outcomes. I define schools as treated if there is a police agency within their school district boundaries that applied for and received a COPS hiring grant for school policing, and my control group consists of school districts that contain an agency that applied for and did not receive the grant. I explore important heterogeneous effects of these grants by student demographics, grade level, and district-level demographic composition. Further, I use natural language processing tools to analyze the content of grant applications to motivate results and investigate heterogeneous effects by grant application content and

[1] Many thanks to Lucy Sorensen, Shawn Bushway, and Chris Curran for their helpful insight, guidance, and feedback during the development and writing of this paper. Thank you also to the staff at the COPS Office at the U.S. Department of Justice for their assistance with procuring much of the data used in this analysis.

[2] The school-to-prison pipeline refers to the disproportionate tendency of minors from disadvantaged backgrounds to be incarcerated due to harsh school discipline policies, such as zero tolerance policies.

sentiment. I find that, on average, suspension rates for Hispanic students increase 10-15% relative to the mean following grant receipt. I also show that grant application content can influence these heterogeneous effects, and that applications from school districts with a greater share of Black or Hispanic students are more likely to include negative language, which trickles down to worse outcomes for all students in those districts.

Extensive research on policing has shown that increased exposure to police presence can reduce overall crime rates (Levitt 2002; Evans and Owens 2007; Weisburst 2019b). However, direct interactions with police can cause decreased physical and mental well-being for civilians (Geller et al. 2014; Ang 2020). Public confidence in the police has decreased over time, and Black and Hispanic Americans are more likely to report low confidence in policing (E. Owens and Ba 2021), although nonwhite Americans are more likely to be victims of a serious violent crime (Harvey 2019). Despite potential benefits of policing in the form of deterrence of crime, a growing body of research documents persistent racial disparities in policing (Gelman, Fagan, and Kiss 2007; Goel, Rao, and Shroff 2016). Exposure to police at younger ages can have negative consequences for psychological well-being, as research has shown that high-school aged students exposed to police violence experience reductions in GPA, high school graduation, and college enrollment (Ang 2020). This finding prompts a further need to understand how increased interactions with police within the school environment can influence young people.

My findings build on a growing literature on police presence in schools, which has shown that interactions with police in the controlled school environment can increase students' positive attitudes about police in general (Theriot 2016), but students differ in whether they feel safer with a School Resource Officer on campus by demographic background, with black students reporting they feel less safe (Theriot and Orme 2016). Other studies have shown that the presence of an SRO does not influence students' perceptions or attitudes toward police (Jackson 2002). Police presence in schools has been associated with increased numbers of crimes recorded at schools (Na and Gottfredson 2013; Gottfredson et al. 2020), and more drug and weapons violations (Zhang 2019; Gottfredson et al. 2020), which could be due to increased detection. Additionally, while school police are typically not authorized to prescribe student disciplinary sanctions, SROs are often involved in influencing school discipline practices (Kupchik 2010; Curran et al. 2019), and several studies have shown that increased presence of SROs increases disciplinary rates within the school (Weisburst 2019a; Sorensen et al. 2023), as well as referrals of students to law enforcement (E. G. Owens 2017). Further, several studies have shown that minority students are more likely than their White peers to be disciplined and referred to law enforcement as SRO presence increases (Weisburst 2019a; Sorensen, Shen, and Bushway 2021). Despite a growing body of literature which demonstrates that SRO presence can have the unintended consequence of increased student disciplinary rates, several studies also show that SRO presence is often correlated with improved school safety in the form of reductions in lower-level disruptions and offenses on school grounds (E. G. Owens 2017; Sorensen et al. 2023).

Policymakers have continued to push for increased law enforcement presence in schools as a popular policy lever in improving school safety (Viano, Curran, and Fisher 2021). While several states have elected to require SROs in schools as a response to school shooting incidents, some studies have shown that SRO presence is not correlated with increased school safety in the form of reduced crime rates or reductions in shooting incidents

(Sorensen et al. 2023), and that the presence of an armed guard on campus during a school shooting incident in fact predicts a higher casualty rate (Peterson, Densley, and Erickson 2021). Identifying the underlying motivations for introducing School Resource Officers into the school environment is important for understanding how the SRO views and performs their role within the school, and how those motivations can trickle down to student outcomes. Prior research has demonstrated that district-level racial composition can influence how SROs perceive their primary role within their schools. In a study of two large school districts, SROs in a predominantly white school district report primarily focusing their efforts on threats to student safety from outside the school (i.e. intruders), whereas SROs in a predominantly minority school district report their role as being primarily focused on threats from within the school (i.e. student misbehavior, fights) (Fisher et al. 2022). Studies have also shown that students in districts with a greater share of minority enrollment can face heightened blameworthiness for identical misbehavior as districts with fewer minority students, due to heterogeneity in teacher and administrator climates surrounding discipline (J. Owens 2022).

School disciplinary practices are historically complex, and policies and practices regarding discipline vary across schools, districts, and states. On average, the annual use of exclusionary discipline in U.S. schools can generate social costs of around $35 billion (Rumberger and Losen 2016). Juveniles who are arrested are less likely to complete high school (Hjalmarsson 2008), more likely to be arrested in the future (Aizer and Doyle 2015), and potentially ineligible for some federal grants and loans to assist in college-going (Lovenheim and Owens 2014). School discipline, while less severe than formal interactions with law enforcement, could lead to future involvements with the criminal justice system via the school-to-prison pipeline. Half of all police agencies employing SROs that were surveyed by the U.S. DOJ in 2019 reported that school resource officers are authorized to interview students without parental permission. If increased law enforcement presence through the use of SROs on campus increases juvenile arrest rates or school discipline rates, it could have ripple effects on educational attainment, human capital formation, and future law enforcement involvement for students.

Despite this renewed policy relevance, widespread adoption of SROs, and public interest in both school safety and student well-being, research on the causal impacts of police presence in schools is somewhat limited by data constraints and the inherent endogeneity in SRO staffing decisions. Because of the nature of SRO employment (SROs are employed by local police agencies rather than schools), School Resource Officer staffing information at the school level is difficult to obtain, making it a data challenge to identify which schools employ SROs on their campus for large-scale studies. In recent years, the Civil Rights Data Collection (CRDC) from the U.S. Department of Education, a survey of every public school in the U.S., has asked respondents to report SRO presence on campus. While this CRDC data on SRO presence is known to be underreported when compared to national-level estimates of SRO presence, it is one of the first comprehensive data sources identifying SRO presence at the school-level. In addition to data limitations, quasi-experimental research on the impacts of SRO presence has been limited by a lack of plausibly exogenous variation in SRO placements, as the decision to place an SRO on campus can be influenced by a number of factors, making simple comparisons of schools with and without an SRO difficult to interpret. Several recent studies have exploited plausibly exogenous variation in SRO presence at the school and district level imposed by grant programs aimed at providing funds to local law enforcement agencies to place SROs in schools (E.

G. Owens 2017; Weisburst 2019a; Sorensen et al. 2023). Owens (2017) and Weisburst (2019a) leverage the “Cops in Schools” program funded through the U.S. Department of Justice from 1999-2009 and exploit variation in timing and size of grants to examine impacts on crime, student misbehavior, and school discipline. Owens (2017) utilizes the School Survey on Crime and Safety (SSOCS) to provide evidence that receipt of increased funds for SROs increases school safety, but that this comes at the cost of increased arrest rates in schools. Weisburst (2019a) utilizes student-level data in the state of Texas and variation in grant receipt timing to analyze student discipline, and finds that grant receipt is associated with an increase in middle school discipline rates, which is larger for Black students. Weisburst (2019a) also finds that exposure to police in schools results in a reduction in high school graduation rates and college-going rates. The most recent study, Sorensen et al. (2023) uses variation in application scoring cutoffs for COPS hiring grants in a regression discontinuity design to find that schools who receive grant funds experience increased rates of suspensions, police referrals, and arrests, and that these effects are larger for Black students than for White students. Owens (2017) and Sorensen et al. (2023) both find that while SRO presence increases school safety in the form of reductions in lower level offenses and disruptions recorded in schools, this could come at the cost of increasingly harsh disciplinary responses for students.

While the above papers contribute significantly to our understanding of the causal effects of SROs in schools through the use of quasi-experimental methods, Sorensen et al. (2023) is unable to investigate school-level heterogeneity due to their limited sample size, and Owens (2017) is unable to investigate effects on formally recorded school discipline. While Weisburst (2019a) and Sorensen et al. (2023) investigate heterogeneous effects of SRO policies on student outcomes by student demographic characteristics, none of these studies consider school or district-level demographic composition or sentiment around policing as a margin for heterogeneity, despite the fact that other qualitative research has indicated that these factors can play in important role in the implementation of SRO programs in schools. Additionally, public perceptions of police and policy surrounding school policing have changed dramatically in the last decade, prompting the need for further research on the consequences of school policing in the current landscape.

In the current study, I leverage variation in district-level SRO funding generated by participation in the COPS Hiring Program grant application cycles from 2014-2017. To understand the underlying motivations for participation in this program and SRO placement, I first use natural language processing tools to analyze sentiment and key themes of applications. I find that while sentiment does not vary significantly across funded and unfunded applications, application sentiment is negatively correlated with the share of minority students in a district. In my causal analysis, I find that this pattern of negative sentiment and extreme language utilized by districts with greater shares of minority students can trickle down to worse outcomes for students in those districts. In my primary specification, I utilize a difference-in-differences approach that accounts for variation in timing of grant receipt. I compare student discipline outcomes within school districts with successful and unsuccessful grant applications to estimate the effect of increased funding and placement of SROs on student disciplinary outcomes. I find no evidence of increases in disciplinary rates on average for all students within a district which receives this increased SRO funding, but these average effects mask important underlying heterogeneity by student, school, and district characteristics. On average, suspension rates for Hispanic students increase significantly following grant receipt, while suspension rates for Black high school students decrease. I find evidence that these results are driven by district-level differences in student racial composition and themes in application content. In school districts with mostly Hispanic students, all students experience increases in suspension rates following grant receipt. However, in school districts with mostly Black students, all students experience reductions in the probability of having multiple suspensions, and high school students are more likely to be arrested or referred to law enforcement. Taken together, these results suggest that police presence in schools can shift the threshold for punishment severity in the form of exclusionary discipline or law enforcement involvement, particularly in schools with a large share of minority students.

1.2 Background

A School Resource Officer is defined by the National Association of School Resource Officers (NASRO) as "a career law enforcement officer with sworn authority who is deployed by an employing police department or agency in a community-oriented policing assignment to work in collaboration with one or more schools". SROs are generally armed, always uniformed police officers who are assigned to patrol local schools, and are responsible for safety and crime prevention at the schools they serve. NASRO estimates that there are between 14,000 and 20,000 SROs currently in service nationwide. On average, around 78% of SROs are White, and 80% of SROs are male (BJS 2019). SRO exposure is increasing in student age; in 2019 the National Center for Education Statistics (NCES) reported that approximately 40% of elementary school students, 68% of middle school students, and 70% of high school students in public schools attended a school with sworn law enforcement officers present. Exposure is also increasing in school enrollment size; approximately 83% of public schools with 1,000 or more students enrolled report law enforcement presence in 2019. Schools can have multiple SROs on campus, and official NASRO recommendations for SRO staffing suggest that one officer per 1,000 students is an optimal staffing ratio (NASRO 2023).

School Resource Officers are trained police officers, but SRO-specific training requirements vary by state and jurisdiction. In a 2019 survey administered by the U.S. Department of Justice, agencies reported varying degrees of training requirements of SROs including training on crisis preparedness planning (82% of agencies), security assessments of campuses (73% of agencies), and de-escalation strategy training (93% of agencies). Additionally, the role of a School Resource Officer within the school environment varies across schools. Many schools utilize SROs beyond their law enforcement responsibilities in assisting with student misbehavior, as well as in a mentor-like capacity to host assemblies on law-related topics, or to generally educate students on police and the law. Selection and recruitment of SROs varies greatly by school and district need, law enforcement interest, and training requirements of the officers (Finn et al. 2005).

The Community Oriented Policing Services (COPS) Hiring Program, administered through the COPS Office at the U.S. Department of Justice, provides three-year grant awards to local law enforcement agencies for the recruitment, retention, and hiring of career law enforcement officers. All state, local, territorial, and tribal law enforcement agencies are eligible to apply for this program. Across fiscal years 2014-2017, the COPS office distributed 183 grants to local law enforcement agencies specifically for school-based policing requests from a total of 1,222 school-based policing applications. Over $45 Million was distributed by the COPS office for school policing efforts over this time period. Conditional on grant receipt, the average law enforcement agency was awarded $255,000 and was funded for an average of 2 school resource officers.

Figure 1.1 shows the geographic distribution of COPS hiring program applicants and award recipients for years 2014-2017 throughout the contiguous United States.[3] There is considerable geographic variation in grant applicants and recipients, illustrating the nationwide visibility of this program as well as the potential for heterogeneous effects of program implementation. Figure 1.2 displays the timing in grant applications by their funded status, and illustrates the considerable variation in grant distribution timing across these years.

While the School-Based Policing grants from the COPS office are awarded specifically to staff SROs in local schools, the applications are fielded from and awarded to local law enforcement agencies, rather than schools or school districts. The COPS Office scores applications on several criteria, including community policing score, crime score, fiscal need score, and miscellaneous bonus points[4]. Figure 1.3 displays the distribution of application scores by funded status, and shows that funded status was not solely determined by the highest overall scoring applications in the sample. In each application year, part of the COPS funding formula requires that 0.5% of total funds be allocated to applicants in each state, to ensure there are not clusters of grant awards in specific states. Additionally, the funding formula ensures that funds are awarded equally to both large and small localities, requiring that half of funds are allocated to agencies serving populations of more than 150,000 residents and half of funds are allocated to agencies serving populations of less than 150,000 residents. Local law enforcement agencies that receive grant funds for school-based policing are eligible for the award funds for three years, and must submit a formalized Memorandum of Understanding (MOU) between the law enforcement agency and their local school district outlining the school-based policing partnership between the two entities. School Resource Officers that are funded by the COPS Hiring Program are required to complete the NASRO Basic School Resource Officer course (at no cost to the agency), a 40-hour training course designed to prepare school resource officers to fulfill their role in the school setting effectively. In some states, this training requirement is above and beyond the baseline training requirements for SROs that are not funded with federal grant money.

Table 1.1 presents summary statistics on student demographics and staffing ratios at the school district level for school districts whose local law enforcement agency applied for COPS school-based policing grants, and police agency demographics at the agency level. School districts whose local law enforcement agency received funding over the study period tended to have higher shares of minority students, English Language Learners (ELL), higher student-to-teacher ratios and more schools within the district. Funded police agencies generally have more total SROs employed in 2016, fewer white officers as a percentage of total police force, and more female officers.

[3] There are several applicants and recipients in Alaska and Hawaii that are not displayed in Figure 1.1 but are included in the main analysis.
[4] The purpose and definition of bonus points can vary across years, and often reflect current preferences of Department of Justice leadership (Sorensen, Lucy C. et al. 2023).

1.3 Application Text Analysis

In grant application data received via a Freedom of Information Act request to the COPS Office at the U.S. DOJ, I observe text fields from every application received for school-based policing funds from the COPS Hiring Program in years 2014-2017. In each grant application, the applicant agency was asked to describe the "Problem/Focus Area" of the grant, which was primarily used to describe the ways the agency intended to utilize the funded SRO, and any issues or programming the agency planned to address in local schools. In order to understand the motivations and general themes underlying agencies' applications for this federal assistance, I employ natural language processing techniques to analyze the grant application content.

First, I use a sentiment analysis tool to assess the overall emotional tone of the grant applications. This approach is implemented in Python, and produces two main output variables for each grant application: a binary variable for whether the overall grant application sentiment is positive or negative, and a running variable between 0 and 100 which represents the overall percent of an application which is positive. The sentiment analysis approach in Python uses natural language processing techniques and a pre-programmed library to read each grant application text field and assign positive or negative weights to each word within the application. A grant application with a descriptor of 'positive' sentiment is one wherein the majority of words used in the application text would be weighted with positive sentiment by the processing tool. The continuous sentiment variable allows for a more nuanced understanding of the level of positive sentiment expressed in each application, and allows me to look at the distribution of sentiment across various applicant characteristics. I link sentiment output from this tool to district-level characteristics of applicant agencies to analyze how sentiment differs across demographics and funded status.

I summarize the output for both the binary sentiment variable and the running variable in Table 1.2 by funded status and by district-level racial composition. As shown, while I do not observe statistically significant differences in overall sentiment between funded and unfunded applications, I find that sentiment varies considerably by the demographic composition of the students in the partner school district. In particular, applications fielded from agencies whose school districts have a greater share of Black or Hispanic students are more likely to have negative sentiment, whereas applications from predominantly white districts contain more positive sentiment. While I find differences in sentiment by district racial composition, I find that this does not affect selection into funding. In Figures 1.4 and 1.5, I plot the distribution of sentiment by funded status and demographic composition, which further illustrates this pattern. While I am not able to reject the null hypothesis that the distribution of sentiment is different between funded and unfunded applications in Figure 1.4, I reject the null hypothesis that sentiment distributions are the same when grouping by school-district racial composition in Figure 1.5.

To further understand the key themes underlying these patterns in application sentiment and overarching motivations for applying to the COPS grant program, I utilize a keyword analysis in Python to extract top keywords from the application Problem/Focus Area field. Using language processing tools in Python, I extract the most frequently used keywords by total mentions in all applications within a certain subgroup. I export files which include the keyword, number of total counts across all applications, as well as number of applications mentioning that keyword.[5]

I analyze keywords and themes for funded and unfunded applications, and by district student demographic enrollment. This allows me to determine which keywords are more predictive of funding receipt, as well as compare application themes by demographic composition to understand how key sentiment surrounding school policing could differ by enrollment demographics. I summarize the primary themes in applications fielded for the school-based policing grants in Tables 1.3-1.6. In these tables, I present the top keywords by difference in relative frequency between two subgroups of applications, highlighting the major thematic differences between these groups of applicants. As shown in Table 1.3, I find that funded applications are more likely to use 'extreme' language such as 'gang' and 'violent' than applications that do not receive funding. While I do not find that funded applications are significantly different in overall sentiment on average than unfunded applications (Figure 1.4), I find evidence that there are more frequent mentions of extreme language. This could indicate that funded applications do a better job of coupling extreme or negative themes with positive language surrounding other aspects of the application, such as solutions or programming. Further, I find considerable differences in themes and language by district racial composition. Applications from school districts where a majority of students enrolled are Hispanic are more likely to mention themes around 'policing' students, including major keywords such as 'gang', 'members', 'violence', and 'poverty'. School districts with more Black students enrolled are more likely to point to specific campus disturbances as a theme, highlighting keywords such as 'gang', 'fights', and 'theft'. Conversely, predominantly White districts are more likely to mention themes surrounding relationship building with the school community, calling out keywords such as 'relationship', 'assist', and 'children'. Given that the overall distribution of sentiment also varies significantly for these applicant subgroups, we can conclude that the content of these applications is considerably different on several dimensions.

This application text analysis provides motivation for investigating heterogeneous effects of grant receipt on student discipline by district racial composition. While many previous studies have investigated the heterogeneous effects of school policing on student discipline by student race (Weisburst 2019a; Sorensen et al. 2023), few have considered how district racial composition and sentiment surrounding policing could further influence the downstream effects of SRO presence. Prior research has shown that SROs view their primary role within the school environment differently based on district racial composition of the student body (Fisher et al. 2022), and that minority students in minority-dominated schools face heightened blameworthiness as compared to students in schools with fewer minority students overall (J. Owens 2022). To my knowledge, this study is the first to attempt to connect these two bodies of literature by investigating the downstream causal effects of differences in sentiment surrounding school police on student discipline outcomes. If the district-level sentiment and key themes surrounding the motivations for placing police in schools can differentially influence the downstream effects for students within those schools, it could have important policy implications for understanding the heterogeneous effects of school policing programs.

[5] For example, if the keyword 'school' was used three times each in two separate applications, my code would assign the keyword 'school' to a score of 6 for total count, and a score of 2 for total mentions in applications.

1.4 Data

This project utilizes data obtained through several Freedom of Information Act (FOIA) requests to the COPS Office at the U.S. Department of Justice. The data provided includes detailed information on all applicants to the 2014-2017 COPS hiring program grant cycles who indicated their application was for School-Based Policing funds. I obtain information on Law Enforcement Agency Name, Originating Agency Identifier (ORI), year of application, written responses indicating the "Problem/Focus Area" to be addressed by this grant, number of officers requested, number of officers funded, total grant award amount, and final application score. I utilize data from the Bureau of Justice Statistics on the location and administrative attributes of these local law enforcement agencies to identify exact geographic location of grant applicants. I then link LEAs to the school district(s)[6] in which they are located to connect them to the schools that are most likely to be impacted by increased SRO presence after grant receipt.[7]

In addition to program data obtained from the COPS office, I incorporate law enforcement survey data from the 2007 and 2016 waves of the Law Enforcement Management and Administrative Statistics (LEMAS) Survey to observe various agency-level outcomes and characteristics. The LEMAS 2007 and 2016 surveys include information on the size and demographics of the police force by assignment, including a variable for the number of SROs employed by the agency. LEMAS is a survey of all law enforcement agencies that employ 100 or more officers, and a representative sample of all other agencies. From the subsample of law enforcement agencies that applied for COPS Hiring Program assistance, I observe 271 unique LEAs in the 2016 LEMAS survey wave, 352 unique LEAs in the 2007 survey wave, and 154 in both survey waves.

To observe discipline outcomes at the school and district level, I use data from the Civil Rights Data Collection (CRDC) from the U.S. Department of Education. This national-level dataset includes school-level information on discipline rates for in-school and out-of-school suspensions, expulsions, referrals to law enforcement, and arrests on school grounds. The CRDC is collected for all public schools in the U.S. in every other school year, and in this project I utilize data from the 2011-12, 2013-14, 2015-16, and 2017-18 data collections. In my analytic sample, I observe only traditional public schools and drop charters, magnets, or other non-traditional public schools.

The CRDC contains enrollment information and demographics at the school level, as well as faculty and supporting staff information. Beginning in 2013, the CRDC asked respondents to indicate whether there were any law enforcement officers present on campus during the school year. This dataset is one of the first to collect this comprehensive school-level information on SRO staffing. The data also include information on staffing of other school security guards, guidance counselors, school psychologists, librarians, and teachers. To supplement this data, I draw more detailed school and district-level characteristics and enrollment demographics from the NCES Common

[6] In the case where law enforcement agencies reside within multiple school districts (such as separate elementary and high school districts), I link them to all school districts in which they reside and observe them separately.

[7] From the data provided from the COPS office, I drop 63 police agencies over the 7 years of data that are unable to be linked by Agency Name or ORI to any other datasets to provide geographic information. Of the applicants to the grant program, 51 were fielded from school district police departments (i.e. Detroit Public Schools Police Department) that I hand-link directly to their partner school district for both geographic information and outcome data. I drop 11 applications from college or university police departments as the SROs funded by these districts would directly serve their college campuses and would not be placed in local public schools.

Core of Data (CCD) files[8].

1.5 Empirical Strategy

1.5.1 First Stage Analysis

In my primary specification, I study the effect of receipt of a COPS Hiring Program grant on student discipline outcomes to understand the causal impact of school policing. For these estimates to be credibly driven by an increased police presence in local schools, I estimate a first-stage equation to provide evidence that receipt of a COPS grant increased police presence in nearby schools. I use data from the Law Enforcement Management and Administrative Statistics (LEMAS) 2007 and 2016 survey waves on the number of SROs employed at the police agency level, as well as Civil Rights Data Collection (CRDC) data on the number of SROs staffed at the school level. I use these datasets to separately estimate two-way fixed effects models to analyze the impact of grant receipt on SRO staffing.

Within the LEMAS data, which is collected at the police agency level, I define a police agency as treated if they are ever funded for a school-based policing grant within the study period, and define the control group as police agencies who apply and never receive funding. Because the LEMAS data is collected in 2007 and 2016, I only include agencies who applied for school-based policing grants prior to 2016.[9] My primary outcome variables are the number of SROs employed per 1,000 students in the nearest school district[10], a binary indicator that takes the value of 1 if the police agency has any officers employed as SROs, and a variable for the percentage of police force within the agency that is classified as an SRO. I estimate the following equation:

$$y_{it} = \alpha + \beta CHP_{it} + \gamma ORI + \lambda Year + \epsilon_{it} \quad (1)$$

Where y_{it} represents the outcome of interest for law enforcement agency *i* at time *t*. CHP_{it} is an interaction variable equal to 1 for agencies that are ever funded in the post-period, and 0 otherwise, and ORI represents law enforcement agency fixed effects.[11] Because the LEMAS data is a cross-sectional survey and not a balanced panel, I run specifications with the full set of survey respondents as well as specifications with a balanced panel of survey respondents who appear in both survey waves, including law enforcement agency (ORI) level fixed effects. I estimate both unweighted models as well as models which include weighting by district-level student enrollment.

In a separate model, I utilize the CRDC 2015 and 2017 waves, which are the first data years to include both a binary indicator variable for SRO presence at the school level as well as a running variable representing the number of full-time equivalent law enforcement officers on campus.[12] This dataset also includes a variable

[8] Civil Rights Data Collection (CRDC) data and NCES Common Core of Data files are downloaded and accessed through the Urban Institute Education Data Portal.
[9] In the first-stage estimation using LEMAS survey data, I exclude applicants for the 2017 grant application cycle as the LEMAS data precedes their application.
[10] For LEAs that connect to multiple school districts (such as separate elementary and high school districts), I calculate the enrollment sum in all partner school districts to calculate the Number of SROs per 1000 students across all districts.
[11] ORI stands for Originating Agency Identifier, a unique identifier used by Law Enforcement Agencies.
[12] The 2015 CRDC experienced some data anomalies in the collection of the running variable for full-time law enforcement on campus. According to their documentation, the question was skipped over for about 69,000 participants. To circumvent this issue for my sample, if the

representing the number of full-time equivalent security guards on campus, which I use to estimate how funding for school police can increase total surveillance on school grounds. My primary outcomes of interest at the school level include the number of SROs per 1,000 students, a binary indicator for SRO presence, as well as the number of total security personnel (including SROs and security guards) per 1,000 students. In this specification, I define treatment at the school district level, and characterize all schools within a district as treated if any law enforcement agency within their attendance boundary lines has received a school-based policing grant within the study period.[13] I estimate the following equation:

$$y_{it} = \alpha + \beta CHP_{it} + \gamma District + \lambda Year + \epsilon_{it} \quad (2)$$

Where y_{it} represents the outcome of interest for school *i* at time *t,* and CHP_{it} is equal to 1 for all schools within a district that contains a police agency that is ever funded in the post period, and 0 otherwise. In all specifications, standard errors are clustered at the school district level, and in certain specifications I include weighting by school-level enrollment. Further, I estimate the model for schools that reported a lack of SRO presence in the pre-period to evaluate the heterogeneous impact of grant receipt for schools who previously report having no law enforcement presence. For both models outlined above, the key identifying assumption is that in the absence of treatment (grant receipt), trends in SRO staffing for treated districts would evolve similarly to their pre-treatment trends, and in parallel to trends in the control districts.

1.5.2 Causal Analysis

To estimate the effects of receipt of a COPS Hiring Program grant on student discipline, I analyze treatment data from the 2014-2017 award cycles for the COPS Hiring Program and outcome data from the 2011, 2013, 2015, and 2017 waves of the CRDC. I define treatment at the school district level, and define every school within a district as treated if there is a law enforcement agency within their district geographic boundary lines that applied for and received a COPS Hiring Grant for school-based policing. I use all schools within later-treated school districts and never-treated school districts (that applied for COPS Hiring grants and did not receive any award) as a comparison group. The primary outcomes of interest are percentages of students experiencing various types of exclusionary discipline. From counts variables provided in the CRDC, I calculate percent of students within a school who receive at least one in-school suspension, any out-of-school suspensions, expulsions, or any arrests or referrals to law enforcement. I exploit variation in treatment timing at the school district level in a difference-in-differences approach following Callaway and Sant'Anna (2021) to estimate the causal impact of receipt of a COPS hiring program school-based policing grant on student discipline outcomes. This method calculates 2x2 difference-in-differences estimations for each combination of periods *t* and treatment timing groups *g* such that it does not use already treated units as controls. The resulting parameters can be combined to provide an average treatment effect

---

school reported '0' officers on campus in the collection of the binary indicator and is missing a value for full-time equivalent officers, I impute that they have '0' full time equivalent officers on campus. Otherwise, I leave the number of full-time officers as 'missing'.

[13] Because the CRDC data is collected only in 2015 and 2017, in this specification I only study the first-stage effect for grant applicants to the 2016 and 2017 award cycles.

across all periods and timing groups, as follows:

$$ATT^{simple} = \frac{\sum_g^G \sum_t^T \mathbb{1}[g \leq t] P(G=g) ATT_{gt}}{\sum_g^G \sum_t^T \mathbb{1}[g \leq t] P(G=g)} \tag{3}$$

The key identifying assumption underlying this approach is that the precise timing of school-based policing grant receipt is uncorrelated with trends in student discipline outcomes, and that in the absence of treatment (grant receipt), trends in student discipline rates in treated school districts would evolve similarly to their pre-treatment trends, and in parallel to trends in the control districts. In my main specification, I weight observations by student enrollment at the school level and standard errors are clustered at the school district level.[14] Further, I include robustness checks in the Appendix where I estimate a stacked difference-in-differences design following Deshpande and Li (2019) utilizing the same treatment and outcome data, and find largely similar patterns of results[15].

1.6 Results

1.6.1 First Stage

To provide evidence that receipt of a school-based policing grant increases the presence of SROs in nearby schools, Tables 1.7 and 1.8 present difference-in-differences estimates of the effect of grant receipt on SRO presence. Table 1.7 displays the estimates utilizing the Law Enforcement Management and Administrative Statistics data, while Table 1.8 displays estimates using SRO staffing data from the Civil Rights Data Collection.

I find evidence that receipt of a COPS Hiring grant increases police presence in nearby schools. In Table 1.7, I provide evidence at the police-agency level using agency employment data that grant receipt increases the chance of a police agency employing any SROs by 18% relative to the mean, and increases the percentage of the total police force that serves as an SRO in their primary assignment by 25% relative to the mean. I also find evidence that the number of SROs employed per 1,000 students in the nearest school district increases, although this result is less precise. In a second specification, I find further evidence that receipt of a COPS hiring grant increases SRO presence in nearby schools from school-level data provided by the CRDC. While my estimates are not very precise, this specification provides suggestive evidence that police and total security presence increases in local schools after grant receipt. Given these results, we can credibly conclude that receipt of a COPS grant for school-based policing increases the level of police presence in nearby schools as compared to those schools within districts who did not receive the grant. Further, receipt of a COPS grant for school policing can increase the total surveillance of students in nearby schools, including both school police presence and security guard presence.

---

[14] I include robustness checks using unweighted estimates and outcomes measured at the district level in the Appendix.

[15] In this alternative design, I estimate the following equation: $Y_{ist} = \beta SBP_{st} + \gamma_{is} + \gamma_{it} + \epsilon_{ist}$, where Y_{ist} measures student outcomes in school i, year t, and stack s. SBP_{st} is an indicator variable equaling one for schools in a school district after they receive a COPS grant for school-based policing. γ_{is} and γ_{it} are school and year fixed effects, respectively. Standard errors are clustered at the school-district level, and estimates are weighted by student enrollment. I stack observations by treatment year, resulting in four total stacks.

1.6.2 Causal Analysis

Over the last decade, use of exclusionary discipline has been trending downward over time as a result of a nationwide push to reduce the use of these sanctions given recent research that has outlined the negative long-term impacts of exclusionary discipline. In Figure 1.6, I show that this pattern is true for my subsample of school districts who apply for grant funds for school police. However, while rates of exclusionary discipline have decreased over time on average, racial disparities in exposure to exclusionary discipline continue to persist. In this section, I present results from my primary specification which follows the difference-in-differences approach proposed by Callaway and Sant'Anna (2021) to estimate the causal effect of receipt of a COPS Hiring Program grant for school-based policing on student discipline outcomes in nearby schools. The primary outcomes of interest in the current study include annual school-level outcomes which calculate: the percentage of students who receive at least one in-school suspension, the percentage of students who receive one out of school suspension, multiple out of school suspensions, and a combined measure of out of school suspension, percentage of students who are expelled from school, and percentage of students who are arrested or referred to law enforcement.

As outlined in Table 1.9, I find that discipline rates remain unchanged for the average student at a given school within a treated district after receipt of a COPS Hiring grant. However, this average effect masks important underlying heterogeneity by student race, school level, and district racial composition. On average, I find that suspension rates for Hispanic students increase between 0.3-0.5 percentage points with more exposure to school policing, and that these effects are particularly prominent for Hispanic high school students, as shown in Table 1.10. Further, I find that Black high school students experience a 2-3 percentage point reduction in suspension rates following receipt of a grant for school-based policing. In Figures 1.7 and 1.8, I estimate event study specifications for the overall average as well as for Hispanic students to illustrate that these findings are not driven by differences in trends of discipline rates prior to grant receipt.

Motivated by prior research which outlines the potential for differing implementation of school policing programs by district racial composition, as well as my findings in the current study which highlight significant differences in application sentiment and key themes surrounding policing, I investigate heterogeneous treatment effects by district racial composition in Tables 1.11 and 1.12. I define district-level racial composition in the pre-period, and this measure is held constant throughout the study period. Further, I define a school district as a Majority White District if more than 50% of students enrolled are white, and a Minority White District otherwise. I follow the same methodology for identifying Majority Black and Majority Hispanic districts.

I find that in Majority Hispanic districts, all students (including both Hispanic and Black students) experience significant increases in likelihood of suspension with increased police presence after grant receipt. These effects are larger in magnitude for high school students in Majority Hispanic districts, who experience a 1.3 percentage point, or 31% increase in the likelihood of experiencing multiple suspensions in a school year relative to the mean, and a 0.6 percentage point, or 9.5% increase in the likelihood of experiencing a single suspension relative to the mean. For students in Majority Black districts, I find that while all students are less likely to experience multiple suspensions, this could be due to a shift in the severity of punishment. All students in predominantly black high schools are less likely to be suspended multiple times, but 1.2 percentage points more likely to be referred to law enforcement or arrested throughout the school year, or an 85% increase relative to the mean. In Majority White schools, students are more likely to be subjected to an in-school suspension following grant receipt. In Table 1.13, I present evidence that there is no change in district-level racial composition of students as a result of grant receipt, indicating that these heterogeneous effects are not driven by student sorting.

This pattern of results points to a potentially unintended consequence of police presence in schools on student outcomes. In school districts with a majority of students who are Hispanic or Black, agencies are more likely to list more negative sentiment in grant applications and to point to more extreme themes surrounding the need for police funding, such as ‘gang’, ‘violence’, and ‘fights’. I find that in these school districts, all students experience a shift in the threshold for punishment severity following grant receipt, but that these patterns differ slightly across district composition. In Hispanic districts, I find that this negative sentiment found in the grant applications and themes surrounding ‘policing’ the students can trickle down to worse outcomes for all students within the district in the form of increased suspension rates. While estimates are imprecise, I also find evidence that all students in these districts are more likely to be arrested or referred to law enforcement. On the other hand, in Black school districts I find that dominant application themes surrounding specific within-school disturbances such as ‘fights’ and ‘theft’ in these districts could result in reductions in the share of students who are subject to multiple out-of-school suspensions, but that this can come at the cost of increases in formal interactions with law enforcement, particularly for high school students, who are significantly more likely to be arrested or referred to law enforcement. In predominantly White school districts, while application sentiment is more positive and key themes relate to community relationship building, increased police presence can have the unintended consequence of increasing the share of students who experience the lowest level of exclusionary discipline.

In addition to differences in application themes surrounding policing, the Majority Hispanic and Majority Black districts in my sample differ in other meaningful ways. In Figure 1.9, I plot the geographic location of these districts from my sample of grant applicants. I find that Majority Black districts are more likely to be in urban centers and are concentrated in the eastern U.S., while Majority Hispanic districts are concentrated in the western U.S., particularly in California and Texas. In Table 1.14, I show that Majority Black districts are more urban, have a greater share of students on free or reduced price meals, have a higher baseline percentage of student suspensions, and have a larger per-student police presence in schools on average. I find that these baseline differences in district level characteristics, as well as district level sentiment surrounding policing, can result in meaningful differences in downstream outcomes for students.

1.6.3 Robustness to Alternative Specifications

Outlined in the Appendix, I run various alternative specifications to confirm that my results are robust to a number of different choices made in the primary analysis. I find that my results are largely robust to unweighted estimations, and to running my analysis at the district-level. I employ a secondary identification strategy, following the stacked difference-in-differences design proposed by Deshpande and Li (2019), and find largely similar patterns of results to my main specification, although some of these results are attenuated in this alternative design.

While I find that my pattern of results are robust to alternative specifications within my own research design, I find meaningfully different patterns of results than have been uncovered in the prior literature. Most notably, my analysis finds a reduction in suspension rates for Black high school students after grant receipt, in contrast to what has been shown in closely related studies. In another study of the COPS Hiring Program, Sorensen et al. (2023) find that receipt of a grant for school-based policing increases suspension rates for Black students in nearby schools. In their paper, the authors use a two-stage least squares regression discontinuity design to study the effect of receipt of a COPS grant on SRO presence in schools. In a first-stage estimation, the authors find that being just above the cut score for application acceptance significantly increases the number of SROs in local schools. They find that this increase in SRO presence translates to reductions in lower level disruptions or offenses on school grounds, but that this comes at the cost of increases in suspension, expulsion, and arrest rates for students, particularly for Black students.

There are several notable differences in the methodological approach between our two studies that could explain these differences in downstream outcomes. First, Sorensen and her coauthors leverage a regression discontinuity design, which identifies the treatment effect for a different margin than the difference-in-differences approach used in the current study. There are also several differences in our sampling approach, given that the other authors use grant applicants from fewer years of the program and limit their analysis to applications within 20 points of the cutoff for funding. Additionally, the authors analyze the treatment effect for a different measure of student discipline than used in the current study[16], and rely on the Civil Rights Data Collection variable for School Resource Officer presence as their treatment variable in the post-period. While I find suggestive evidence of an increase in police presence using this data, I find that my first-stage estimation of effects on this variable is imprecise, which could suggest that any disparities between effects could be driven by baseline differences in SRO presence in the pre-period. Finally, the authors are not able to investigate heterogeneous effects by district-level student racial composition, and I find that my pattern of results is likely driven by differences in treatment effects by district-level racial composition. While our overall findings for suspension rates for Black high school students are different, the interpretation of our results remains largely similar, in that increased police presence shifts the threshold for punishment severity for students. The observation that similarly related studies find meaningfully different effects when studying a different margin provides further evidence that implementation of school policing programs should consider the range of possibilities for heterogeneity of downstream effects.

1.7 Discussion & Conclusions

In this study, I investigate the effect of school-based policing on student discipline outcomes using evidence from a federal grant program which provided funding for local law enforcement agencies to place police officers in schools. First, I investigate important community-level sentiment and themes surrounding policing in application

[16] In their paper, the authors measure student discipline outcomes as counts per 100 students transformed by the inverse hyperbolic sine.

text data provided by the COPS Office. In this analysis, using natural language processing tools, I show that while I find significant differences in sentiment by district racial composition, this does not affect selection into funding. I show that districts with a student enrollment that is predominantly Hispanic tend to focus on themes surrounding ‘policing’ the student-body, while Black districts tended to highlight themes surrounding ‘incidents’ on school grounds. By contrast, predominantly White school districts were more likely to reference themes surrounding ‘relationship-building’. In my causal analyses, I show that these district-level differences in sentiment and key themes correspond to heterogeneous treatment effects.

While I find no evidence of increases in disciplinary rates on average for all students within a district which receives this targeted SRO funding, these average effects mask important underlying heterogeneity by student, school, and district characteristics. On average, suspension rates for Hispanic students increase between 10-15% relative to the mean following grant receipt, while suspension rates for Black high school students decrease. In school districts with mostly Hispanic students, all students experience 0.5-0.6 percentage point increases in suspension rates following grant receipt. However, in school districts with mostly Black students, all students experience reductions in the probability of having multiple suspensions, and high school students are 85% more likely to be arrested or referred to law enforcement relative to the mean. This pattern of results points to a potentially unintended consequence of police presence in schools on student outcomes, wherein police presence shifts the threshold for usage of more extreme forms of punishment in schools. Although students in predominantly Black school districts benefit from reductions in suspension rates, this could come at the cost of more formal interactions with law enforcement in the form of arrests or referrals.

Prior literature has shown that while school police can increase school safety in the form of reductions in lower-level disruptions or offenses, this could come at the cost of increases in formal discipline, particularly for minority students (Sorensen, Lucy C. et al. 2023; E. G. Owens 2017). Research has shown that school police view their primary roles and responsibilities differently based on the demographic composition of the student body they serve (Fisher et al. 2022), and that minority students in schools with greater minority enrollment face a double jeopardy when receiving disciplinary sanctions from administration (J. Owens 2022). In the current study, I find evidence that this school and district-level sentiment surrounding policing and student composition can have heterogeneous implications for the downstream effects of implementation of a school policing program. Taken together, this body of literature coupled with the results presented in this paper provide evidence that school policing programs are not a one-size-fits-all policy approach.

1.8 References

[1] Aizer, Anna, and Joseph J. Doyle. 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics* 130 (2): 759–803. https://doi.org/10.1093/qje/qjv003.

[2] Anderson, Kenneth Alonzo. 2018. "Policing and Middle School: An Evaluation of a Statewide School Resource Officer Policy." *Middle Grades Review* 4 (2).

[3] Ang, Desmond. 2020. "The Effects of Police Violence on Inner-City Students." *The Quarterly Journal of Economics* 136 (1): 115–68. https://doi.org/10.1093/qje/qjaa027.

[4] Callaway, Brantly, and Pedro H.C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

[5] Counts, Jennifer, Kristina N. Randall, Joseph B. Ryan, and Antonis Katsiyannis. 2018. "School Resource Officers in Public Schools: A National Review." *Education and Treatment of Children* 41 (4): 405–30. https://doi.org/10.1353/etc.2018.0023.

[6] Curran, F. Chris. 2016. "Estimating the Effect of State Zero Tolerance Laws on Exclusionary Discipline, Racial Discipline Gaps, and Student Behavior." *Educational Evaluation and Policy Analysis* 38 (4): 647–68. https://doi.org/10.3102/0162373716652728.

[7] Curran, F. Chris. 2019. "The Law, Policy, and Portrayal of Zero Tolerance School Discipline: Examining Prevalence and Characteristics Across Levels of Governance and School Districts." *Educational Policy* 33 (2): 319–49. https://doi.org/10.1177/0895904817691840.

[8] Curran, F. Chris, Benjamin W. Fisher, Samantha Viano, and Aaron Kupchik. 2019. "Why and When Do School Resource Officers Engage in School Discipline? The Role of Context in Shaping Disciplinary Involvement." *American Journal of Education* 126 (1): 33–63. https://doi.org/10.1086/705499.

[9] Curran, F. Chris, Samantha Viano, Aaron Kupchik, and Benjamin W. Fisher. 2021. "Do Interactions With School Resource Officers Predict Students' Likelihood of Being Disciplined and Feelings of Safety? Mixed-Methods Evidence From Two School Districts." *Educational Evaluation and Policy Analysis* 43 (2): 200–232. https://doi.org/10.3102/0162373720985904.

[10] Davis, Elizabeth J. 2022. "Law Enforcement Agencies That Employ School Resource Officers, 2019." *U.S. Department of Justice*, November. https://bjs.ojp.gov/library/publications/law-enforcement-agencies-employ-school-resource-officers-2019.

[11] Deshpande, Manasi, and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy* 11 (4): 213–48. https://doi.org/10.1257/pol.20180076.

[12] Eden, Max. "School Discipline Reform and Disorder: Evidence from New York City Public Schools, 2012-16."

Manhattan Institute.

[13] Evans, William N., and Emily G. Owens. 2007. "COPS and Crime." *Journal of Public Economics* 91 (1–2): 181–201. https://doi.org/10.1016/j.jpubeco.2006.05.014.

[14] Finn, Peter, Jack McDevitt, William Lassiter, Michael Shively, and Tom Rich. 2005. "Case Studies of 19 School Resource Officer (SRO) Programs." https://doi.org/10.1037/e510372006-001.

[15] Fisher, Benjamin W., and Emily A. Hennessy. 2016. "School Resource Officers and Exclusionary Discipline in U.S. High Schools: A Systematic Review and Meta-Analysis." *Adolescent Research Review* 1 (3): 217–33. https://doi.org/10.1007/s40894-015-0006-8.

[16] Fisher, Benjamin W, Ethan M Higgins, Aaron Kupchik, Samantha Viano, F Chris Curran, Suzanne Overstreet, Bryant Plumlee, and Brandon Coffey. 2022. "Protecting the Flock or Policing the Sheep? Differences in School Resource Officers' Perceptions of Threats by School Racial Composition." *Social Problems* 69 (2): 316–34. https://doi.org/10.1093/socpro/spaa062.

[17] Geller, Amanda, Jeffrey Fagan, Tom Tyler, and Bruce G. Link. 2014. "Aggressive Policing and the Mental Health of Young Urban Men." *American Journal of Public Health* 104 (12): 2321–27. https://doi.org/10.2105/AJPH.2014.302046.

[18] Gelman, Andrew, Jeffrey Fagan, and Alex Kiss. 2007. "An Analysis of the New York City Police Department's 'Stop-and-Frisk' Policy in the Context of Claims of Racial Bias." *Journal of the American Statistical Association* 102 (479): 813–23. https://doi.org/10.1198/016214506000001040.

[19] Goel, Sharad, Justin M. Rao, and Ravi Shroff. 2016. "Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy." *The Annals of Applied Statistics* 10 (1). https://doi.org/10.1214/15-AOAS897.

[20] Gottfredson, Denise C., Scott Crosse, Zhiqun Tang, Erin L. Bauer, Michele A. Harmon, Carol A. Hagen, and Angela D. Greene. 2020. "Effects of School Resource Officers on School Crime and Responses to School Crime." *Criminology & Public Policy* 19 (3): 905–40. https://doi.org/10.1111/1745-9133.12512.

[21] Harvey, Anna, and Taylor Mattia. 2022. "Reducing Racial Disparities in Crime Victimization: Evidence from Employment Discrimination Litigation." *Journal of Urban Economics*, May, 103459. https://doi.org/10.1016/j.jue.2022.103459.

[22] Hjalmarsson, Randi. 2008. "Criminal Justice Involvement and High School Completion." *Journal of Urban Economics* 63 (2): 613–30. https://doi.org/10.1016/j.jue.2007.04.003.

[23] Holt, Stephen, Katie Vinopal, Heasun Choi, and Lucy Sorensen. 2022. "Strictly Speaking: Examining Teacher Use of Punishment and Student Outcomes." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4093118.

[24] Jackson, Arrick. 2002. "Police-school Resource Officers' and Students' Perception of the Police and Offending." *Policing: An International Journal of Police Strategies & Management* 25 (3): 631–50.

https://doi.org/10.1108/13639510210437078.

[25] James, Nathan, and Gail McCallion. 2013. "School Resource Officers: Law Enforcement Officers in Schools." *Congressional Research Service*.

[26] Kupchik, Aaron. 2010. *Homeroom Security: School Discipline in an Age of Fear*. New York: New York University Press.

[27] Kupchik, Aaron, F. Chris Curran, Benjamin W. Fisher, and Samantha L. Viano. 2020. "Police Ambassadors: Student-Police Interactions in School and Legal Socialization." *Law & Society Review* 54 (2): 391–422. https://doi.org/10.1111/lasr.12472.

[28] Lacoe, Johanna, and Matthew P. Steinberg. 2019. "Do Suspensions Affect Student Outcomes?" *Educational Evaluation and Policy Analysis* 41 (1): 34–62. https://doi.org/10.3102/0162373718794897.

[29] Levitt, Steven D. 2002. "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply." *American Economic Review* 92 (4): 1244–50. https://doi.org/10.1257/00028280260344777.

[30] Lovenheim, Michael F., and Emily G. Owens. 2014. "Does Federal Financial Aid Affect College Enrollment? Evidence from Drug Offenders and the Higher Education Act of 1998." *Journal of Urban Economics* 81 (May): 1–13. https://doi.org/10.1016/j.jue.2014.01.005.

[31] Lynch, Caitlin G., Randy R. Gainey, and Allison T. Chappell. 2016. "The Effects of Social and Educational Disadvantage on the Roles and Functions of School Resource Officers." *Policing: An International Journal of Police Strategies & Management* 39 (3): 521–35. https://doi.org/10.1108/PIJPSM-02-2016-0021.

[32] Madfis, Eric. 2016. "'It's Better to Overreact': School Officials' Fear and Perceived Risk of Rampage Attacks and the Criminalization of American Public Schools." *Critical Criminology* 24 (1): 39–55. https://doi.org/10.1007/s10612-015-9297-0.

[33] Na, Chongmin, and Denise C. Gottfredson. 2013. "Police Officers in Schools: Effects on School Crime and the Processing of Offending Behaviors." *Justice Quarterly* 30 (4): 619–50. https://doi.org/10.1080/07418825.2011.615754.

[34] "National Association of School Resource Officers." Accessed March 31, 2023. https://www.nasro.org/.

[35] "National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education." Accessed March 31, 2023. https://nces.ed.gov/.

[36] Owens, Emily, and Bocar Ba. 2021. "The Economics of Policing and Public Safety." *Journal of Economic Perspectives* 35 (4): 3–28. https://doi.org/10.1257/jep.35.4.3.

[37] Owens, Emily G. 2017. "Testing the School-to-Prison Pipeline: Testing the School-to-Prison Pipeline." *Journal of Policy Analysis and Management* 36 (1): 11–37. https://doi.org/10.1002/pam.21954.

[38] Owens, Jayanti. 2022. "Double Jeopardy: Teacher Biases, Racialized Organizations, and the Production of Racial/Ethnic Disparities in School Discipline." *American Sociological Review* 87 (6): 1007–48.

https://doi.org/10.1177/00031224221135810.

[39] Peterson, Jillian, James Densley, and Gina Erickson. 2021. "Presence of Armed School Officials and Fatal and Nonfatal Gunshot Injuries During Mass School Shootings, United States, 1980-2019." *JAMA Network Open* 4 (2): e2037394. https://doi.org/10.1001/jamanetworkopen.2020.37394.

[40] Rumberger, Russell W., and Daniel J. Losen. 2016. "The High Cost of Harsh Discipline and Its Disparate Impact." June 2, 2016. https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/the-high-cost-of-harsh-discipline-and-its-disparate-impact/UCLA_HighCost_6-2_948.pdf.

[41] Sorensen, Lucy C., Montserrat Avila-Acosta, John B. Engberg, and Shawn D. Bushway. 2023. "The Thin Blue Line in Schools: New Evidence on School-based Policing across the U.S." *Journal of Policy Analysis and Management* 42 (4): 941–70. https://doi.org/10.1002/pam.22498.

[42] Sorensen, Lucy C., Shawn D. Bushway, and Elizabeth J. Gifford. 2022. "Getting Tough? The Effects of Discretionary Principal Discipline on Student Outcomes." *Education Finance and Policy* 17 (2): 255–84. https://doi.org/10.1162/edfp_a_00341.

[43] Sorensen, Lucy C, Yinzhi Shen, and Shawn D Bushway. 2021. "Making Schools Safer and/or Escalating Disciplinary Response: A Study of Police Officers in North Carolina Schools." *Educational Evaluation and Policy Analysis* 43 (3): 495–519. https://doi.org/10.3102/01623737211006409.

[44] Theriot, Matthew T. 2016. "The Impact of School Resource Officer Interaction on Students' Feelings About School and School Police." *Crime & Delinquency* 62 (4): 446–69. https://doi.org/10.1177/0011128713503526.

[45] Theriot, Matthew T., and John G. Orme. 2016. "School Resource Officers and Students' Feelings of Safety at School." *Youth Violence and Juvenile Justice* 14 (2): 130–46. https://doi.org/10.1177/1541204014564472.

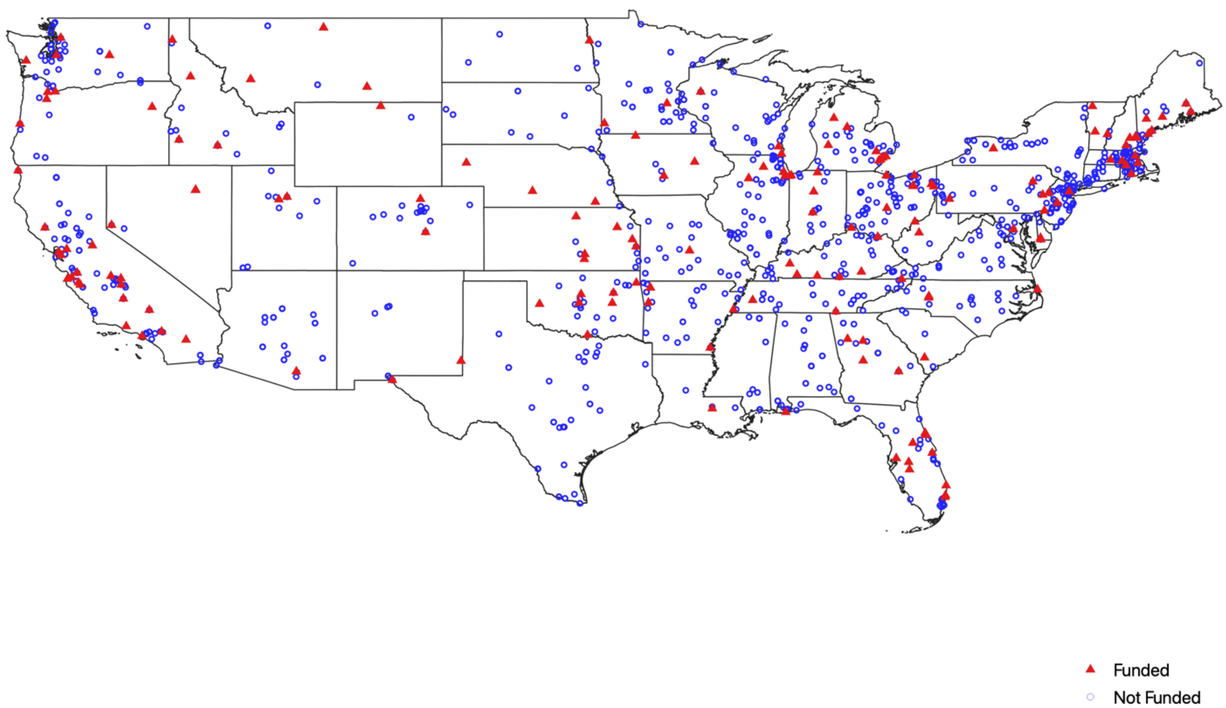
[46] Viano, Samantha, F. Chris Curran, and Benjamin W. Fisher. 2021. "Kindergarten Cop: A Case Study of How a Coalition Between School Districts and Law Enforcement Led to School Resource Officers in Elementary Schools." *Educational Evaluation and Policy Analysis* 43 (2): 253–79. https://doi.org/10.3102/0162373721989290.

[47] Weisburst, Emily K. 2019a. "Patrolling Public Schools: The Impact of Funding for School Police on Student Discipline and Long-Term Education Outcomes: Patrolling Public Schools." *Journal of Policy Analysis and Management* 38 (2): 338–65. https://doi.org/10.1002/pam.22116.

[48] Weisburst, Emily K. 2019b. "Safety in Police Numbers: Evidence of Police Effectiveness from Federal COPS Grant Applications." *American Law and Economics Review* 21 (1): 81–109. https://doi.org/10.1093/aler/ahy010.

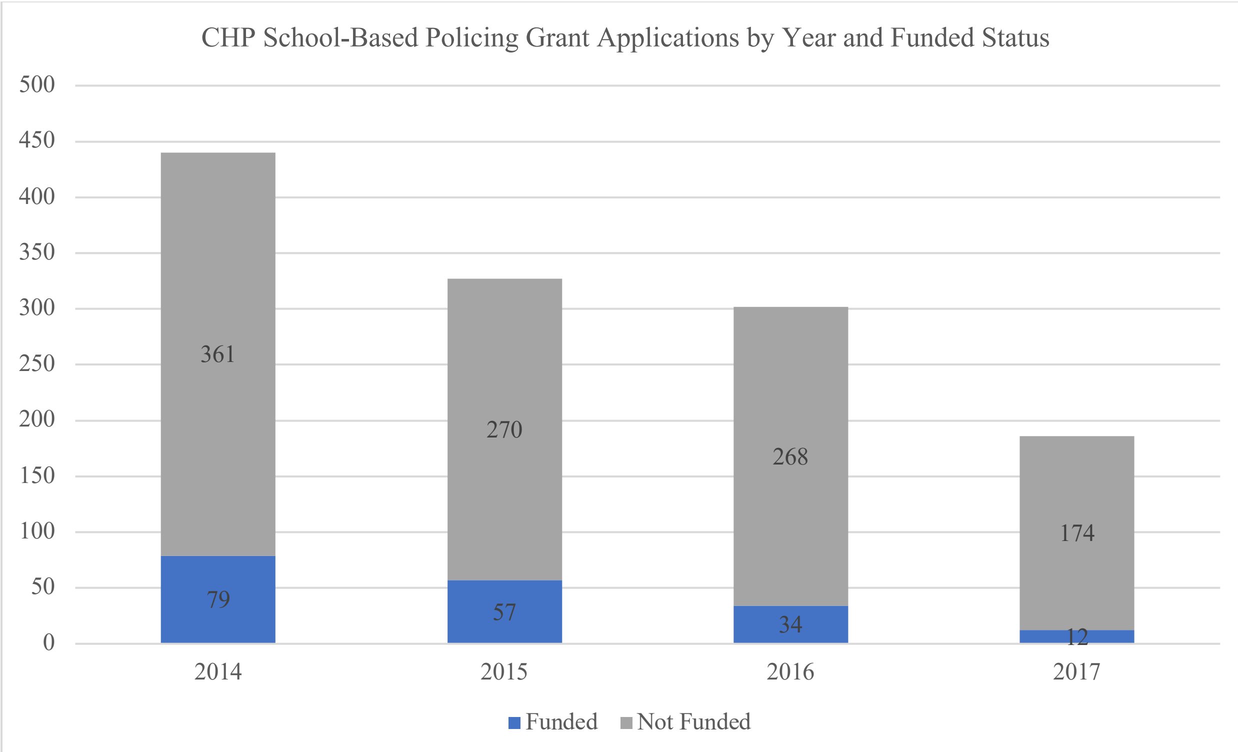
[49] Zhang, Gary. 2019. "The Effects of a School Policing Program on Crime, Discipline, and Disorder: A Quasi-Experimental Evaluation." *American Journal of Criminal Justice* 44 (1): 45–62. https://doi.org/10.1007/s12103-018-9440-z.

Figure 1.1 COPS Grant Application Map by Funded Status (2014-2017)



Note: This map plots the latitude and longitude of police agencies that applied for COPS Hiring Program grants for school-based policing in the 2014-2017 grant cycles by their award status. Red triangles represent agencies that received an award, while blue circles represent agencies that did not receive an award. Geographic data is sourced from the Bureau of Justice Statistics and the NCES Common Core of Data. Not pictured here are grant applicants/recipients in Alaska and Hawaii.

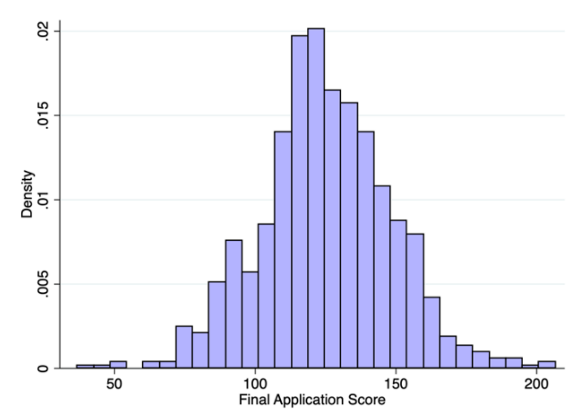
Figure 1.2 Histogram of School-District Level Treatment Timing



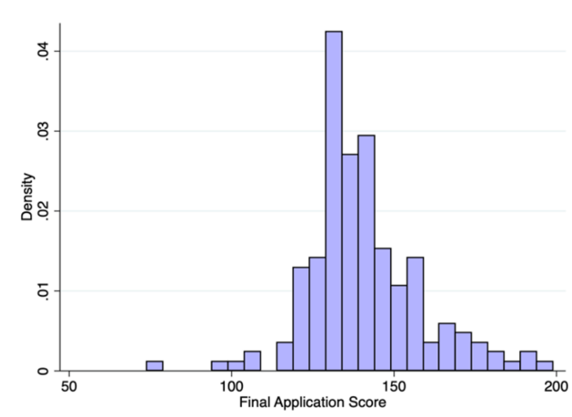
Note: Data obtained through a Freedom of Information Act request to the COPS Office at the U.S. Department of Justice is used to create this histogram of grant application timing by award status. Applying agencies are linked to the school district(s) in which they reside and treatment status is defined above at the school district level.

Figure 1.3 Histogram of Grant Application Scores by Funded Status

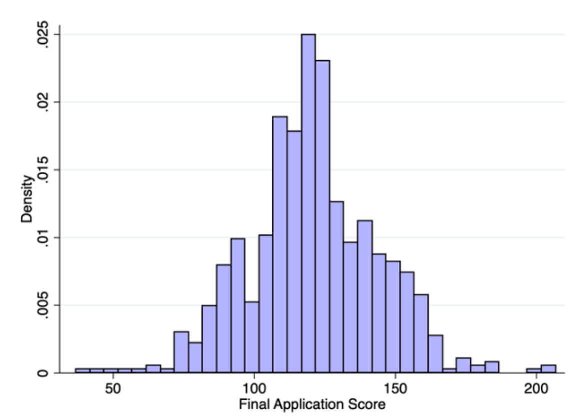
(a) All Applications



(b) Funded Applications



(c) Unfunded Applications



Data Source: FOIA request to the COPS Office of the U.S. Department of Justice

Table 1.1 School District and Police Agency Summary Statistics by Grant Funded Status

	All School Districts	**Never Funded**	**Funded**
	Mean/sd	Mean/sd	Mean/sd
Student : Guidance Counselor Ratio	485.46	489.11	467.22
	(259.01)	(268.46)	(205.32)
Student : Librarian Ratio	1535.57	1469.89	1863.46
	(2273.59)	(2025.66)	(3229.59)
Student : Teacher Ratio	15.37	15.25	16.00
	(3.40)	(3.09)	(4.62)
Number of Schools in District	25.84	21.42	47.91
	(70.56)	(57.09)	(114.27)
Observations	821	684	137
Percent White	0.62	0.64	0.54
	(0.29)	(0.28)	(0.31)
Percent Black	0.11	0.10	0.14
	(0.18)	(0.17)	(0.22)
Percent Hispanic	0.19	0.18	0.24
	(0.24)	(0.24)	(0.27)
Percent Male	0.52	0.52	0.52
	(0.01)	(0.01)	(0.01)
Percent Female	0.49	0.49	0.49
	(0.01)	(0.01)	(0.01)
Percent Special Education	0.14	0.15	0.14
	(0.07)	(0.07)	(0.04)
Percent English Language Learner	0.07	0.06	0.10
	(0.10)	(0.10)	(0.13)
Observations	1,206	991	215
	All Police Agencies	**Never Funded**	**Funded**
	Mean/sd	Mean/sd	Mean/sd
Number of SROs Employed	5.70	5.17	7.94
	(10.62)	(9.87)	(13.28)
Number of SROs Per 1000 Students	0.50	0.51	0.44
	(1.10)	(1.20)	(0.50)
Percent White FTS Officers	0.82	0.85	0.70
	(0.23)	(0.21)	(0.29)
Percent Black FTS Officers	0.06	0.05	0.09
	(0.11)	(0.09)	(0.18)
Percent Hispanic FTS Officers	0.08	0.08	0.09
	(0.16)	(0.17)	(0.11)
Percent Male FTS Officers	0.89	0.90	0.83
	(0.14)	(0.11)	(0.22)
Observations	167	135	32

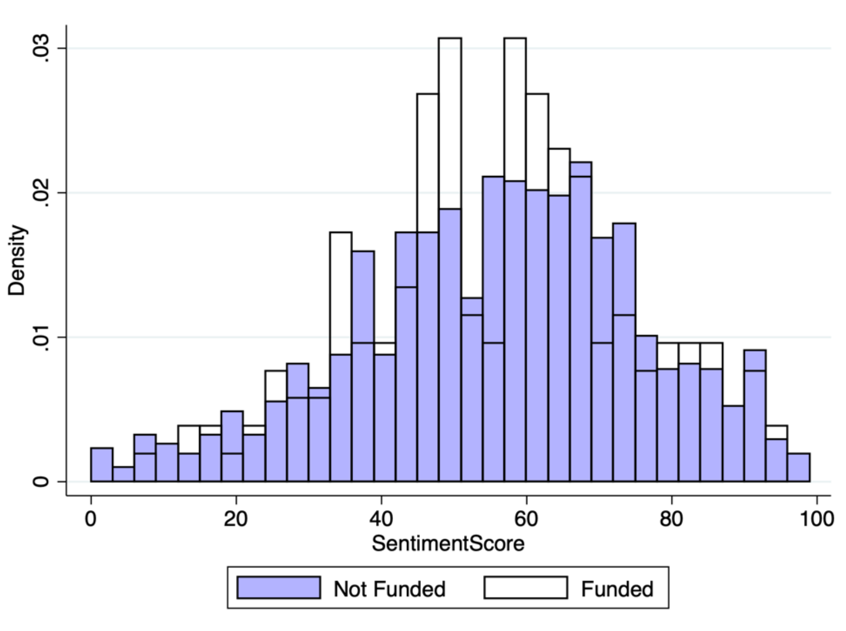
Note: School District level summary statistics data sourced from the NCES Common Core of Data 2014-2017 enrollment reports. Police Agency level summary statistics data sourced from the Bureau of Justice Statistics Law Enforcement Management and Administrative Statistics 2016 survey. "FTS" stands for Full-Time Sworn.

Table 1.2 Sentiment Analysis Descriptive Statistics by Funded Status and District Racial Composition

	Not Funded	Funded	Diff
	Mean	Mean	b
Average Positive Sentiment Percentage per Application	55.97	55.46	0.52
Proportion of Applications with Overall Positive Sentiment	0.62	0.59	0.03
Observations	1,027	174	1,201
	Minority Hispanic Districts	**Majority Hispanic Districts**	**Diff**
	Mean	Mean	b
Average Positive Sentiment Percentage per Application	56.76	49.79	6.97***
Proportion of Applications with Overall Positive Sentiment	0.64	0.44	0.19***
Observations	1,058	133	1,191
	Minority White Districts	**Majority White Districts**	**Diff**
	Mean	Mean	b
Average Positive Sentiment Percentage per Application	54.58	56.57	-1.99
Proportion of Applications with Overall Positive Sentiment	0.55	0.65	-0.09**
Observations	355	836	1,191
	Minority Black Districts	**Majority Black Districts**	**Diff**
	Mean	Mean	b
Average Positive Sentiment Percentage per Application	56.08	54.08	2.00
Proportion of Applications with Overall Positive Sentiment	0.63	0.48	0.15*
Observations	1,128	63	1,191

Note: District-level racial composition is calculated using data from the NCES Common Core of Data. A district with a subgroup racial majority contains a student body where greater than 50% of students belong to that subgroup.

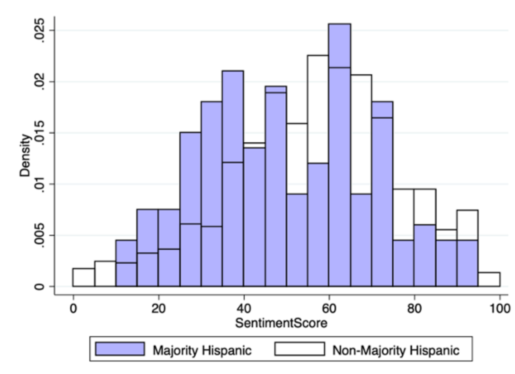
Figure 1.4 Histogram of Sentiment Score by Application Funded Status



Note: Sentiment score is calculated as a value between 0 and 100 for proportion of a single application which is considered to have positive sentiment. A score between 50-100 would be categorized as an application with ‘positive’ sentiment, while a score between 0-50 would be categorized as an application with ‘negative’ sentiment. The p-value from a Kolmogorov-Smirnov test of differences between the above distributions is 0.228.

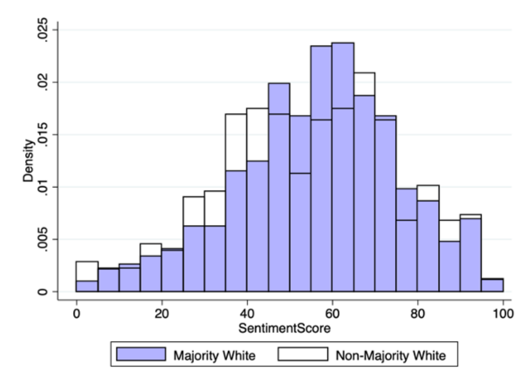
Figure 1.5 Histogram of Sentiment Score by District Racial Composition

(a) By Hispanic Composition



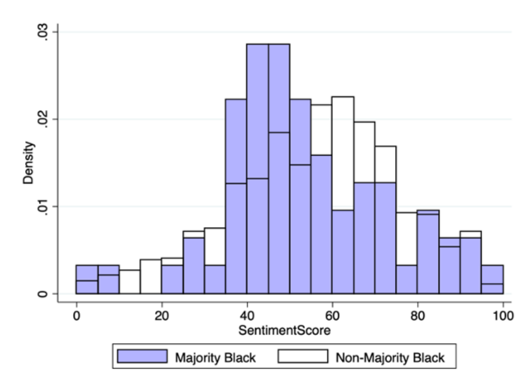
Note: The p-value from a Kolmogorov-Smirnov test of differences between the above distributions is 0.000.

(b) By White Composition



Note: The p-value from a Kolmogorov-Smirnov test of differences between the above distributions is 0.014.

(c) By Black Composition



Note: The p-value from a Kolmogorov-Smirnov test of differences between the above distributions is 0.048.

Notes: Sentiment score is calculated as a value between 0 and 100 for percent of a single application which is considered to have positive sentiment. A score between 50-100 would be categorized as an application with ‘positive’ sentiment, while a score between 0-50 would be categorized as an application with ‘negative’ sentiment. District-level racial composition is calculated using data from the NCES Common Core of Data. A district with a subgroup racial majority contains a student body where greater than 50% of students belong to that subgroup.

Table 1.3 Top Application Keywords by Relative Frequency Difference between Funded and Unfunded Applications

Keyword [A]	Frequency in Funded Applications [B]	Frequency in Unfunded Applications [C]	Difference [B]-[C]
City	0.52	0.31	0.21
Gang	0.32	0.17	0.15
Continue	0.25	0.13	0.12
Violent	0.20	0.08	0.12
Youth	0.52	0.39	0.13
Focus	0.17	0.26	-0.09
Allow	0.19	0.26	-0.07
Basis	0.09	0.15	-0.06
Full	0.14	0.19	-0.05
Serve	0.18	0.23	-0.05

Note: [B] is calculated as the percentage of Funded applications which contain at least one mention of the keyword in [A] in their text submission. [C] is calculated similarly for Unfunded applications.

Table 1.4 Top Application Keywords by Relative Frequency Difference between Majority Hispanic and Minority Hispanic Applications

Keyword [A]	Frequency in Majority Hispanic Applications [B]	Frequency in Minority Hispanic Applications [C]	Difference [B]-[C]
Gang	0.52	0.17	0.35
Members	0.37	0.19	0.18
Violence	0.49	0.32	0.17
Rate	0.25	0.09	0.16
Poverty	0.22	0.06	0.16
Help	0.21	0.37	-0.16
Service	0.13	0.28	-0.15
Position	0.10	0.23	-0.13
Issues	0.33	0.44	-0.11
Safety	0.52	0.62	-0.10

Note: [B] is calculated as the percentage of applications from Majority Hispanic districts which contain at least one mention of the keyword in [A] in their text submission. [C] is calculated similarly for applications from Minority Hispanic districts.

Table 1.5 Top Application Keywords by Relative Frequency Difference between Majority Black and Minority Black Applications

Keyword [A]	Frequency in Majority Black Applications [B]	Frequency in Minority Black Applications [C]	Difference [B]-[C]
Gang	0.44	0.19	0.25
Security	0.52	0.32	0.20
Fights	0.21	0.04	0.17
Theft	0.22	0.09	0.13
Act	0.22	0.09	0.13
Children	0.11	0.36	-0.25
Abuse	0.08	0.23	-0.15
Alcohol	0.06	0.20	-0.14
Drug	0.32	0.44	-0.12
Program	0.38	0.49	-0.11

Note: [B] is calculated as the percentage of applications from Majority Black districts which contain at least one mention of the keyword in [A] in their text submission. [C] is calculated similarly for applications from Minority Black districts.

Table 1.6 Top Application Keywords by Relative Frequency Difference between Majority White and Minority White Applications

Keyword [A]	Frequency in Majority White Applications [B]	Frequency in Minority White Applications [C]	Difference [B]-[C]
Issues	0.46	0.34	0.12
Administration	0.17	0.07	0.10
Relationship	0.21	0.12	0.09
Assist	0.31	0.22	0.09
Children	0.37	0.29	0.08
Gang	0.12	0.41	-0.29
Activity	0.22	0.31	-0.09
Residents	0.08	0.17	-0.09
Campuses	0.09	0.18	-0.09
Violence	0.32	0.39	-0.07

Note: [B] is calculated as the percentage of applications from Majority White districts which contain at least one mention of the keyword in [A] in their text submission. [C] is calculated similarly for applications from Minority White districts.

Table 1.7 COPS Hiring Program School-Based Policing Grants and SRO Employment, evidence from LEMAS Survey Data

	Number of SROs per 1000 Students			SRO Binary			SRO Rate		
SBP Grant Awarded	0.00	0.10	0.05	0.04	0.07	0.14*	0.01	0.02**	0.01**
	(0.21)	(0.09)	(0.05)	(0.11)	(0.06)	(0.08)	(0.01)	(0.01)	(0.01)
Weighted by District Enrollment		X	X		X	X		X	X
ORI Fixed Effects			X			X			X
Mean	0.40	0.40	0.40	0.75	0.75	0.75	0.04	0.04	0.04
Observations	369	369	152	373	370	152	372	369	152

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: I utilize LEMAS survey data from 2007 and 2016 to estimate a difference-in-differences specification where treatment equals 1 if a police agency was ever funded in the 2014-2016 application cycles and 0 if they were never funded within the 2014-2016 application cycles despite submitting an application. I calculate the outcome variable Number of SROs per 1000 students by connecting police agencies to their nearest school district enrollment numbers. SRO Binary is equal to 1 if the police agency lists any officers as serving as School Resource Officers and zero otherwise, and SRO Rate represents the percent of the police force at that agency serving as School Resource Officers.

Table 1.8 COPS Hiring Program School-Based Policing Grants and SRO Placement, evidence from the Civil Rights Data Collection

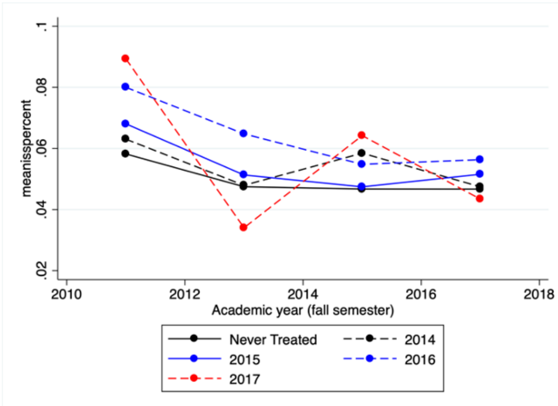
	Number of SROs per 1000 Students			SRO Binary			Total Security Personnel per 1000 Students		
SBP Grant Awarded	0.30	0.19	0.54	0.03	-0.01	0.15	0.17	0.11	0.39
	(0.39)	(0.19)	(0.39)	(0.09)	(0.08)	(0.11)	(0.41)	(0.20)	(0.40)
School District Fixed Effects	X	X	X	X	X	X	X	X	X
Weighted by School Enrollment		X	X		X	X		X	X
No SRO Presence in 2015			X			X			X
Mean	0.43	0.43	0.12	0.26	0.26	0.06	1.08	1.08	0.76
Observations	9411	9208	6899	9482	9264	6943	9397	9195	6886

Standard errors in parentheses

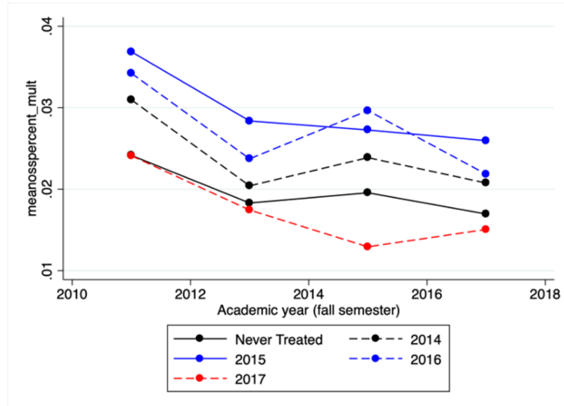
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: I utilize CRDC respondent data from 2015 and 2017 to estimate a difference-in-differences specification where treatment equals 1 if a school resides in a district that contains a police agency that was ever funded in the 2016-2017 application cycles and 0 if they were never funded within the 2016-2017 application cycles despite submitting an application. I calculate the outcome variable Number of SROs per 1000 students using CCD enrollment numbers at the school level. SRO Binary is equal to 1 if the school reports any SRO presence during the school year and 0 otherwise, and Total Security Personnel per 1000 Students represents the total number of law enforcement officers and security guards present during the school year at the school level. I drop observations from the state of Florida, as they underreport SRO presence in the CRDC.

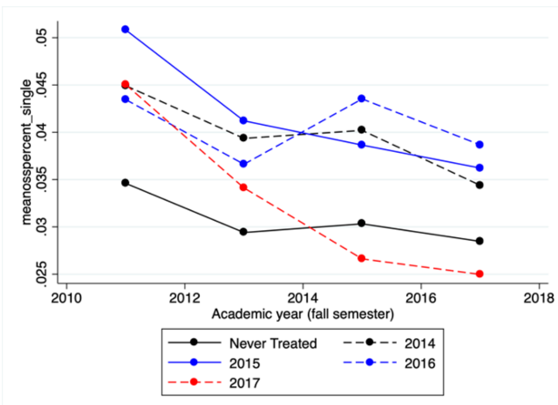
Figure 1.6 Time-Trends in Means for Student Discipline Outcomes by Grant Receipt Year



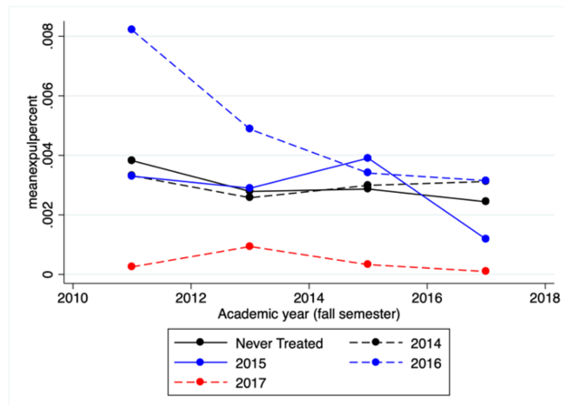
(a) In-School Susp. Rate by Treat Year



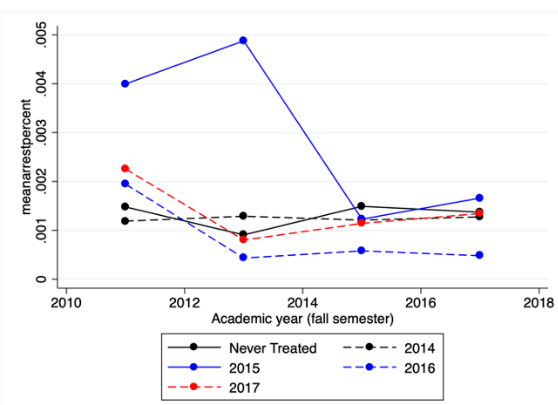
(b) Mult. Out-of-School Susp. Rate by Treat Year



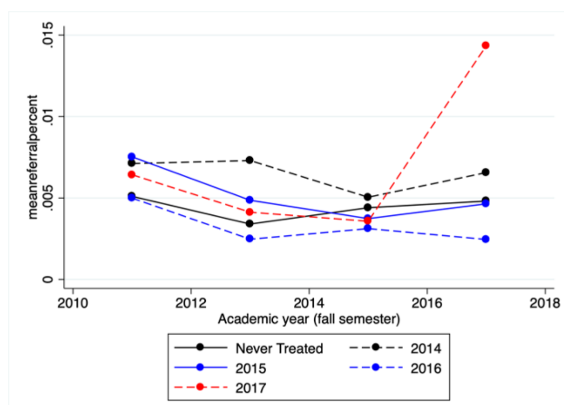
(c) Single Out-of-School Susp. Rate by Treat Year



(d) Expulsion Rate by Treat Year



(e) Arrest Rate by Treat Year



(f) Referral Rate by Treat Year

Data Source: Civil Rights Data Collection (2011-2017)

Notes: This figure graphs time trends in the primary outcome variables for grant applicants who were never treated, and by treatment year, where treatment is defined as receipt of funding for school-based policing. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest Rate or Referral Rate).

Table 1.9 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls

	(1) In-School Susp Rate	(2) Out-of-School Susp Rate	(3) Mult. Out-of-School Susp Rate	(4) Single Out-of-School Susp Rate	(5) Expulsion Rate	(6) Arrest or Referral Rate
PANEL A: ALL STUDENTS						
ATT^{simple}	-0.001	0.000	-0.001	0.001	0.000	-0.008
	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.008)
Mean	0.043	0.051	0.021	0.031	0.002	0.006
Observations	40,830	40,857	40,880	40,892	40,592	40,718
PANEL B: WHITE STUDENTS						
ATT^{simple}	0.003	0.003	0.001	0.001	0.000	-0.004
	(0.004)	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)
Mean	0.037	0.046	0.018	0.029	0.002	0.006
Observations	39,967	39,959	40,004	40,004	39,712	39,201
PANEL C: BLACK STUDENTS						
ATT^{simple}	-0.010	-0.013	-0.012	-0.001	0.000	-0.002
	(0.008)	(0.010)	(0.008)	(0.004)	(0.001)	(0.003)
Mean	0.077	0.097	0.042	0.057	0.004	0.012
Observations	37,421	37,410	37,513	37,535	37,312	36,353
PANEL D: HISPANIC STUDENTS						
ATT^{simple}	0.003	0.005**	0.003**	0.003*	0.000	-0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.000)	(0.003)
Mean	0.042	0.049	0.020	0.031	0.002	0.007
Observations	40,002	40,002	40,049	40,066	39,771	39,037

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

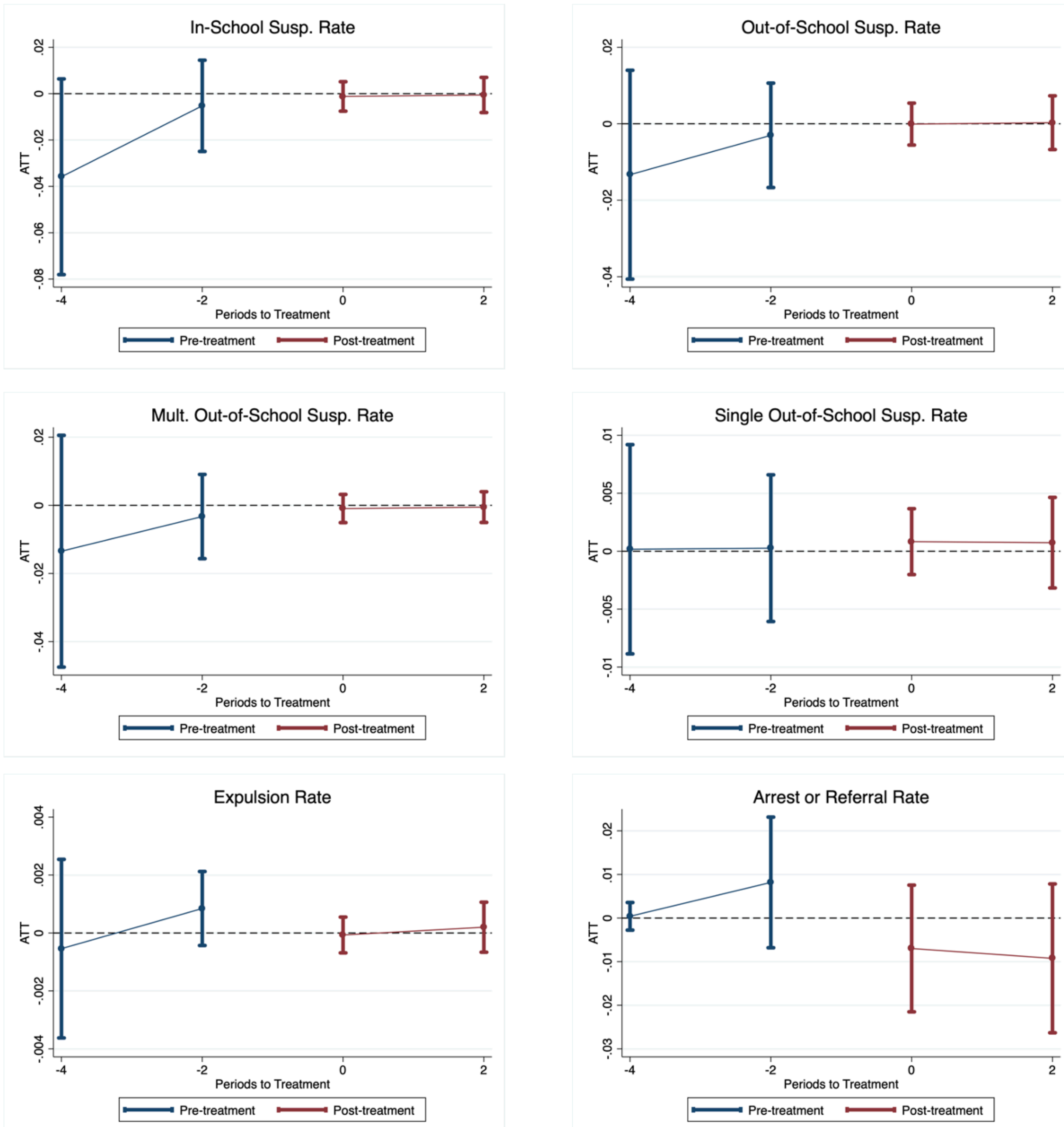
Table 1.10 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated high schools as controls

	HIGH SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
ATT^{simple}	-0.001	-0.001	-0.002	0.001	0.000	-0.009
	(0.006)	(0.005)	(0.003)	(0.003)	(0.001)	(0.010)
Mean	0.081	0.093	0.038	0.057	0.007	0.018
Observations	6,965	6,988	7,000	7,003	6,935	6,945
	PANEL B: WHITE STUDENTS					
ATT^{simple}	0.005	0.005	0.002	0.004	0.000	-0.002
	(0.006)	(0.004)	(0.002)	(0.003)	(0.001)	(0.003)
Mean	0.070	0.077	0.030	0.050	0.006	0.015
Observations	6,808	6,813	6,837	6,837	6,772	6,690
	PANEL C: BLACK STUDENTS					
ATT^{simple}	-0.020	-0.033**	-0.024*	-0.009	0.002	-0.001
	(0.015)	(0.017)	(0.013)	(0.007)	(0.002)	(0.005)
Mean	0.133	0.157	0.066	0.096	0.009	0.030
Observations	6,297	6,302	6,342	6,348	6,289	6,160
	PANEL D: HISPANIC STUDENTS					
ATT^{simple}	0.004	0.009*	0.005*	0.004	-0.001	0.000
	(0.009)	(0.005)	(0.003)	(0.003)	(0.001)	(0.006)
Mean	0.085	0.098	0.040	0.062	0.007	0.019
Observations	6,803	6,818	6,840	6,841	6,775	6,682

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

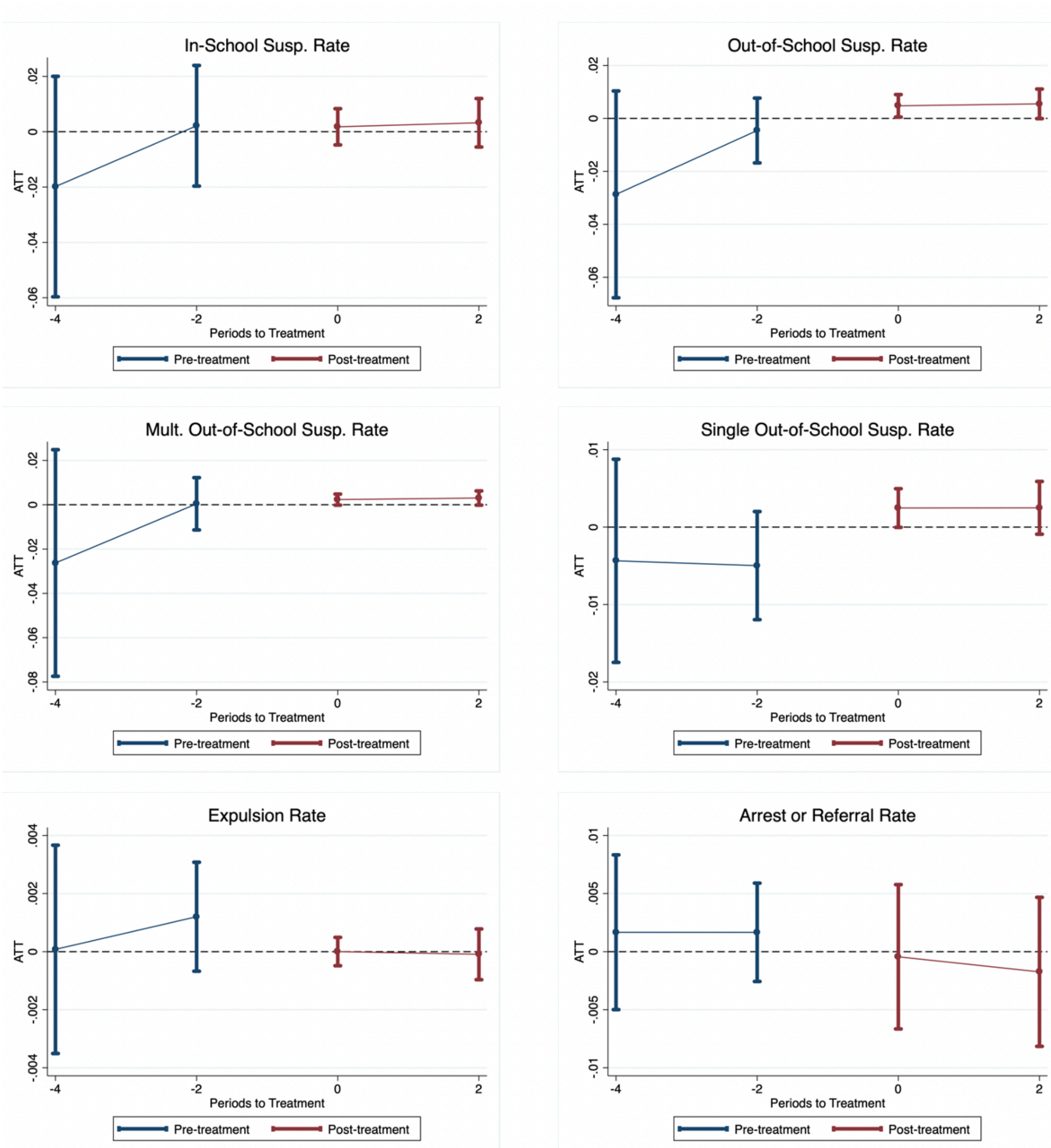
Figure 1.7 Callaway and Sant'Anna Event Study Aggregations for School-Level Discipline Outcomes



Data Source: Civil Rights Data Collection 2011, 2013, 2015, 2017

Notes: This figure plots the event study aggregations following Callaway and Sant'Anna (2021) for all students in a treated school. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Figure 1.8 Callaway and Sant'Anna Event Study Aggregations for Discipline Outcomes for Hispanic Students



Data Source: Civil Rights Data Collection 2011, 2013, 2015, 2017

Notes: This figure plots the event study aggregations following Callaway and Sant'Anna (2021) for Hispanic students in a treated school. Outcome measures include percent of Hispanic students with any in-school suspension (In-School Susp Rate), percent of Hispanic students with any out-of-school suspension (Out-of-School Susp Rate), percent of Hispanic students with single or multiple out of school suspensions, percent of Hispanic students expelled (Expulsion Rate), and percent of Hispanic students either arrested or referred to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by Hispanic student enrollment at the school level.

Table 1.11 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, by district racial composition

	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate	(4) Arrest or Referral Rate	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate	(4) Arrest or Referral Rate
	Panel 1a: All Students in Majority White Districts				Panel 2a: All Students in Minority White Districts			
ATT^{simple}	0.013**	0.002	0.002	0.000	-0.003	-0.001	0.002	-0.011
	(0.006)	(0.003)	(0.003)	(0.001)	(0.003)	(0.001)	0.002	(0.011)
Mean	0.042	0.017	0.027	0.005	0.043	0.026	0.034	0.008
Observations	21,285	21,311	21,313	21,156	19,545	19,569	19,579	19,562
	Panel 1b: All Students in Majority Black Districts				Panel 2b: All Students in Minority Black Districts			
ATT^{simple}	-0.005	-0.032***	-0.005	0.002	0.000	0.002	0.001	-0.009
	(0.020)	(0.009)	(0.005)	(0.002)	(0.003)	(0.001)	(0.002)	(0.009)
Mean	0.065	0.040	0.053	0.006	0.041	0.020	0.029	0.006
Observations	2,448	2,448	2,448	2,448	38,382	38,432	38,444	38,270
	Panel 1c: All Students in Majority Hispanic Districts				Panel 2c: All Students in Minority Hispanic Districts			
ATT^{simple}	0.005	0.006***	0.005**	0.003	-0.001	-0.003	0.000	-0.013
	(0.006)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.002)	(0.011)
Mean	0.030	0.019	0.028	0.007	0.046	0.022	0.031	0.006
Observations	8,294	8,294	8,302	8,295	32,536	32,586	32,590	32,423

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with single (Single Out-of-School Susp Rate) or multiple (Mult. Out-of-School Susp Rate) out of school suspensions, and percent of students with any arrests or referrals to law enforcement (Arrest or Referral Rate). Racial composition defined by pre-treatment percentage of student body at the district level. A district with a subgroup racial majority contains a student body where greater than 50% of students belong to that subgroup. All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Table 1.12 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated High Schools as controls, by district racial composition

	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate	(4) Arrest or Referral Rate	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate	(4) Arrest or Referral Rate
	Panel 1a: High School Students in Majority White Districts				Panel 2a: High School Students in Minority White Districts			
ATT^{simple}	0.018*	0.004	0.004	0.002	-0.005	-0.002	0.002	-0.014
	(0.009)	(0.004)	(0.005)	(0.002)	(0.008)	(0.005)	(0.003)	(0.015)
Mean	0.077	0.030	0.049	0.014	0.087	0.048	0.068	0.023
Observations	3,994	4,015	4,016	3,971	2,971	2,985	2,987	2,974
	Panel 1b: High School Students in Majority Black Districts				Panel 2b: High School Students in Minority Black Districts			
ATT^{simple}	-0.005	-0.059***	-0.008	0.012**	-0.001	0.002	0.002	-0.010
	(0.033)	(0.023)	(0.012)	(0.006)	(0.006)	(0.002)	(0.003)	(0.011)
Mean	0.130	0.078	0.092	0.014	0.078	0.035	0.055	0.018
Observations	388	388	388	388	6,577	6,612	6,615	6,557
	Panel 1c: High School Students in Majority Hispanic Districts				Panel 2c: High School Students in Minority Hispanic Districts			
ATT^{simple}	0.009	0.013***	0.006**	0.004	-0.001	-0.005	0.000	-0.013
	(0.013)	(0.003)	(0.003)	(0.006)	(0.007)	(0.004)	(0.003)	(0.013)
Mean	0.060	0.042	0.063	0.024	0.086	0.036	0.056	0.016
Observations	1,338	1,342	1,344	1,337	5,627	5,658	5,659	5,608

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with single (Single Out-of-School Susp Rate) or multiple (Mult. Out-of-School Susp Rate) out of school suspensions, and percent of students with any arrests or referrals to law enforcement (Arrest or Referral Rate). Racial composition defined by pre-treatment percentage of student body at the district level. A district with a subgroup racial majority contains a student body where greater than 50% of students belong to that subgroup. All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Table 1.13 Effect of COPS Hiring Grants Treatment on District Level Observables

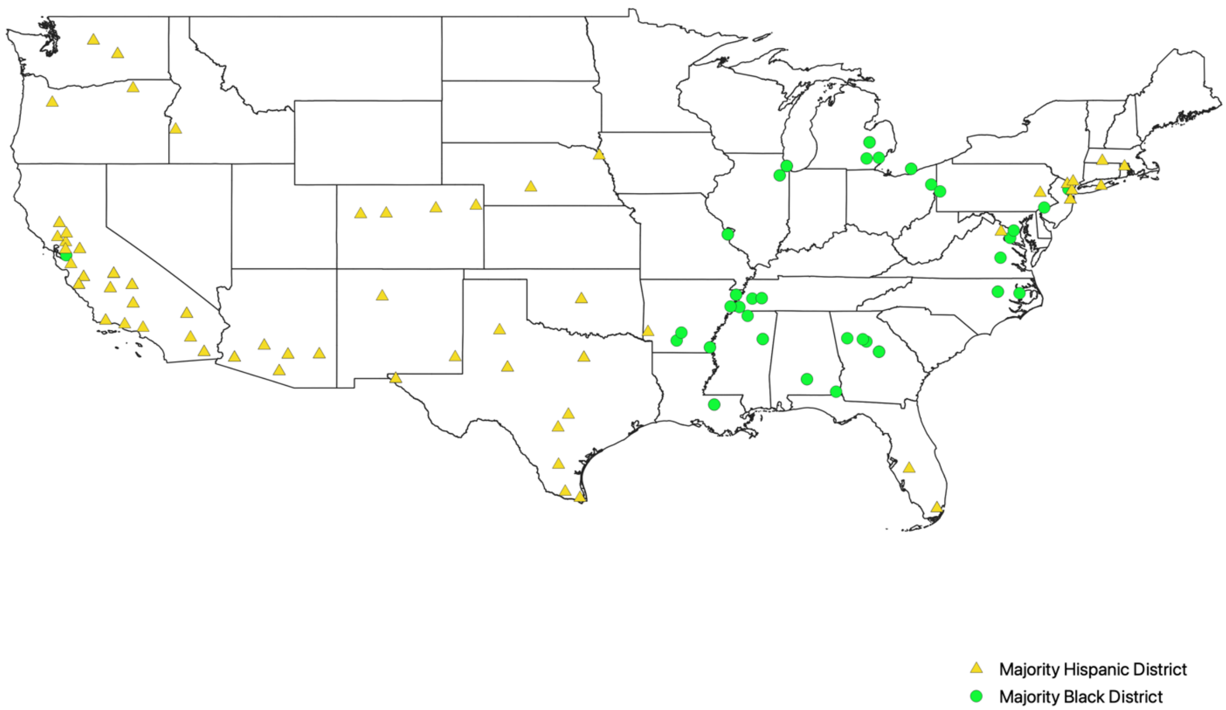
	(1) Percent White	(2) Percent Black	(3) Percent Hispanic	(4) Percent FRPL	(5) Percent SPED	(6) Percent ELL	(7) Teacher: Student Ratio	(8) Guidance Counselor: Student Ratio
ATT	0.006	-0.003	0.003	0.012	0.002	-0.014	-0.391*	-36.226
	(0.005)	(0.002)	(0.004)	(0.008)	(0.001)	(0.013)	(0.213)	(23.001)
Mean	0.645	0.095	0.179	0.416	0.141	0.064	16.080	552.975
Observations	3,344	3,344	3,344	3,344	3,330	3,291	3,316	3,154

Standard errors clustered at the school district level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: All outcomes are measured at the district level, and all specifications are weighted by student enrollment at the district level.

Figure 1.9 COPS Grant Applicants by Hispanic and Black School District Composition



Note: This map plots the latitude and longitude of police agencies that applied for COPS Hiring Program grants for school-based policing in the 2014-2017 grant cycles by their school district racial composition. Yellow triangles represent agencies that reside within a school district with a majority Hispanic students, while green circles represent agencies that reside within a school district with a majority Black students. Geographic data is sourced from the Bureau of Justice Statistics and the NCES Common Core of Data.

Table 1.14 School District Summary Statistics by Hispanic and Black School District Composition

	All School Districts	Majority Hispanic Districts	Majority Black Districts
	Mean/sd	Mean/sd	Mean/sd
Student : Guidance Counselor Ratio	554.67	833.84	456.29
	(461.40)	(988.19)	(202.41)
Student : Teacher Ratio	16.12	20.63	15.27
	(3.96)	(4.77)	(2.41)
Number of Schools in District	21.99	50.40	33.61
	(62.24)	(135.81)	(59.13)
Percent Urban Districts	0.56	0.58	0.66
	(0.50)	(0.50)	(0.48)
Observations	1,081	127	64
Percent White	0.62	0.15	0.17
	(0.29)	(0.14)	(0.12)
Percent Black	0.11	0.05	0.71
	(0.18)	(0.07)	(0.13)
Percent Hispanic	0.19	0.76	0.08
	(0.24)	(0.16)	(0.08)
Percent Male	0.52	0.51	0.51
	(0.01)	(0.01)	(0.01)
Percent Female	0.49	0.49	0.49
	(0.01)	(0.01)	(0.01)
Percent Special Education	0.14	0.11	0.14
	(0.07)	(0.03)	(0.04)
Percent English Language Learner	0.07	0.26	0.04
	(0.10)	(0.15)	(0.05)
Observations	1,206	142	67
Number of SROs Requested	1.68	2.19	2.53
	(1.62)	(2.39)	(2.92)
Number of Full Time Sworn Officers	53.14	66.46	102.11
	(119.85)	(111.73)	(241.92)
Observations	1,143	124	62
Percent Free or Reduced Price Lunch	0.43	0.49	0.74
	(0.23)	(0.28)	(0.14)
In-School Suspension Percent	0.06	0.06	0.10
	(0.06)	(0.07)	(0.09)
Multiple Out-of-School Suspension Percent	0.03	0.04	0.06
	(0.03)	(0.05)	(0.06)
Single Out-of-School Suspension Percent	0.04	0.06	0.07
	(0.03)	(0.05)	(0.05)
Observations	834	85	37
Number of SROs Per 1000 Students	0.50	0.28	0.44
	(1.10)	(0.39)	(0.33)
Percent White FTS Officers	0.82	0.55	0.59
	(0.23)	(0.30)	(0.29)
Percent Black FTS Officers	0.06	0.03	0.28
	(0.11)	(0.03)	(0.25)
Percent Hispanic FTS Officers	0.08	0.33	0.05
	(0.16)	(0.29)	(0.05)
Percent Male FTS Officers	0.89	0.86	0.81
	(0.14)	(0.20)	(0.26)
Observations	168	22	13

Note: School District level summary statistics data sourced from the NCES Common Core of Data 2014-2017 enrollment reports and the 2011 Civil Rights Data Collection. Police Agency level summary statistics data sourced from the Bureau of Justice Statistics Law Enforcement Management and Administrative Statistics 2016 survey. “FTS” stands for Full-Time Sworn.

Table 1.15 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, by Application Text Sentiment

	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate	(1) In-School Susp Rate	(2) Mult. Out-of-School Susp Rate	(3) Single Out-of-School Susp Rate
	PANEL 1: APPLICATIONS WITH POSITIVE SENTIMENT			PANEL 2: APPLICATIONS WITH NEGATIVE SENTIMENT		
	Panel 1a: White Students			Panel 2a: White Students		
ATT^{simple}	0.002	0.001	0.002	0.002	0.001	-0.001
	(0.004)	(0.002)	(0.003)	(0.003)	(0.001)	(0.001)
Mean	0.036	0.016	0.026	0.039	0.022	0.034
Observations	23,604	23,625	23,626	11,839	11,854	11,857
	Panel 1b: Black Students			Panel 2b: Black Students		
ATT^{simple}	-0.005	0.000	-0.003	0.001	-0.020**	-0.007
	(0.010)	(0.005)	(0.008)	(0.009)	(0.010)	(0.005)
Mean	0.076	0.040	0.056	0.079	0.047	0.062
Observations	22,274	22,344	22,352	10,719	10,726	10,741
	Panel 1c: Hispanic Students			Panel 2c: Hispanic Students		
ATT^{simple}	0.002	0.002	0.000	0.005	0.005*	0.004**
	(0.005)	(0.001)	(0.002)	(0.004)	(0.003)	(0.002)
Mean	0.041	0.019	0.029	0.043	0.023	0.033
Observations	23,609	23,629	23,640	11,810	11,831	11,833

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), and percent of students with single (Single Out-of-School Susp Rate) or multiple (Mult. Out-of-School Susp Rate) out of school suspensions.

1.9 Appendix Tables & Figures

Table 1.16 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, unweighted estimates

	(1) In-School Susp Rate	(2) Out-of-School Susp Rate	(3) Mult. Out-of-School Susp Rate	(4) Single Out-of-School Susp Rate	(5) Expulsion Rate	(6) Arrest or Referral Rate
			PANEL A: ALL STUDENTS			
ATT^{simple}	0.000	0.000	0.000	0.000	0.000	-0.008
	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.008)
Mean	0.043	0.051	0.021	0.031	0.002	0.006
Observations	40,830	40,857	40,880	40,892	40,592	40,718
			PANEL B: WHITE STUDENTS			
ATT^{simple}	-0.001	0.000	-0.001	-0.001	0.001**	-0.011
	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.010)
Mean	0.037	0.046	0.018	0.029	0.002	0.006
Observations	39,967	39,959	40,004	40,004	39,712	39,201
			PANEL C: BLACK STUDENTS			
ATT^{simple}	0.000	-0.002	0.000	-0.001	-0.001	-0.012
	(0.005)	(0.005)	(0.003)	(0.004)	(0.001)	(0.011)
Mean	0.077	0.097	0.042	0.057	0.004	0.012
Observations	37,421	37,410	37,513	37,535	37,312	36,353
			PANEL D: HISPANIC STUDENTS			
ATT^{simple}	0.001	0.005*	0.002	0.003*	0.000	-0.009
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.009)
Mean	0.042	0.049	0.020	0.031	0.002	0.007
Observations	40,002	40,002	40,049	40,066	39,771	39,037

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data Source: Civil Rights Data Collection 2011, 2013, 2015, 2017

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). Percentages are calculated at the school level.

Table 1.17 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated districts as controls, district-level

	(1) In-School Susp Rate	(2) Out-of-School Susp Rate	(3) Mult. Out-of-School Susp Rate	(4) Single Out-of-School Susp Rate	(5) Expulsion Rate	(6) Arrest or Referral Rate
PANEL A: ALL STUDENTS						
ATT^{simple}	0.000	0.001	0.000	0.001	0.000	-0.009
	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.008)
Mean	0.051	0.053	0.021	0.032	0.003	0.006
Observations	3,342	3,342	3,342	3,342	3,340	3,341
PANEL B: WHITE STUDENTS						
ATT^{simple}	0.002	0.002	0.001	0.001	0.000	-0.007
	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.004)
Mean	0.045	0.046	0.018	0.030	0.003	0.005
Observations	3,342	3,340	3,342	3,342	3,340	3,339
PANEL C: BLACK STUDENTS						
ATT^{simple}	-0.009	-0.010	-0.009	-0.001	0.000	-0.003
	(0.007)	(0.007)	(0.006)	(0.004)	(0.001)	(0.003)
Mean	0.097	0.103	0.043	0.063	0.005	0.012
Observations	3,260	3,261	3,262	3,261	3,258	3,257
PANEL D: HISPANIC STUDENTS						
ATT^{simple}	0.004	0.006**	0.003**	0.003	0.000	-0.002
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.003)
Mean	0.053	0.055	0.022	0.035	0.003	0.006
Observations	3,301	3,300	3,301	3,301	3,297	3,298

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Data Source: Civil Rights Data Collection 2011, 2013, 2015, 2017

Note: Percentages are calculated at the district level across all schools within a district. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). All specifications are weighted by student enrollment at the district level.

Table 1.18 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, elementary schools

	ELEMENTARY SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
ATT^{simple}	-0.001	0.000	0.000	0.000	0.000	-0.006
	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.005)
Mean	0.017	0.029	0.012	0.018	0.001	0.002
Observations	26,511	26,509	26,513	26,516	26,321	26,421
	PANEL B: WHITE STUDENTS					
ATT^{simple}	0.001	0.001	0.001	0.000	0.000	-0.003
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.003)
Mean	0.016	0.029	0.011	0.018	0.000	0.002
Observations	25,866	25,862	25,870	25,872	25,682	25,349
	PANEL C: BLACK STUDENTS					
ATT^{simple}	-0.007	-0.005	-0.005	0.000	0.000	-0.002
	(0.004)	(0.007)	(0.005)	(0.004)	(0.000)	(0.002)
Mean	0.035	0.060	0.025	0.036	0.001	0.004
Observations	24,334	24,308	24,336	24,338	24,209	23,521
	PANEL D: HISPANIC STUDENTS					
ATT^{simple}	0.001	0.003	0.002	0.001	0.000	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.002)
Mean	0.014	0.024	0.010	0.015	0.001	0.002
Observations	25,958	25,947	25,961	25,967	25,774	25,271

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Table 1.19 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, middle schools

	MIDDLE SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
ATT^{simple}	-0.002	0.000	-0.003	0.003	0.000	-0.016
	(0.007)	(0.006)	(0.005)	(0.002)	(0.001)	(0.015)
Mean	0.099	0.092	0.040	0.053	0.004	0.012
Observations	7,278	7,284	7,291	7,297	7,260	7,276
	PANEL B: WHITE STUDENTS					
ATT^{simple}	0.005	0.004	0.002	0.002	0.000	-0.007
	(0.007)	(0.004)	(0.002)	(0.002)	(0.001)	(0.005)
Mean	0.084	0.078	0.032	0.047	0.004	0.010
Observations	7,219	7,210	7,223	7,221	7,184	7,088
	PANEL C: BLACK STUDENTS					
ATT^{simple}	-0.009	-0.015	-0.025	0.008	0.000	-0.006
	(0.019)	(0.022)	(0.020)	(0.006)	(0.002)	(0.007)
Mean	0.174	0.173	0.081	0.097	0.007	0.023
Observations	6,718	6,728	6,763	6,777	6,742	6,600
	PANEL D: HISPANIC STUDENTS					
ATT^{simple}	-0.002	0.006	0.001	0.005	0.000	-0.001
	(0.009)	(0.006)	(0.003)	(0.003)	(0.001)	(0.006)
Mean	0.099	0.093	0.039	0.056	0.004	0.012
Observations	7,167	7,163	7,174	7,184	7,148	7,010

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Table 1.20 Simple aggregation of Callaway and Sant'Anna estimates of COPS Hiring Grants on student discipline using never treated and later treated schools as controls, by race category

	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
PANEL A: ASIAN						
ATT^{simple}	-0.001	0.001	0.001	0.000	0.000	-0.029*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.016)
Mean	0.019	0.021	0.006	0.015	0.001	0.003
Observations	34,389	34,395	34,403	34,403	34,120	34,138
PANEL B: NATIVE AMERICAN						
ATT^{simple}	-0.006	0.004	0.002	0.004	0.000	-0.005
	(0.009)	(0.010)	(0.005)	(0.006)	(0.002)	(0.007)
Mean	0.053	0.066	0.025	0.043	0.003	0.009
Observations	24,412	24,372	24,444	24,463	24,319	24,185
PANEL C: PACIFIC ISLANDER						
ATT^{simple}	0.004	0.000	-0.007	0.005	0.002	0.000
	(0.005)	(0.009)	(0.007)	(0.005)	(0.001)	(0.002)
Mean	0.034	0.041	0.012	0.030	0.001	0.006
Observations	14,308	14,297	14,328	14,325	14,320	14,266
PANEL D: TWO OR MORE RACES						
ATT^{simple}	0.006	0.007	0.002	0.004	0.000	-0.024
	(0.006)	(0.005)	(0.002)	(0.003)	(0.001)	(0.015)
Mean	0.055	0.070	0.029	0.044	0.003	0.009
Observations	34,581	34,533	34,647	34,646	34,365	34,283

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

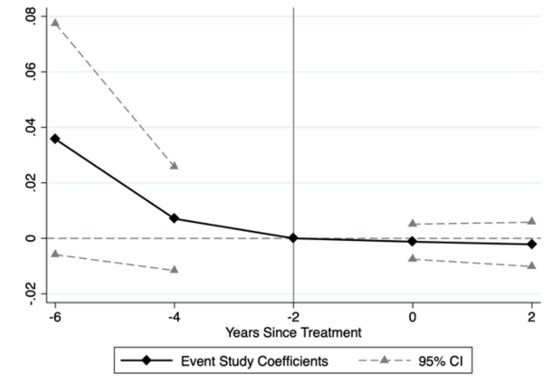
Note: All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate). In Panel C, I exclude the state of Hawaii.

Table 1.21 School District Summary Statistics by Hispanic and Black District Composition and Funded Status

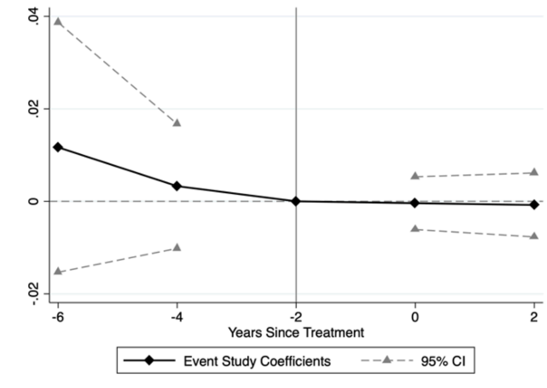
	Majority Hispanic Districts		Majority Black Districts	
	Funded	**Unfunded**	**Funded**	**Unfunded**
	Mean/sd	Mean/sd	Mean/sd	Mean/sd
Student : Guidance Counselor Ratio	1203.49	738.68	436.35	461.88
	(1915.88)	(517.78)	(175.70)	(210.57)
Student : Teacher Ratio	21.82	20.33	14.60	15.46
	(4.12)	(4.90)	(2.69)	(2.32)
Number of Schools in District	68.92	45.63	36.57	32.78
	(197.03)	(115.85)	(58.77)	(59.80)
Percent Urban Districts	0.54	0.59	0.79	0.62
	(0.51)	(0.49)	(0.43)	(0.49)
Observations	26	101	14	50
Percent White	0.13	0.16	0.13	0.18
	(0.08)	(0.15)	(0.13)	(0.11)
Percent Black	0.05	0.04	0.76	0.70
	(0.07)	(0.07)	(0.15)	(0.13)
Percent Hispanic	0.77	0.76	0.09	0.08
	(0.14)	(0.17)	(0.08)	(0.08)
Percent Male	0.51	0.51	0.51	0.51
	(0.01)	(0.01)	(0.01)	(0.01)
Percent Female	0.49	0.49	0.49	0.49
	(0.01)	(0.01)	(0.01)	(0.01)
Percent Special Education	0.11	0.11	0.14	0.14
	(0.02)	(0.04)	(0.05)	(0.04)
Percent English Language Learner	0.31	0.25	0.05	0.04
	(0.14)	(0.15)	(0.05)	(0.05)
Observations	28	114	14	53
Number of SROs Requested	2.81	2.03	2.64	2.50
	(3.21)	(2.11)	(2.93)	(2.95)
Number of Full Time Sworn Officers	104.50	56.37	95.36	104.08
	(161.26)	(92.94)	(169.13)	(260.82)
Observations	26	98	14	48
Percent Free or Reduced Price Lunch	0.45	0.50	0.80	0.72
	(0.30)	(0.28)	(0.11)	(0.15)
In-School Suspension Percent	0.03	0.06	0.12	0.09
	(0.04)	(0.08)	(0.11)	(0.08)
Multiple Out-of-School Suspension Percent	0.03	0.04	0.04	0.07
	(0.02)	(0.05)	(0.04)	(0.07)
Single Out-of-School Suspension Percent	0.05	0.06	0.06	0.07
	(0.04)	(0.05)	(0.05)	(0.05)
Observations	19	66	9	28
Number of SROs Per 1000 Students	0.34	0.24	0.41	0.45
	(0.33)	(0.42)	(0.55)	(0.32)
Percent White FTS Officers	0.58	0.53	0.36	0.63
	(0.27)	(0.33)	(0.43)	(0.27)
Percent Black FTS Officers	0.03	0.02	0.58	0.23
	(0.04)	(0.02)	(0.37)	(0.20)
Percent Hispanic FTS Officers	0.20	0.41	0.05	0.04
	(0.15)	(0.33)	(0.06)	(0.06)
Percent Male FTS Officers	0.78	0.91	0.79	0.82
	(0.32)	(0.06)	(0.16)	(0.28)
Observations	8	14	2	11

Note: School District level summary statistics data sourced from the NCES Common Core of Data 2014-2017 enrollment reports and the 2011 Civil Rights Data Collection. Police Agency level summary statistics data sourced from the Bureau of Justice Statistics Law Enforcement Management and Administrative Statistics 2016 survey. "FTS" stands for Full-Time Sworn.

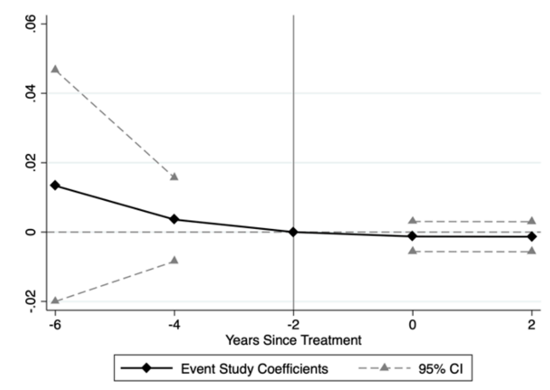
Figure 1.10 Stacked Event Study Aggregations for School-Level Discipline Outcomes



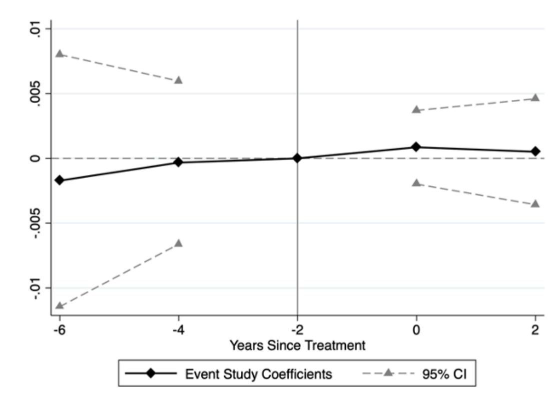
(a) In-School Susp. Rate



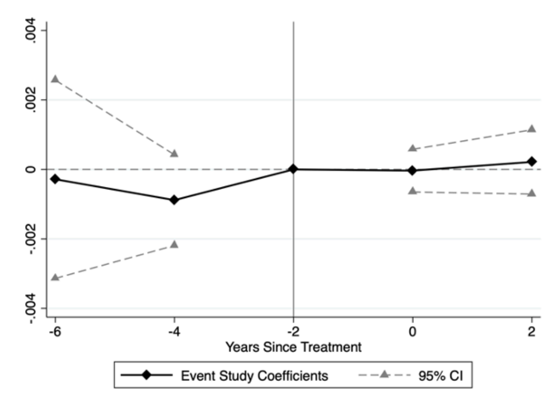
(b) Out-of-School Susp. Rate



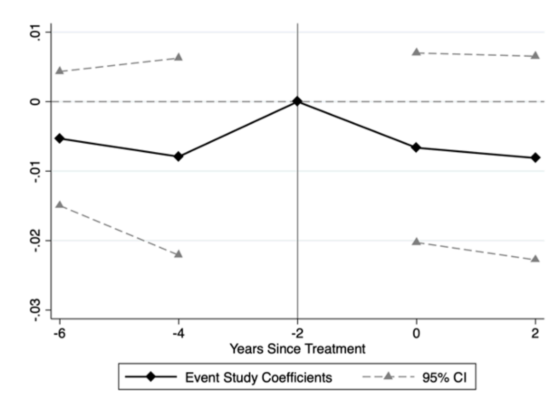
(c) Mult. Out-of-School Susp. Rate



(d) Single Out-of-School Susp. Rate



(e) Expulsion Rate



(f) Arrest or Referral Rate

Notes: This figure plots the event study aggregations following Deshpande and Li (2019) for students in a treated school. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students expelled (Expulsion Rate), and percent of students either arrested or referred to law enforcement (Arrest or Referral Rate). All outcomes are measured at the school level, and all specifications are weighted by student enrollment at the school level.

Table 1.22 Stacked Difference-in-Differences estimates of COPS Hiring Grants on student discipline using never treated schools as controls

	(1) In-School Susp Rate	(2) Out-of-School Susp Rate	(3) Mult. Out-of-School Susp Rate	(4) Single Out-of-School Susp Rate	(5) Expulsion Rate	(6) Arrest or Referral Rate
			PANEL A: ALL STUDENTS			
SBP x Post	-0.007	-0.003	-0.004	0.001	0.001	-0.003
	(0.006)	(0.004)	(0.003)	(0.002)	(0.000)	(0.004)
Mean	0.043	0.050	0.021	0.030	0.002	0.006
Observations	122,424	122,481	122,545	122,563	122,242	122,077
			PANEL B: WHITE STUDENTS			
SBP x Post	-0.008	-0.002	-0.005*	0.002*	0.000	-0.001
	(0.007)	(0.003)	(0.003)	(0.001)	(0.000)	(0.001)
Mean	0.038	0.044	0.017	0.028	0.002	0.005
Observations	120,989	120,975	121,078	121,080	120,749	120,019
			PANEL C: BLACK STUDENTS			
SBP x Post	-0.019	-0.008	-0.010	0.001	0.002	-0.001
	(0.013)	(0.008)	(0.006)	(0.006)	(0.001)	(0.002)
Mean	0.080	0.097	0.042	0.057	0.004	0.011
Observations	112,770	112,712	112,948	113,001	112,769	111,853
			PANEL D: HISPANIC STUDENTS			
SBP x Post	0.002	0.001	0.001	0.000	0.001	0.000
	(0.007)	(0.004)	(0.003)	(0.002)	(0.001)	(0.002)
Mean	0.043	0.049	0.020	0.031	0.002	0.006
Observations	120,315	120,270	120,400	120,429	120,125	119,337

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Percentages are calculated at the school level, all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate).

Table 1.23 Stacked Difference-in-Differences estimates of COPS Hiring Grants on student discipline using never treated schools as controls, elementary schools

	ELEMENTARY SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
SBP x Post	-0.001	-0.002	-0.001	-0.001	0.000	-0.003
	(0.002)	(0.003)	(0.002)	(0.002)	(0.000)	(0.002)
Mean	0.018	0.029	0.012	0.018	0.001	0.002
Observations	78,250	78,241	78,242	78,254	78,065	77,907
	PANEL B: WHITE STUDENTS					
SBP x Post	0.000	-0.001	-0.001	0.000	0.000	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)	(0.001)
Mean	0.016	0.028	0.011	0.018	0.001	0.002
Observations	77,048	77,036	77,049	77,057	76,867	76,435
	PANEL C: BLACK STUDENTS					
SBP x Post	-0.010	-0.005	-0.004	-0.001	0.000	-0.001
	(0.006)	(0.008)	(0.005)	(0.004)	(0.000)	(0.001)
Mean	0.038	0.060	0.026	0.036	0.001	0.003
Observations	71,921	71,827	71,916	71,922	71,809	71,232
	PANEL D: HISPANIC STUDENTS					
SBP x Post	0.003**	0.001	0.002	0.000	0.000	-0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)
Mean	0.015	0.025	0.010	0.016	0.001	0.001
Observations	76,742	76,700	76,734	76,753	76,571	76,094

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Percentages are calculated at the school level, and all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate).

Table 1.24 Stacked Difference-in-Differences estimates of COPS Hiring Grants on student discipline using never treated schools as controls, middle schools

	MIDDLE SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
SBP x Post	-0.017	-0.007	-0.012	0.004	0.000	-0.008
	(0.015)	(0.008)	(0.008)	(0.004)	(0.001)	(0.007)
Mean	0.095	0.085	0.036	0.049	0.004	0.010
Observations	22,775	22,790	22,809	22,811	22,782	22,758
	PANEL B: WHITE STUDENTS					
SBP x Post	-0.018	-0.007	-0.011**	0.004*	0.000	-0.005**
	(0.013)	(0.005)	(0.005)	(0.002)	(0.001)	(0.002)
Mean	0.080	0.071	0.029	0.044	0.004	0.008
Observations	22,646	22,625	22,654	22,654	22,617	22,474
	PANEL C: BLACK STUDENTS					
SBP x Post	-0.023	-0.008	-0.023	0.010	0.001	-0.004
	(0.030)	(0.017)	(0.015)	(0.010)	(0.003)	(0.005)
Mean	0.171	0.164	0.074	0.093	0.007	0.019
Observations	21,208	21,226	21,281	21,316	21,290	21,138
	PANEL D: HISPANIC STUDENTS					
SBP x Post	-0.009	-0.003	-0.003	0.001	0.001	-0.001
	(0.021)	(0.010)	(0.008)	(0.006)	(0.001)	(0.004)
Mean	0.095	0.088	0.035	0.054	0.004	0.010
Observations	22,496	22,470	22,508	22,517	22,493	22,339

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Percentages are calculated at the school level, and all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate).

Table 1.25 Stacked Difference-in-Differences estimates of COPS Hiring Grants on student discipline using never treated schools as controls, high schools

	HIGH SCHOOLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
	PANEL A: ALL STUDENTS					
SBP x Post	-0.016	-0.005	-0.007	0.002	0.001*	-0.002
	(0.012)	(0.006)	(0.005)	(0.004)	(0.001)	(0.005)
Mean	0.082	0.091	0.037	0.057	0.007	0.016
Observations	21,212	21,263	21,307	21,311	21,208	21,225
	PANEL B: WHITE STUDENTS					
SBP x Post	-0.016	-0.003	-0.008*	0.005**	0.001	0.001
	(0.013)	(0.005)	(0.004)	(0.002)	(0.001)	(0.002)
Mean	0.070	0.075	0.029	0.049	0.006	0.014
Observations	21,108	21,127	21,188	21,182	21,078	20,923
	PANEL C: BLACK STUDENTS					
SBP x Post	-0.046*	-0.021	-0.020**	-0.001	0.004*	0.001
	(0.024)	(0.013)	(0.010)	(0.009)	(0.002)	(0.005)
Mean	0.137	0.157	0.066	0.095	0.009	0.029
Observations	19,459	19,477	19,569	19,581	19,488	19,301
	PANEL D: HISPANIC STUDENTS					
SBP x Post	-0.001	-0.001	0.001	-0.002	0.001	0.001
	(0.015)	(0.007)	(0.006)	(0.006)	(0.001)	(0.004)
Mean	0.087	0.098	0.039	0.062	0.007	0.018
Observations	20,894	20,917	20,975	20,976	20,878	20,721

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Percentages are calculated at the school level, and all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate).

Table 1.26 Stacked Difference-in-Differences estimates of COPS Hiring Grants on student discipline using never treated schools as controls, treatment defined as number of officers requested per 1,000 students

	(1)	(2)	(3)	(4)	(5)	(6)
	In-School Susp Rate	Out-of-School Susp Rate	Mult. Out-of-School Susp Rate	Single Out-of-School Susp Rate	Expulsion Rate	Arrest or Referral Rate
			PANEL A: ALL STUDENTS			
SBP x Post	-0.555	-0.352*	-0.340	-0.013	-0.008	0.001
	(0.412)	(0.208)	(0.267)	(0.148)	(0.040)	(0.102)
Mean	0.043	0.050	0.021	0.030	0.002	0.006
Observations	122,424	122,481	122,545	122,563	122,242	122,077
			PANEL B: WHITE STUDENTS			
SBP x Post	-1.015*	-0.411*	-0.580**	0.167	0.051*	0.005
	(0.577)	(0.220)	(0.269)	(0.130)	(0.027)	(0.062)
Mean	0.038	0.044	0.017	0.028	0.002	0.005
Observations	120,989	120,975	121,078	121,080	120,749	120,019
			PANEL C: BLACK STUDENTS			
SBP x Post	-1.051	-0.787	-0.985**	0.116	0.046	0.069
	(0.897)	(0.589)	(0.456)	(0.364)	(0.113)	(0.123)
Mean	0.080	0.097	0.042	0.057	0.004	0.011
Observations	112,770	112,712	112,948	113,001	112,769	111,853
			PANEL D: HISPANIC STUDENTS			
SBP x Post	-0.023	0.025	0.080	-0.066	-0.024	0.081
	(0.392)	(0.165)	(0.181)	(0.111)	(0.036)	(0.072)
Mean	0.043	0.049	0.020	0.031	0.002	0.006
Observations	120,315	120,270	120,400	120,429	120,125	119,337

Standard errors clustered at the school district level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Percentages are calculated at the school level, and all specifications are weighted by student enrollment at the school level. Outcome measures include percent of students with any in-school suspension (In-School Susp Rate), percent of students with any out-of-school suspension (Out-of-School Susp Rate), percent of students with single or multiple out of school suspensions, percent of students with any expulsions (Expulsion Rate), and percent of students with either an arrest or a referral to law enforcement (Arrest or Referral Rate). In the above specification, treatment is defined as 0 for agencies who received no funding, and as the number of officers requested per 1,000 students for agencies who received funding.

CHAPTER 2

2 Examining the Effects of Tennessee's Third-grade Retention Policy on Student Achievement, Discipline, and Attendance

Dr. Susan Kemper Patrick and Kaitlyn Elgart[17]

2.1 Introduction

Many schools choose to retain students in the early elementary grades who do not demonstrate adequate academic progress and are deemed unprepared to enter the next grade. National estimates suggest that between five and ten percent of American students are retained in either Kindergarten or first grade (Child Trends, 2015; Frederick & Hauser, 2008; Warren et al., 2014). Since retaining students requires paying for an additional year of schooling, retention is considered one of the most expensive educational interventions. West (2012) estimated that the annual cost of grade retention in the U.S. exceeds $12 billion per year.

Retention has been used as an academic intervention throughout the United States since at least the early twentieth century, although retention decisions have traditionally been left up to school personnel (Jackson, 1975). While retention decisions are often left to the discretion of teachers or principals, some districts and states have adopted 'retention policies', which often mandate the retention of elementary students who do not demonstrate reading proficiency alongside targeted interventions for students. As of 2020, at least 17 states had adopted retention policies targeting third-grade students who are not proficient in reading (Education Commission of the States, 2020), with at least ten states passing retention legislation within the past decade (Workman, 2014). Third grade has often been a target year for statewide policies on retention and other literacy interventions because it has been identified as a critical age for building a solid foundation in literacy, and it is often the first grade in which students are tested using a statewide standardized assessment (Cummings et al., 2021; Annie E. Casey Foundation, 2010; Council of Chief State School Officers, 2019). This renewed policy focus surrounding third-grade literacy and retention suggests that these retention policies may influence increasing numbers of elementary students.

Research and policy debate around grade retention is contentious, with inconclusive evidence about the short and long-term effects of retention (Allen et al., 2009; Jimerson, 2001; Lorence, 2006; Valbuena et al., 2021). A main criticism of past retention research surrounds the design of the studies, with many scholars suggesting that retention research has traditionally compared the outcomes of retained students to inappropriate or misleading comparison groups (Hong & Raudenbush, 2005; Jackson, 1975; Jimerson, 2001; Lorence, 2006). More recent analyses of retention introduced more advanced quantitative methods to ameliorate this comparison group problem (e.g., Martorell & Mariano, 2018; Schwerdt et al., 2017). In particular, researchers have leveraged the fact that many high-profile retention policies, such as those in Florida and Chicago, required retention decisions to be based on

[17] This project was completed in partnership with the Tennessee State Board of Education (TSBE), with funding provided by Tennessee SCORE. The data was provided by the Tennessee Education Research Alliance (TERA), a partnership between the Tennessee Department of Education and Vanderbilt University. We particularly thank Amy Owen and Erika Leicht at TSBE as well as Erin O'Hara, Jessica Holter, and Laura Booker at TERA for their valuable feedback and support throughout the project.

proficiency cut-offs on standardized tests.

While prior descriptive research has found negative associations between retention and student outcomes, some more recent studies indicate that retention policies, which include both retention directives and academic interventions for targeted students, may have more positive effects than previous studies have shown (Greene & Winters, 2007, 2009; Jacob & Lefgren, 2004; McCombs et al., 2009; Roderick & Nagaoka, 2005; Winters & Greene, 2012). Many of these policies, including those in Florida, Chicago, and New York, mandate robust literacy interventions in addition to retention, including summer school, reading remediation, or other supplemental resources for students who fell below a certain reading proficiency cut-point. These analyses typically estimated the effects of retention policies using a regression discontinuity design, thus highlighting the policy effect only for those students scoring slightly below the proficiency cutoff, and typically cannot disentangle the effects of the retention policy from these other academic interventions partnered with mandated retention (Valbuena et al., 2021). Furthermore, the policies most studied in the prior literature left limited discretion to district-level or school-level decision-makers. Given that many states do not employ such strict retention policies, and many students targeted by these policies are not those with borderline scores right below the proficiency cutoff, it is important to understand the effects of retention policies in more generalizable policy contexts as well as for a larger subset of students.

This paper examines retention patterns and student outcomes in Tennessee, where legislation passed in 2011 requiring that third-grade students demonstrate reading proficiency before being promoted to fourth grade. Tennessee's 2011 retention law, which provided exemptions for students participating in research-based interventions and students with an Individualized Education Program (IEP), gave significant discretion to schools and districts to determine how to best implement the policy and decide which types of research-based interventions can serve students at risk of retention. As a result, the context surrounding Tennessee's retention policy varied significantly from the stricter policies frequently studied in most prior analyses on retention. In this paper, we leveraged a difference-in-differences approach to estimate the effect of the retention policy on achievement, attendance, and disciplinary outcomes for students targeted by the law (i.e., third-grade students who are not demonstrating proficiency in reading). While other papers studying retention policies often use a regression discontinuity approach to estimate the "local" effects of retention policy on marginal students (i.e., students right below the proficiency score cut-off), we employ a difference-in-differences strategy to examine the "global" effects of this policy change on a broader subset of Tennessee students. Our estimates therefore reflect the policy effect for not only the marginal student, but also the lowest performing students, a group that has not been previously studied in this context. We also explore heterogenous effects of the retention policy by students' demographic background and schooling history. Our analysis indicates that the 2011 Tennessee retention law had small but lasting positive effects on the reading achievement and disciplinary outcomes of students targeted by the law, and did not deter students from attending school in later grades. Further, we find evidence of heterogeneous effects for certain student subgroups.

2.2 Prior Research on Retention and Retention Policies

Elementary school retention rates are typically highest in Kindergarten and first grade, and estimated retention rates in later elementary grades have remained low even after the passage of numerous state laws targeting retention in third grade (Aud et al., 2013; Child Trends, 2015; Warren et al., 2014). Numerous meta-analyses, systematic reviews, and empirical studies have been published on the effects of retention (Allen et al. 2009; Holmes 1989; Holmes & Matthews, 1984; Jackson, 1975; Jimerson, 2001; Lorence, 2006; Valbuena et al., 2021). Taken together, the results are fairly inconsistent about whether retention has positive, negative, or no effects on the academic, socio-emotional, or behavioral outcomes of retained students. These inconsistencies are likely due to differences across time and across schools, districts, or states in how retention policies have been enacted as well as methodological differences in how research studies measured the effects of retention or retention policies. Some descriptive studies that estimate the naïve differences in outcomes between students who have experienced retention and those who have not typically conclude that retention has largely negative effects on students (Holmes, 1989; Holmes & Matthews, 1984) or that there is insufficient evidence that retention improved outcomes for students (Jackson, 1975; Jimerson, 2001). More recent observational studies have also found negative associations between retention and various outcomes (e.g., Andrew, 2014; Giano et al., 2021; Hong & Raudenbush, 2005; Hong & Yu, 2008; Hughes et al., 2018).

2.2.1 Retention Policy

Most recent research, including this study, has focused on estimating the effects of retention policies rather than retention itself. Despite prior literature that has shown largely negative impacts of grade retention, in the past two decades an increasing number of districts and states have implemented policies mandating retention or other interventions in certain grades, typically based on students' standardized reading assessment scores. These policies, in which retention is often paired with other mandates targeting reading instruction, literacy diagnostic screening, and academic interventions, are part of broader legislative efforts to improve early literacy (Education Commission of the States, 2020; Council of Chief City School Officers, 2019). These state policies sometimes exempt students from retention based on participation in certain interventions (e.g., summer school) or exempt English language learners and students with disabilities from these mandates. Quasi-experimental research estimating the effects of these newer retention policies often rely on regression discontinuity designs in which researchers compare students on either side of a proficiency cut-off (i.e., similarly performing students, only some of whom are targeted by the retention policy) while other studies leverage differences across time before and after the policy change (Greene & Winters, 2007, 2009; Hwang & Koedel, 2022; Jacob & Lefgren, 2004; Roderick & Nagaoka, 2005; Slungaard Mumma & Winters, 2023; Winters & Greene, 2012). These studies have historically focused on retention policies in a few districts and states, including Chicago, Florida, and New York City, although a few recent studies have expanded the scope of this research to more states (e.g., Hwang & Koedel, 2022; Slungaard Mumma & Winters, 2023; Cummings et al., 2021).

Most of these retention policies mandate both retention and additional academic interventions, such as summer school or supplemental literacy instruction during the retained year, based on whether students score below a proficiency cut-off on standardized reading assessments. Florida's third-grade retention policy is frequently cited in research and media because it represented one of the first and farthest reaching mandatory retention programs (Cummings et al., 2021). Starting in 2002, Florida mandated the retention of students who did not perform at a proficient level on the state's third-grade reading assessment, and third grade retention rates increased to 14% in the first year of implementation (from 3% the year before). Florida's retention law also mandated that retained students should attend summer reading camp, be assigned to a "high performing teacher," and receive an additional 90 minutes of reading instruction during their retained year. Analyses of Florida's third-grade retention policy have found that retained students have slightly better (but not significantly different) achievement outcomes one year after being retained (Greene & Winters, 2007, 2009), and long-run analyses found that positive achievement effects persisted through middle school (Schwerdt et al., 2015; Winters & Greene, 2012). Özek (2015) found that Florida's retention policy increased the likelihood of disciplinary incidents and suspensions in the first two years after retention, but that these differences attenuated over time.

Chicago Public Schools were one of the nation's first large districts that tied retention policy decisions directly to standardized test scores for 3rd, 6th, and 8th grade students. To study the impact of this policy, Jacob and Lefgren (2004) used a regression discontinuity approach estimating the net effect of attending both summer school and being retained in the subsequent year and found that "summer school and grade retention increased student achievement roughly 20 percent of a year's worth of learning". To isolate the effects of this policy, Roderick and Nagaoka (2005) limited their analysis only to students who originally qualified for summer school to isolate the effect of being retained an additional year. Their analysis found very small positive effects on reading growth after the first year that disappeared by the second year, and the authors concluded, "retention did not proffer any academic benefits to third graders who were retained nor did it have any substantial negative effect on their reading achievement". Taken together, this research on the retention policy implemented in Chicago indicates that for marginal students at risk of being retained, grade retention does not offer as much academic benefit in the long-run as targeted academic intervention, such as summer school.

In recent years, state-level policies regarding retention have expanded alongside research on the impacts of these policies on student outcomes. Researchers in Michigan have found that implementation of a third-grade retention policy improved student performance in high-stakes reading testing in elementary grades (Westall & Cummings, 2023). In 2014, Mississippi enacted the Literacy-Based Promotion Act, which required third-graders to be proficient in reading before being promoted to the fourth grade (Burk, 2020). As a result, standardized test scores on national assessments improved across the state for students impacted by this policy change (Burk, 2020; Slungaard Mumma & Winter, 2023). Additionally, there is no evidence that students impacted by this policy were deterred from attending school in later grades, or more likely to receive a special education designation in later grades (Slungaard Mumma & Winter, 2023). Research on a retention policy in Indiana finds that students targeted by the policy benefit from gains in academic achievement, while finding no sustained adverse impacts on attendance or disciplinary outcomes (Hwang & Koedel, 2022). In New York City, researchers further find that elementary school students impacted by retention requirements do not reduce attendance behavior and do not increase disciplinary incidence (Martorell & Mariano, 2018). These policies all couple retention mandates with additional academic interventions provided to at-risk students, thus we can interpret the results as an overall policy effect rather than the direct effect of retention.

2.2.2 Heterogeneous Effects

Descriptive studies of retention in the early grades consistently find that retention disproportionately impacts certain students. Students who are younger for their grade cohort are more likely to experience retention (Valbuena et al., 2021), boys are more likely to be retained than girls (Aud et al., 2013; Corman, 2003; Hong & Yu, 2007; Lorence & Dworkin, 2006), and retention rates are higher for students from low-income families (Aud et al., 2013; Child Trends, 2015; Corman, 2003; Frederick & Hauser, 2008; Lorence & Dworkin, 2006). Some studies find that Black and Hispanic students are more likely to be retained than their White peers, even when accounting for measures of academic achievement (Aud et al., 2013; Child Trends, 2015; Greene & Winters, 2009; Hong & Yu, 2007; Lorence & Dworkin, 2006; Warren et al., 2014). Time series estimates suggest that some of these gender, racial, and socioeconomic gaps in retention rates have been narrowing over the last decade (Brey et al., 2019; Warren et al., 2014).

Only a handful of studies have considered whether retention policies have heterogenous effects across student subgroups. Quasi-experimental research on the achievement effects of third grade retention policies in Florida and Indiana found little evidence of meaningful heterogeneity by race/ethnicity or gender (Hwang & Koedel, 2022; Schwerdt et al., 2017; Winters & Greene, 2012), although analyses on disciplinary outcomes find that increased disciplinary incidents are concentrated among male students and students from low-income families (Özek, 2015). A recent analysis of Louisiana's third grade retention policy found that positive effects on subsequent reading achievement were driven by Black and Hispanic students (Slungaard Mumma & Winters, 2023). Further research is needed to understand whether retention policies have heterogenous effects across student subgroups, and whether heterogeneity may depend on the specific policy or context. Additionally, these studies are limited by their reliance on regression discontinuity approaches, thus estimating the heterogeneous effects only for marginal students at risk of being targeted by retention policies. Further research is needed to disentangle the global effect of retention policy implementation for student subgroups.

2.3 Context of the Study

This study examined the 2011 third-grade retention policy in Tennessee, which has not been previously studied in analyses of state-level retention policies. As of 2009-10, approximately 4% of Tennessee's Kindergarten students were retained, 3% of first-grade students, 1% of second-grade students, and 0.6% of third-grade students. At this time, Tennessee's retention rates were slightly lower than reported national averages for early elementary grades (Aud et al., 2013; Warren & Saliba, 2012). In 2011, the Tennessee legislature passed Public Chapter 351 requiring that "a student in the third grade shall not be promoted to the next grade level unless the student has shown a basic understanding of curriculum and ability to perform the skills required in the subject of reading as demonstrated by the student's grades or standardized test results." The law went into effect in the 2011-2012 academic year. This policy exempted students who participated in a district-approved, research-based literacy intervention and students with disabilities served by individualized education plans (Tennessee Code Annotated § 49-6-3115(a)). Unlike prior retention policies (i.e., Florida, Chicago), which mandated a combination of retention and academic interventions for targeted students, Tennessee's policy allowed for students to either be retained or to participate in an academic intervention in lieu of retention. The legislation gave considerable autonomy to local districts to determine which students who qualified should be retained or receive an academic intervention, and to decide which research-based interventions to provide for their students in lieu of retention. Based on guidance from the Tennessee State Board of Education, local school districts were authorized to create their own guidelines on how to make these decisions. Given relatively low passing rates of third-grade standardized reading exams across the years before and after this policy change as displayed in Appendix Figure 2.7, this policy targeted a majority of Tennessee third-grade students, a much larger subset of students than other state-level retention policies have historically impacted. This policy was subsequently updated in 2021, however this analysis focuses only on the 2011 retention policy.

The current study focuses on Tennessee students entering third grade for the first time in the 2009-10 to 2013-14 academic years. During this time, the state of Tennessee served approximately 900,000 students in Kindergarten through 12th grade within its public school system, with approximately 70,000 students per grade cohort. We use student-level statewide administrative data to examine the effect of the third-grade retention law change on subsequent student outcomes for students targeted by the law. We specifically examined the retention law's influence on the subsequent achievement, attendance, and disciplinary outcomes of third-grade students targeted by the law (which we define based on their third-grade reading proficiency). We also explored whether these effects vary across student demographic subgroups and prior history with retention.

2.4 Data and Methods

This study uses student-level data in the state of Tennessee for school years 2006-2019 provided by the Tennessee Education Research Alliance (TERA), a research partnership between Vanderbilt University and the Tennessee Department of Education. Our analysis focused specifically on Tennessee students who were in third grade during the academic years immediately preceding and following the passage of the legislation. Because of the longitudinal nature of the data, we can follow these five cohorts of students as they continue through their academic careers in Tennessee schools. The analytic sample included all Tennessee students who entered third grade in academic years 2009-10 through 2013-14 who received a test score for their third-grade reading standardized assessment (N=350,642 students). Because we use reading assessment data to identify students targeted by the law, third-grade students without a score are excluded (N=32,876 students, or 9% of all third graders in these five cohorts). Certain student subgroups (notably, English language learners and students with documented disabilities) are more likely to be missing a third-grade test score, thus any heterogeneous effects for those subgroups should be interpreted with caution. Table 2.1 includes descriptive statistics for the analytic sample, which consists of Tennessee students that are 48% female, 65% White, and 60% economically disadvantaged.

2.4.1 Measures

2.4.1.1 Student-Level Characteristics

We used student-level characteristics as covariates in certain analytic models and to separate students into subgroups to analyze heterogenous effects. Because several of these characteristics can change over time (i.e., Economically Disadvantaged status and English Learner Status), all student characteristics included in the models are measured as of a student's first year in third grade. Table 2.1 reports the sample averages of these characteristics pooled across years.

2.4.1.2 Outcome Measures

There are three primary outcome measures in this analysis: annual attendance rate, disciplinary record, and reading achievement. We observed each outcome annually for fourth through eighth grade. The annual attendance rate captured the percent of school days that a student was recorded as in attendance during a given academic year. The average annual attendance rate for the analytic sample was 95.3% (sd=5.5), with slight variation across grades. The measure of discipline used in this analysis indicated whether a student has any suspensions (in-school or out-of-school) within an academic year. This discipline measure excluded expulsions or any minor infractions that are not formally recorded. In elementary grades, students are less likely to receive disciplinary infractions than in middle school grades. In our sample, 6% of fourth-grade students had at least one suspension, as compared to 16% of eighth-grade students.

Finally, we used state-administered standardized exams in English language arts to identify reading proficiency level and reading achievement as measured by scale scores. Prior to the 2015-16 academic year, Tennessee administered the Tennessee Comprehensive Assessment Program (TCAP) annually to students in third through eighth grade. After the 2015-16 school year, Tennessee transitioned to a new testing regime, TNReady. Due to state-wide system complications, there was no testing in the 2015-16 school year and thus we have no achievement outcome measures in that school year. While the two testing regimes used similar proficiency levels for scoring (i.e., which students were scoring at or above proficient in English language arts), the scale scores varied significantly across grades and across testing regimes. For this reason, we standardized the scores within grade-year such that the average score for each grade-year test administration is 0 with a standard deviation of 1.

In all specifications, we compare student outcomes within a particular grade. Some prior literature on retention policies has compared students across grades to account for changes in student-grade composition induced by increases in retention of students which could impact comparison groups within-grade (Greene & Winters, 2007b; Valbuena et al., 2021b). However, in the policy context in Tennessee, retention rates remain consistently low before and after the policy change which allows us to avoid any complications in comparison groups that could be induced by an increase in student retention. Further, the Tennessee standardized tests used in these years assessed grade-specific standards and scoring is not comparable across grades, which makes comparisons of achievement outcomes across grade levels difficult to interpret.

2.4.1.3 Retention Measures

Because student-level retention decisions are not directly measured by Tennessee's administrative data system, we created a retention measure using annual student enrollment records to study descriptive patterns in grade retention. The enrollment records included a minimum and maximum grade for each student in each school year, and we identified a student as retained if their enrollment records indicated that they spent two full consecutive academic years in the same grade. This is a conservative measure of retention because it does not include any students who may have switched grades over the course of the academic year or students who were promoted mid-year.

Table 2.2 shows K-2 and third-grade retention rates for the five cohorts of third-grade students included in this analysis. Overall, 10.5% of third graders had been retained before entering third grade and 0.9% were retained in third grade. As in prior research, retention rates prior to third grade vary by gender, economic disadvantage, and disability status, although differences across subgroups shrink for third-grade retention rates. Notably, students scoring below proficient on the third-grade reading assessment are more than three times as likely to have been retained before third grade than students who scored at or above proficient, but only 1.4% of students scoring below proficient are retained in third grade.

Figure 2.1 illustrates third-grade retention rates by proficiency status across years. Across all years of the study, students at or above proficient on the third-grade reading assessment are almost never retained in third grade. Among students scoring below proficient, statewide retention rates in third grade remain below 2% consistently, even after the passage of the 2011 retention policy. As noted in Section III, the retention law gives districts the autonomy to determine research-based interventions that can be offered in lieu of retention, thus we can infer that in lieu of an increase in retention practices in the wake of this retention law passage, it is more likely that these students experienced increased exposure to targeted academic interventions. We further explored variation across schools and districts in third-grade retention rates, and Figure 2.2 includes a histogram of school-level retention rates for third-grade students scoring below proficient on their third-grade reading assessment for the school years immediately preceding and following the third-grade retention law change. Notably, the majority of schools in Tennessee did not retain any third-graders scoring below proficient, but retention rates and practices appear to vary considerably at the school level.

2.4.2 Analytic Approach

To evaluate the policy effect of Tennessee's 2011 third-grade retention law, we used a difference-in-differences approach to estimate an overall effect on the students targeted by the law (which we define as third-grade students who are performing below grade-level reading proficiency according to their third-grade standardized assessment). The third-grade retention law specifically targeted students who are not demonstrating proficiency in reading, thus a below-proficient third-grade reading score could trigger students' retention or inclusion in "research-based interventions" as determined by schools and districts. As such, this model compared outcomes for third-grade students who scored below the proficiency cut-off on the Tennessee third-grade standardized reading assessment to those students who scored at or above proficient in the periods right before and after the law was implemented. We estimated the effect of the third-grade retention law with pooled cross-sectional data using the following empirical specification:

$$y_{it} = \beta_0 + \beta_1 NotProf_i + \beta_2 Year_t + \beta_3 PolicyEffect_i + \beta_4 X_i + \theta_s + \epsilon$$

where y_{it} is the outcome of interest for student i at time t (i.e., annual attendance rate, disciplinary record, or standardized achievement score for a given grade), $NotProf_i$ is equal to 1 if student i scored below proficient on their third-grade reading assessment and 0 if student i scored at or above proficient, $Year_t$ represents year fixed effects, and $PolicyEffect_i$ is equal to 1 if a student was in third grade after the policy went into effect and scored below proficient on their third grade reading assessments. X_i is a vector of student-level covariates as measured during the students' first year in third grade, and θ_s represent school-level fixed effects.[18] When estimating the effect on the likelihood of having a disciplinary record, we used linear probability models. The coefficient of interest, β_3, captures the effect of the third-grade law change on third-grade students targeted by the policy.

We estimated three specifications: a base model excluding any student covariates or school fixed effects, a model including student-level covariates, and our preferred specification which includes a vector of student-level covariates as well as school-level fixed effects to account for variation in implementation of the third-grade retention policies by student demographics and across schools. Given that districts could also set retention policies, we performed a robustness check using district-level fixed effects and found substantively similar results. To evaluate heterogeneous effects, we used the same models described above while splitting the sample into subgroups based on certain student characteristics (i.e., retention history, race/ethnicity, gender, economic status, disability status). In all models, standard errors are clustered at the school level. The primary analysis included five cohorts of students who entered third grade between 2009-2010 and 2013-2014.

This difference-in-difference approach allows us to study the policy effect of these widely-adopted third-grade retention policies on an under-researched group of students, given that prior literature often focuses on the local effect for students who marginally fail their exams (Valbuena et al., 2021). The estimates from our model measure the global policy effect for all students targeted by the law, and in particular this model allows us to include the lowest-performing subgroup of students in our estimate, a group that is most likely to be impacted by retention or intervention but has not previously been studied in recent retention literature. Further, due to the nature of the 2011 Tennessee retention policy, retention rates for third grade students did not increase dramatically in the years

[18] We assign schools to students based on their enrollment in their first year in third grade. We first restrict the enrollment file to student-by-year-by-school spells in which students are enrolled for at least 20 instructional days and then assign each student to the school with the last enrollment spell for third grade. We anticipated that schools were making decisions about retention or interventions in lieu of retention at the end of third grade.

following implementation, and those students that did experience an increase in retention were those scoring on the lowest end of the achievement distribution, limiting the ability to estimate a discontinuity in retention rates at the proficiency threshold in this context as demonstrated in Appendix Figure 2.8. The key identifying assumption underlying the difference-in-difference approach is that in the absence of the retention policy implementation, outcome trends for the targeted group (i.e., students scoring below proficient on their third-grade assessment) would have developed similarly to the trends of similarly performing students who entered third grade before the law change, and in parallel to the trends of students not targeted by the law (i.e., students scoring proficient or higher on their third-grade assessment). We present Appendix Figures 2.4-2.6 which highlight the time trends in outcomes by third-grade cohort year and illustrate that our outcomes of interest adhere to the necessary parallel trends assumption.

There are important limitations of this methodological approach. First, while we opt to identify the treatment group based on third-grade test scores, the retention law indicated that third-grade students should not be promoted unless they have "shown a basic understanding of curriculum and ability to perform the skills required in the subject of reading as demonstrated by the student's grades or standardized test results." Given that we cannot observe grades in this analysis, it is possible that districts or schools decided to promote students who scored below proficient on their standardized reading assessment because of their classroom grades without offering a research-based intervention. However, this source of bias would cause attenuation in our results, and work against us finding a statistically significant result. Further, assigning treatment in our analytic model based on standardized test scores rather than a teacher-determined grading scheme eliminates any teacher or school-level impact on manipulating the treatment group through grade inflation. We include robustness tables that limit the sample to students scoring within the middle of the distribution of test scores to observe the policy effect for an academically similar group of students who are more likely to be targeted by similar teaching and grading strategies throughout the school year and find that our analyses remain robust to this subgroup analysis. Further, we may be concerned that teachers are able to adjust teaching strategy over time to better target struggling students that could be impacted by this policy change, thus changing the average treatment effect for students in subsequent years of policy implementation. We include a robustness table limiting the sample to observing the first cohort of students targeted by this policy change, and find that our results are robust to this analysis.

Second, this analysis can only estimate the effect of the retention law change, not the actual effect of student-level retention or unobservable academic intervention decisions. As shown in Figure 2.1, only a small percentage of students targeted by the law were retained each year. Our analytic approach estimates the effect of the policy change as a whole, and can be interpreted as the intent-to-treat effect of a policy bundle of both direct retention of targeted students and the research-based interventions provided to non-retained students. Finally, this approach cannot distinguish the effect of the retention law change from other statewide policy changes that went into effect during the same academic year which may have differentially impacted third-grade students scoring below proficient in reading. To the best of our knowledge there are no other concurrent statewide policies targeting students confounding these results.[19]

2.5 Results

The primary findings from the difference-in-differences models estimating the effects of the third-grade retention law change on subsequent student outcomes in fourth through eighth grade are shown in Table 2.3 (annual attendance rate), Table 2.4 (disciplinary record), and Table 2.5 (standardized achievement scores). For each outcome, Panel A displays results for the base model, Panel B displays results for model adding student-level covariates, and Panel C displays results for the full model including student-level covariates and school-level fixed effects.

As shown in Table 2.3, we found mixed and mostly insignificant effects of the retention law change on the subsequent attendance of targeted students (i.e., those scoring below proficient on their third-grade reading assessment), with all coefficients close to zero. The statistically significant results for eighth grade are unlikely to be practically meaningful (a 0.20-0.25 difference on a 0-100 scale). This indicates that students targeted by this policy are not deterred from attending school in later grades as a result of their inclusion in either grade retention or academic intervention programs. Further, Table 2.4 illustrates that there were consistent and statistically significant effects on the likelihood of having a disciplinary record, with students targeted by the retention law exhibiting reductions in their probability of receiving a disciplinary action in fifth through eighth grade. The magnitude of these effects varied across grades (a 0.8% decrease in the likelihood of having a disciplinary record in fifth grade to a 2.5% decrease in sixth grade). In elementary grades, students are less likely to receive disciplinary infractions than in middle school grades so the larger magnitudes for later grades should be interpreted with those differences in mind. In the analytic sample, 6% of fourth-grade students had at least one suspension compared to 16% of eighth-grade students. Table 2.5 displays the effects on standardized reading achievement, which shows consistent and statistically significant effects on subsequent reading achievement in fourth through eighth grades. These positive effects on reading achievement (standardized within grade-year) were relatively small in magnitude (between 0.02-0.08 standard deviations). Based on average testing scale scores in the 2013-14 school year, the positive effects calculated here translated to an increase in between 1-3 scale score points, or approximately 3-10% of the standardized reading achievement gap between economically disadvantaged students and those not classified as economically disadvantaged. For both discipline and achievement outcomes, the effects persisted through eighth grade and in some cases grew larger in magnitude in the later grades.

In addition to these average effects, we also examined heterogeneity across certain student demographic characteristics (i.e., sex, race/ethnicity, economic status, disability status, and retention experiences). These subgroup results are shown in Appendix Tables 2.6-2.11. We found larger effects on disciplinary actions and

[19] Tennessee has implemented several statewide changes to education policy during the academic years studied in this analysis, including the implementation of statewide teacher evaluation system in 2011-2012 and the statewide implementation of Response to Intervention (RTI) in 2014-2015. While the statewide teacher evaluations system was implemented in the same year as the third-grade retention policy, we have no reason to suspect it would differentially impact third-grade students scoring below proficient in reading. In contrast, while the statewide implementation of RTI may differentially impact lower-performing students, it was implemented four years after the third-grade retention law. See Gilmour et al (2022) for more on the implementation of RTI.

subsequent reading achievement for students who were identified as economically disadvantaged in third grade, male students, and Black middle school students. We find inconsistent results for students who were identified as having disabilities as of the third grade. While the third-grade retention law specifically exempts students with individualized education plans (IEPs) from inclusion in policy interventions, students with disabilities in Tennessee are more likely to experience retention in the early grades.[20] However, given that students with disabilities are less likely to participate in standardized testing requirements, interpreting results for this subgroup could be problematic given potential for sample selection.

Finally, we examined heterogenous effects based on students' retention history. We categorize students as "never retained" if they were not retained in Kindergarten through third grade and "ever retained" if they were retained at least once between Kindergarten and third grade. In this analysis (shown in Figure 2.3), we find that the effects on subsequent reading achievement and discipline hold for students who have never experienced retention, suggesting that the main effects described above may not be driven by retention alone.

To test the robustness of our analysis, we analyzed several alternative specifications to account for potential confounders which are shown in Appendix Tables 2.12-2.14. First, we restricted our sample to the third-grade cohorts immediately preceding and following implementation of the retention law (students in third grade in 2010-11 and 2011-12) to account for any other targeted education programs that could have been implemented around the same time as well as potentially heterogeneous treatment effects by third-grade cohort. Next, we ran a robustness check in which we limited the sample to students who have test scores and enrollment records in all years of analysis (3rd through 8th grade) to account for any attrition or sample bias from students or parents specifically selecting out of Tennessee public schools as a result of this policy change. Finally, we restricted the sample to students just below the proficiency cutoff (proficiency level 2 in Tennessee's standardized testing regime) and just above the cutoff (level 3) to examine how this policy change impacted the marginal students in Tennessee who were likely more similar in performance throughout the school year and therefore more likely to benefit from similar instruction leading up to their third grade examinations. Our results were robust to the first two specifications, and we found that the magnitude of our results decreased in the specification that restricts to students just above and below the proficiency cutoff, signaling that students who were performing at the lowest level of reading proficiency in third grade could benefit the most from this retention policy.

2.6 Implications and Conclusions

In this study, we used statewide administrative data from Tennessee to study whether a third-grade retention law passed in 2011 affected short- and medium-term outcomes for students targeted by the law. While most prior studies on third-grade retention policies have focused solely on achievement outcomes, we examined effects on three different outcomes: subsequent reading achievement as measured by standardized reading assessment scores, disciplinary records capturing whether students received a suspension, and annual attendance rate. We found

[20] Notably, the analysis only includes students with disabilities who received a third-grade reading assessment score.

consistent evidence that the law change had small, positive effects for subsequent disciplinary and reading achievement for those students targeted by the law, and that this did not deter students from attending school. The estimated effects for both achievement and disciplinary outcomes persisted through eighth grade.

Prior research on other third-grade retention policies has found positive effects on subsequent reading achievement that persisted into middle school, and studies on retention policies in Florida, Indiana, and Louisiana reported effects that were larger in magnitude than the current analysis (Hwang & Koedel, 2022; Schwerdt et al., 2017; Slungaard Mumma & Winters, 2023). Given that retention policies in these other states included much more restrictive policies and specific mandated interventions (i.e., summer school, tutoring in small groups, assignment of highly effective teacher), the smaller effects in Tennessee may reflect heterogeneity in how districts supported students targeted by the retention policy as well as less intensive interventions that may have been provided in Tennessee compared to other contexts. Unlike prior studies, we find positive effects for student discipline, such that students targeted by the policy are less likely to experience suspensions in subsequent years. Only a few studies have examined the effects of third grade retention policies on discipline, and these studies found either null (Hwang & Koedel, 2022; Martorell & Mariano, 2018) or negative effects in the form of increased discipline rates (Özek, 2015). We also found evidence of heterogenous effects by certain student subgroups, with larger discipline and achievement effects for economically disadvantaged students, male students, and Black middle school students. While other papers studying retention policies often use a regression discontinuity approach to estimate the "local" effects of retention policy on marginal students, we employ a difference-in-differences strategy to examine the "global" effects of this policy change on a broader subset of Tennessee students. Our estimates therefore reflect the policy effect for a broader group of students than has been previously studied in this context.

Tennessee's 2011 third-grade retention law varied considerably from the high-profile policies most commonly studied. Prior studied policies mandated retention and/or additional academic interventions based on reading proficiency score cut-offs without giving much discretion to school personnel. While these policies have received considerable attention in research and media stories about retention, many states employ retention policies that give greater discretion to districts or schools to make decisions about retention (Education Commission of the States, 2020). The enactment of Tennessee's third-grade retention policy did not lead to dramatic increases in third-grade retention rates, suggesting that most Tennessee schools and districts decided to offer academic interventions in lieu of retention for students who did not demonstrate proficiency in reading. In our analysis, the effects on subsequent discipline and reading achievement hold for students targeted by the law (i.e., students scoring below proficient on their third-grade standardized reading assessment) but who had never experienced retention. This pattern could suggest that the law's positive effects may be driven by interventions offered in lieu of retention rather than retention itself.

While our analysis could not specifically observe which interventions or supports were provided to students in lieu of or alongside retention, given prior research on the effects of retention (Valbuena et al., 2021), these academic interventions may be more beneficial for students and more cost-effective than retention itself. For example, there is a robust evidence base that summer school programs and tutoring can have significant and substantial effects on achievement for struggling students (e.g., see Matsudaira, 2008; Zvoch & Stevens, 2013; Nickow et al., 2020; Wanzek et al., 2016). Across the years studied in this analysis, more than half of Tennessee's third-grade students scored below proficient on their standardized reading assessment in third grade, suggesting that earlier interventions for struggling students may be worthwhile. Tennessee has already made considerable investments in early literacy—including statewide implementation of teacher training, universal diagnostic screeners, and additional instruction for struggling students—that have offered many students additional literacy support in the early grades (Tennessee Department of Education, 2017, 2023).

About one in ten Tennessee students were retained between Kindergarten and third grade but retention rates vary considerably across student characteristics. As in prior research on retention rates (Aud et al., 2013; Child Trends, 2015; Lorence & Dworkin, 2006), certain Tennessee students were more likely to experience retention. However, our heterogeneity analyses suggested that the effects of the third-grade retention law were stronger for economically disadvantaged students and male students, groups which have historically had higher retention rates in the early grades. Given that we find our results are most likely driven not by grade retention, but by the additional academic interventions offered to students, this could indicate that students who have traditionally been targeted by grade retention as an intervention may find greater benefit from academic interventions in lieu of retention. Finally, when we limited our analysis to students closer to the proficiency cut-off, we found attenuated effects on reading achievement and discipline. These patterns indicate that retentions policies more generally and the third-grade retention law specifically did not have a uniform effect across the distribution of student achievement. Given that prior research on heterogeneous effects of retention or retention policy has been inconclusive, additional research is needed to better unpack how retention policies may differentially impact certain students.

Retention is one of the most expensive educational interventions that is commonly used in our nation's schools (West, 2012). The most recent national estimates calculated that retention costs exceeded $12 billion per year (West, 2012). A state-level analysis by the Texas Education Agency estimated that their state spent approximately $1.7 billion to retain about 190,000 students for an extra year in the 2000-2001 academic year (Valbuena et al., 2021). Given this cost, the implementation and effects of retention policies are under-researched. Recent research estimating the effects of some high profile retention policies—which tend to use quasi-experimental methods that compare students performing similarly on exams across a scoring cutoff—typically cannot disentangle the effects of retention from other academic interventions offered in lieu of or alongside retention. As with our study in Tennessee, many states do not systematically collect student-level data on academic interventions offered to struggling students. However, given that our findings suggest strong positive gains in academic outcomes for students targeted by academic interventions, future research should consider how to collect more data on these academic interventions over time to better understand the prevalence of these interventions and potentially disentangle the effect of retention from other academic interventions. Given the importance of the early grades in determining the educational trajectory of students, further analysis is needed to better understand the effects of retention and retention policies, which affect thousands of students every year in Tennessee and millions of students across the nation.

2.7 References

[1] Abbott, Mary, Howard Wills, Charles R. Greenwood, Debra Kamps, Linda Heitzman-Powell, and James Selig. 2010. "The Combined Effects of Grade Retention and Targeted Small-Group Intervention on Students' Literacy Outcomes." Reading & Writing Quarterly 26 (1): 4–25. https://doi.org/10.1080/10573560903396876.

[2] Abidin, Richard R., Wendy M. Golladay, and Anna L. Howerton. 1972. "Elementary School Retention: An Unjustifiable, Discriminatory and Noxious Educational Policy." Journal of School Psychology 9 (4): 410–17.

[3] Alexander, Karl L., Doris R. Entwisle, and Susan L. Dauber. 2003. On the Success of Failure: A Reassessment of the Effects of Retention in the Primary School Grades. Cambridge University Press.

[4] Allen, Chiharu S., Qi Chen, Victor L. Willson, and Jan N. Hughes. 2009. "Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-Analytic, Multilevel Analysis." Educational Evaluation and Policy Analysis 31 (4): 480–99.

[5] Andrew, Megan. 2014. "The Scarring Effects of Primary-Grade Retention? A Study of Cumulative Advantage in the Educational Career." Social Forces 93 (2): 653–85.

[6] Bali, Valentina A., Dorothea Anagnostopoulos, and Reginald Roberts. 2005. "Toward a Political Explanation of Grade Retention." Educational Evaluation and Policy Analysis 27 (2): 133–55. https://doi.org/10.3102/01623737027002133.

[7] Barrett-Tatum, Jennifer, Kristen Ashworth, and David Scales. 2019. "Gateway Literacy Retention Policies: Perspectives and Implications from the Field." International Journal of Education Policy and Leadership 15 (10): n10.

[8] Bowman-Perrott, Lisa J., Socorro Herrera, and Kevin Murry. 2010. "Reading Difficulties and Grade Retention: What's the Connection for English Language Learners?" Reading & Writing Quarterly 26 (1): 91–107. https://doi.org/10.1080/10573560903397064.

[9] Brey, Cristobal de, Lauren Musu, Joel McFarland, Sidney Wilkinson-Flicker, Melissa Diliberti, Anlan Zhang, Claire Bransetter, and Xiaolei Wang. 2019. "Status and Trends in the Education of Racial and Ethnic Groups 2018." National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019038.

[10] Burk, Kymyona. 2020. "The Perfect Storm: Mississippi's Momentum for Improving Reading Achievement." The Reading League Journal 1 (2): 28–34.

[11] Byrd, Robert S., and Michael L. Weitzman. 1994. "Predictors of Early Grade Retention among Children in the United States." Pediatrics 93 (3): 481–87.

[12] Corman, H. 2003. "The Effects of State Policies, Individual Characteristics, Family Characteristics, and Neighbourhood Characteristics on Grade Repetition in the United States." Economics of Education Review 22 (4): 409–20. https://doi.org/10.1016/S0272-7757(02)00070-5.

[13] Cummings, Amy, Katharine O. Strunk, and Craig De Voto. 2021. "'A Lot of States Were Doing It': The Development of Michigan's Read by Grade Three Law." Journal of Educational Change, 1–26.

[14] "Early Warning: Why Reading by the End of Third Grade Matters." 2010. Annie E. Casey Foundation. https://www.aecf.org/resources/early-warning-why-reading-by-the-end-of-third-grade-matters.

[15] Eide, Eric R., and Mark H. Showalter. 2001. "The Effect of Grade Retention on Educational and Labor Market Outcomes." Economics of Education Review 20 (6): 563–76.

[16] Entwisle, Doris R., and Karl L. Alexander. 1993. "Entry Into School: The Beginning School Transition and Educational Stratification in the United States." Annual Review of Sociology 19 (1): 401–23. https://doi.org/10.1146/annurev.so.19.080193.002153.

[17] Eren, Ozkan, Michael F. Lovenheim, and H. Naci Mocan. 2022. "The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy." Journal of Labor Economics 40 (2): 361–95. https://doi.org/10.1086/715836.

[18] Figlio, David, and Umut Özek. 2020. "An Extra Year to Learn English? Early Grade Retention and the Human Capital Development of English Learners." Journal of Public Economics 186 (June): 104184. https://doi.org/10.1016/j.jpubeco.2020.104184.

[19] Giano, Zachary, Amanda L. Williams, and Jennifer N. Becnel. 2022. "Grade Retention and School Dropout: Comparing Specific Grade Levels Across Childhood and Early Adolescence." The Journal of Early Adolescence 42 (1): 33–57. https://doi.org/10.1177/02724316211010332.

[20] Gilmour, Allison F., Justin Harper, Blair Lloyd, and Alyssa Van Camp. 2023. "Response to Intervention and Specific Learning Disability Identification: Evidence From Tennessee." Journal of Learning Disabilities, December, 00222194231215013. https://doi.org/10.1177/00222194231215013.

[21] Gottfried, Michael A. 2013. "The Spillover Effects of Grade-Retained Classmates: Evidence from Urban Elementary Schools." American Journal of Education 119 (3): 405–44. https://doi.org/10.1086/669851.

[22] Greene, Jay P., and Marcus A. Winters. 2007. "Revisiting Grade Retention: An Evaluation of Florida's Test-Based Promotion Policy." Education Finance and Policy 2 (4): 319–40. https://doi.org/10.1162/edfp.2007.2.4.319.

[23] Griffith, Catherine Aurentz, John Wills Lloyd, Kathleen Lynne Lane, and Melody Tankersley. 2010. "Grade Retention of Students During Grades K–8 Predicts Reading Achievement and Progress During Secondary Schooling." Reading & Writing Quarterly 26 (1): 51–66. https://doi.org/10.1080/10573560903396967.

[24] Holmes, C. Thomas, and Kenneth M. Matthews. 1984. "The Effects of Nonpromotion on Elementary and Junior High School Pupils: A Meta-Analysis." Review of Educational Research 54 (2): 225–36.

[25] Hong, Guanglei, and Stephen W. Raudenbush. 2005. "Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics." Educational Evaluation and Policy Analysis 27 (3): 205–24.

https://doi.org/10.3102/01623737027003205.

[26] Hong, Guanglei, and Stephen W. Raudenbush. 2006. "Evaluating Kindergarten Retention Policy." Journal of the American Statistical Association 101 (475). http://amstat.tandfonline.com/doi/abs/10.1198/016214506000000447.

[27] Hong, Guanglei, and Bing Yu. 2007. "Early-Grade Retention and Children's Reading and Math Learning in Elementary Years." Educational Evaluation and Policy Analysis 29 (4): 239–61. https://doi.org/10.3102/0162373707309073.

[28] Hong, Guanglei, and Bing Yu. 2008. "Effects of Kindergarten Retention on Children's Social-Emotional Development: An Application of Propensity Score Method to Multivariate, Multilevel Data." Developmental Psychology 44 (2): 407–21.

[29] Huddleston, Andrew P. 2014. "Achievement at Whose Expense? A Literature Review of Test-Based Grade Retention Policies in US Schools." Education Policy Analysis Archives 22: 1–31.

[30] Hughes, Jan N., Qi Chen, Felix Thoemmes, and Oi-man Kwok. 2010. "An Investigation of the Relationship Between Retention in First Grade and Performance on High Stakes Tests in Third Grade." Educational Evaluation and Policy Analysis 32 (2): 166–82. https://doi.org/10.3102/0162373710367682.

[31] Hughes, Jan N., Stephen G. West, Hanjoe Kim, and Shelby S. Bauer. 2018. "Effect of Early Grade Retention on School Completion: A Prospective Study." Journal of Educational Psychology 110 (7): 974.

[32] Hwang, NaYoung, and Cory Koedel. 2022. "Holding Back to Move Forward: The Effects of Retention in the Third Grade on Student Outcomes." Annenberg Institute at Brown University, EdWorkingPaper: 22-688. https://doi.org/10.26300/mmxx-3e82.

[33] Hwang, Sophia HJ, and Elise Cappella. 2018. "Rethinking Early Elementary Grade Retention: Examining Long-Term Academic and Psychosocial Outcomes." Journal of Research on Educational Effectiveness 11 (4): 559–87.

[34] Jimerson, Shane R. 2001. "Meta-Analysis of Grade Retention Research: Implications for Practice in the 21st Century." School Psychology Review 30 (3): 420–37.

[35] Jimerson, Shane R., Gabrielle E. Anderson, and Angela D. Whipple. 2002. "Winning the Battle and Losing the War: Examining the Relation between Grade Retention and Dropping out of High School." Psychology in the Schools 39 (4): 441–57.

[36] Kretschmann, Julia, Miriam Vock, Oliver Lüdtke, Malte Jansen, and Anna Gronostaj. 2019. "Effects of Grade Retention on Students' Motivation: A Longitudinal Study over 3 Years of Secondary School." Journal of Educational Psychology 111 (8): 1432.

[37] LiCalsi, Christina, Umut Ozek, and David Figlio. 2019. "The Uneven Implementation of Universal School Policies: Maternal Education and Florida's Mandatory Grade Retention Policy." Education Finance and Policy 14 (3): 383–413.

[38] Locke, Victoria Nevin, and P. Johnelle Sparks. 2019. "Who Gets Held Back? An Analysis of Grade Retention Using Stratified Frailty Models." Population Research and Policy Review 38 (5): 695–731. https://doi.org/10.1007/s11113-019-09524-3.

[39] Lorence, Jon, and Anthony Gary Dworkin. 2006. "Elementary Grade Retention in Texas and Reading Achievement Among Racial Groups: 1994–2002." Review of Policy Research 23 (5): 999–1033. https://doi.org/10.1111/j.1541-1338.2006.00247.x.

[40] Mariano, Louis T., Francisco (Paco) Martorell, and Tiffany Tsai. 2018. The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools. RAND Santa Monica, CA.

[41] Martorell, Paco, and Louis T. Mariano. 2018. "The Causal Effects of Grade Retention on Behavioral Outcomes." Journal of Research on Educational Effectiveness 11 (2): 192–216. https://doi.org/10.1080/19345747.2017.1390024.

[42] Mattison, Amira, Linda M. Raffaele Mendez, Robert Dedrick, Sarah Dickinson, Emily Wingate, and Camille Hanks. 2018. "Early Elementary Teacher Ratings of Behavior as Predictors of Grade Retention: Race, Gender, and Socioeconomic Status as Potential Moderators." Psychology in the Schools 55 (10): 1171–87.

[43] McCombs, Jennifer Sloan, Sheila Nataraj Kirby, and Louis T. Mariano. 2009. "Ending Social Promotion Without Leaving Children Behind." Product Page. Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs/monographs/MG894.html.

[44] McCormick, Meghan P., Robin Neuhaus, E. Parham Horn, Erin E. O'Connor, Hope I. White, Samantha Harding, Elise Cappella, and Sandee McClowry. 2019. "Long-Term Effects of Social–Emotional Learning on Receipt of Special Education and Grade Retention: Evidence From a Randomized Trial of INSIGHTS." AERA Open 5 (3): 1–21. https://doi.org/10.1177/2332858419867290.

[45] McCoy, Ann R., and Arthur J. Reynolds. 1999. "Grade Retention and School Performance:: An Extended Investigation." Journal of School Psychology 37 (3): 273–98.

[46] Miller, Luke C., and Daphna Bassok. 2019. "The Effects of Universal Preschool on Grade Retention." Education Finance and Policy 14 (2): 149–77. https://doi.org/10.1162/edfp_a_00248.

[47] Moser, Stephanie E., Stephen G. West, and Jan N. Hughes. 2012. "Trajectories of Math and Reading Achievement in Low-Achieving Children in Elementary School: Effects of Early and Later Retention in Grade." Journal of Educational Psychology 104 (3): 603–21. http://dx.doi.org.proxy.library.vanderbilt.edu/10.1037/a0027571.

[48] Murray, Christy S., Althea L. Woodruff, and Sharon Vaughn. 2010. "First-Grade Student Retention Within a 3-Tier Reading Framework." Reading & Writing Quarterly 26 (1): 26–50. https://doi.org/10.1080/10573560903396934.

[49] Nickow, Andre, Philip Oreopoulos, and Vincent Quan. 2020. "The Impressive Effects of Tutoring on PreK-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence." Working Paper. Working

Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27476.

[50] Owings, William A., and Susan Magliaro. 1998. “Grade Retention: A History of Failure.” Educational Leadership 56 (1): 86–88.

[51] Özek, Umut. 2015. “Hold Back to Move Forward? Early Grade Retention and Student Misbehavior.” Education Finance and Policy 10 (3): 350–77.

[52] Penfield, Randall D. 2010. “Test-Based Grade Retention Does It Stand Up to Professional Standards for Fair and Appropriate Test Use?” Educational Researcher 39 (2): 110–19. https://doi.org/10.3102/0013189X10363007.

[53] Range, Bret G., John Pijanowski, Carleton R. Holt, and Suzie Young. 2012. “The Perceptions of Primary Grade Teachers and Elementary Principals about the Effectiveness of Grade-Level Retention.” Professional Educator 36 (1). http://eric.ed.gov/?id=EJ989520.

[54] Reschly, Amy L. 2010. “Reading and School Completion: Critical Connections and Matthew Effects.” Reading & Writing Quarterly 26 (1): 67–90. https://doi.org/10.1080/10573560903397023.

[55] Reynolds, Arthur J. 1992. “Grade Retention and School Adjustment: An Explanatory Analysis.” Educational Evaluation and Policy Analysis 14 (2): 101–21.

[56] Rhodes, Judith, Johanna M. Thomas, and Alesa R. Liles. 2018. “Predictors of Grade Retention among Children in an Elementary School Truancy Intervention.” Journal of At-Risk Issues 21 (1): 1–10.

[57] Roderick, Melissa. 1994. “Grade Retention and School Dropout: Investigating the Association.” American Educational Research Journal 31 (4): 729–59. https://doi.org/10.2307/1163393.

[58] Roderick, Melissa, and Jenny Nagaoka. 2005. “Retention under Chicago’s High-Stakes Testing Program: Helpful, Harmful, or Harmless?” Educational Evaluation and Policy Analysis 27 (4): 309–40.

[59] Schwerdt, Guido, Martin R. West, and Marcus A. Winters. 2017. “The Effects of Test-Based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida.” Journal of Public Economics 152 (August): 154–69. https://doi.org/10.1016/j.jpubeco.2017.06.004.

[60] Shepard, Lorrie A., and Mary Lee Smith. 1990. “Synthesis of Research on Grade Retention.” Educational Leadership 47 (8): 84–88.

[61] Slungaard Mumma, Kirsten, and Marcus A. Winter. 2023. “The Effect of Retention Under Mississippi’s Test-Based Promotion Policy.” Boston University Wheelock Educational Policy Center, Working Paper 2023–1. https://wheelockpolicycenter.org/high-quality-education/ms-read-by-grade-three/.

[62] “State K-3 Policies: Is There a Third Grade Retention Policy?” 2020. Education Commission of the States. https://reports.ecs.org/comparisons/state-k-3-policies-19.

[63] Stearns, Elizabeth, Stephanie Moller, Judith Blau, and Stephanie Potochnick. 2007. “Staying Back and Dropping Out: The Relationship Between Grade Retention and School Dropout.” Sociology of Education 80 (3): 210–40. https://doi.org/10.1177/003804070708000302.

[64] Tennessee Department of Education. 2015. “Annual Statistic Report of the Department of Education.” Nashville, TN: Tennessee Department of Education. https://www.tn.gov/assets/entities/education/attachments/asr_1415.pdf.

[65] Tennessee Department of Education. 2017. “Building the Framework: A Report on Elementary Grades Reading in Tennessee.” https://www.tn.gov/content/dam/tn/education/reports/rpt_bldg_the_framework.pdf.

[66] Tennessee Department of Education. 2023. “Reading 360.” https://www.tn.gov/education/reading-360.html.

[67] “Third Grade Reading Laws: Implementation and Impact.” Council of Chief State School Officers. Accessed October 4, 2022. https://ccsso.org/resource-library/third-grade-reading-laws-implementation-and-impact.

[68] Valbuena, Javier, Mauro Mediavilla, Álvaro Choi, and María Gil. 2021. “Effects of Grade Retention Policies: A Literature Review of Empirical Studies Applying Causal Inference.” Journal of Economic Surveys 35 (2): 408–51.

[69] Warren, John Robert, Emily Hoffman, and Megan Andrew. 2014. “Patterns and Trends in Grade Retention Rates in the United States, 1995–2010.” Educational Researcher 43 (9): 433–43. https://doi.org/10.3102/0013189X14563599.

[70] Warren, John Robert, and Jim Saliba. 2012. “First- Through Eighth-Grade Retention Rates for All 50 States A New Method and Initial Results.” Educational Researcher 41 (8): 320–29. https://doi.org/10.3102/0013189X12457813.

[71] West, Martin R. 2012. “Is Retaining Students in the Early Grades Self-Defeating?” Brookings Institute: Washington, DC. http://www.brookings.edu/research/papers/2012/08/16-student-retention-west.

[72] Westall, John, and Amy Cummings. 2023. “The Effects of Early Literacy Policies on Student Achievement.” Preprint. SSRN. https://doi.org/10.2139/ssrn.4427675.

[73] Westall, John, Andrew Utter, Tara Kilbride, and Katharine O. Strunk. 2022. “Read by Grade Three Law Initial Retention Decisions for the 2021-22 School Year.” Education Policy Innovation Collaborative, Michigan State University. https://epicedpolicy.org/rbg3-retention-decisions-2022-23/.

[74] Winters, Marcus A., and Jay P. Greene. 2012. “The Medium-Run Effects of Florida’s Test-Based Promotion Policy.” Education Finance and Policy 7 (3): 305–30. https://doi.org/10.1162/EDFP_a_00069.

[75] Witmer, Stacie M., Lynn M. Hoffman, and Katharyn E. Nottis. 2004. “Elementary Teachers’ Beliefs and Knowledge About Grade Retention: How Do We Know What They Know?” Education 125 (2): 173–93.

[76] Wu, Wei, Stephen G. West, and Jan N. Hughes. 2010. “Effect of Grade Retention in First Grade on Psychosocial Outcomes.” Journal of Educational Psychology 102 (1): 135.

[77] Xia, Nailing, and Sheila Nataraj Kirby. 2009. “Retaining Students in Grade.” Santa Monica, CA: Rand Corporation. http://www.rand.org/pubs/technical_reports/TR678.html.

[78] Yang, Mi-Youn, Zibei Chen, Judith L. F. Rhodes, and Marmar Orooji. 2018. “A Longitudinal Study on Risk Factors of Grade Retention among Elementary School Students Using a Multilevel Analysis: Focusing on

Material Hardship and Lack of School Engagement." Children and Youth Services Review 88 (May): 25–32. https://doi.org/10.1016/j.childyouth.2018.02.043.

Table 2.1 Descriptive Statistics for Analytic Sample

	Percent
Student's third-grade characteristics	
Female	48%
Asian	2%
Black	24%
Hispanic	9%
Native American	<1%
White	65%
Economically Disadvantaged	60%
English Learner	6%
Student with Disability	14%
Scoring at or above proficient in reading	45%
Total student observations	350,518

Note: Observations are pooled across the five cohorts in the analytic sample (students starting third grade for the first time in 2009-10 through 2013-14). Students were classified as economically disadvantaged in Tennessee's state data system if they qualified for free or reduced-price (FRP) school meals. We defined English learners as students identified as actively receiving English as a Second Language (ESL) instruction. Students with disabilities include those identified as having any documented disability while excluding those identified as gifted and talented students.

Table 2.2 Retention Rates for Analytic Sample

	Retention rate K-2	Retention rate 3rd
By gender		
Female	8.2%	0.8%
Male	12.6%	1.0%
By race/ethnicity		
Asian/Pacific Islander	4.1%	0.4%
Black	10.1%	1.2%
Hispanic	10.1%	0.7%
Native American	11.0%	0.5%
White	10.8%	0.8%
Economic disadvantage		
Econ. disadvantaged	13.4%	1.2%
Not disadvantaged	6.1%	0.5%
Disability identification		
Student with disability	24.9%	1.2%
Student without disability	7.8%	0.9%
English learner status		
English learner	11.6%	0.9%
Not an English learner	10.4%	0.9%
Third-grade reading proficiency		
Not proficient	13.5%	1.4%
Proficient	4.4%	0.0%
All students	10.5%	0.9%

Note: Retention rates are averaged across the five cohorts in the analytic sample (students starting third grade for the first time in 2009-10 through 2013-14). The K-2 retention rates represents the percent of students retained at least once before entering third grade.

Table 2.3 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Attendance

	Effect on Annual Attendance Rate (0-100)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
Panel A. Base Specification	-0.0548	-0.0646	0.0061	0.1055	0.2028*
	(-0.0415)	(0.0419)	(0.0539)	(0.0646)	(0.0734)
Sample Mean	95.63	95.62	95.24	94.83	94.69
N	342272	330767	320074	313186	303916
Panel B. + Student Covariates	-0.0484	-0.0648	0.0099	0.1082	0.2098*
	(0.0415)	(0.0418)	(0.0528)	(0.0623)	(0.0712)
Sample Mean	95.63	95.62	95.24	94.83	94.69
N	339688	328277	317623	310755	301538
Panel C. +Student Covariates + School FEs	-0.0075	-0.0361	0.0338	0.1072	0.2409**
	(0.036)	(0.0382)	(0.0468)	(0.0583)	(0.0662)
Sample Mean	95.63	95.62	95.24	94.83	94.69
N	339657	328235	317524	310708	301501

Notes: The dependent variable is the student's annual attendance rate (0-100) in their 4^{th}-8^{th} grade year. Standard errors are clustered at the school level. Panel A presents results from a basic difference-in-differences specification excluding student covariates and school fixed effects, Panel B presents results from the specification which includes student covariates, and Panel C presents results from the specification which includes both student covariates and school fixed effects. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.4 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Discipline

	Effect on Likelihood of Having a Disciplinary Record (0-1)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
Panel A. Base Specification	-0.004	-0.0082*	-0.0263***	-0.0178***	-0.0129**
	(0.0027)	(0.0032)	(0.0041)	(0.0042)	(0.004)
Sample Mean	0.06	0.09	0.16	0.17	0.17
N	342799	331154	320588	313629	304500
Panel B. + Student Covariates	-0.0037	-0.0076*	-0.0241***	-0.0160***	-0.0125**
	(0.0024)	(0.0028)	(0.0036)	(0.0037)	(0.0037)
Sample Mean	0.06	0.09	0.16	0.17	0.17
N	340215	328664	318137	311198	302122
Panel C. +Student Covariates + School FEs	-0.0039	-0.0087**	-0.0249***	-0.0165***	-0.0140***
	(0.0019)	(0.0024)	(0.0032)	(0.0034)	(0.0035)
Sample Mean	0.06	0.09	0.16	0.17	0.17
N	340184	328621	318037	311151	302085

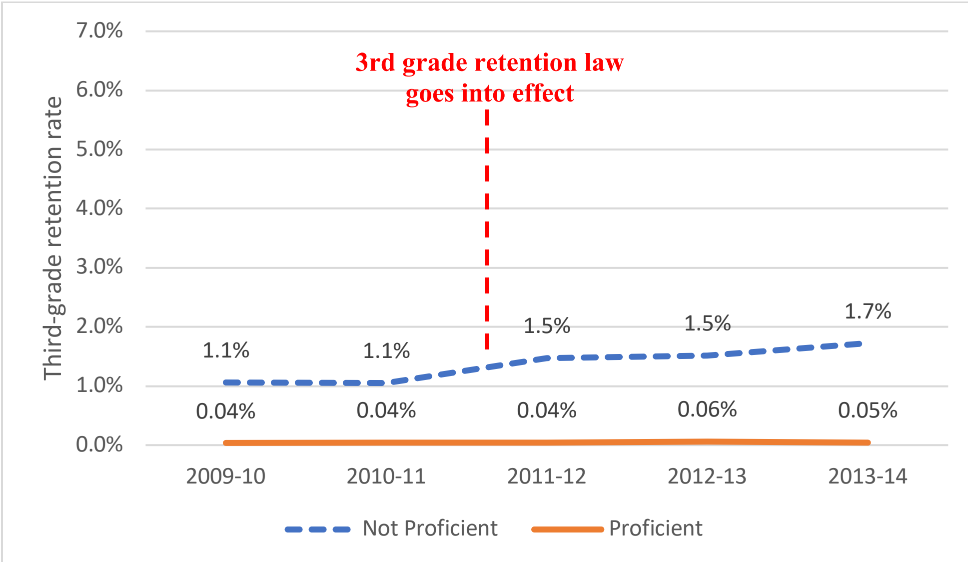
Notes: The dependent variable is the student's annual likelihood of having a disciplinary record in their 4^{th}-8^{th} grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. Standard errors are clustered at the school level. Panel A presents results from a basic difference-in-differences specification excluding student covariates and school fixed effects, Panel B presents results from the specification which includes student covariates, and Panel C presents results from the specification which includes both student covariates and school fixed effects. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.5 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Standardized Reading Achievement

	Effect on Standardized Reading Score				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
Panel A. Base Specification	0.0033	0.0229*	0.0586***	0.0434***	0.0415***
	(0.0081)	(0.0095)	(0.0089)	(0.0098)	(0.0096)
N	325961	250571	244959	240981	233102
Panel B. + Student Covariates	0.0218**	0.0425***	0.0708***	0.0641***	0.0515***
	(0.007)	(0.0081)	(0.0076)	(0.0084)	(0.0087)
N	323511	248754	243098	239426	231127
Panel C. +Student Covariates + School FEs	0.0227***	0.0475***	0.0722***	0.0695***	0.0570***
	(0.0061)	(0.0069)	(0.0065)	(0.0075)	(0.0081)
N	323490	248720	243026	239393	231107

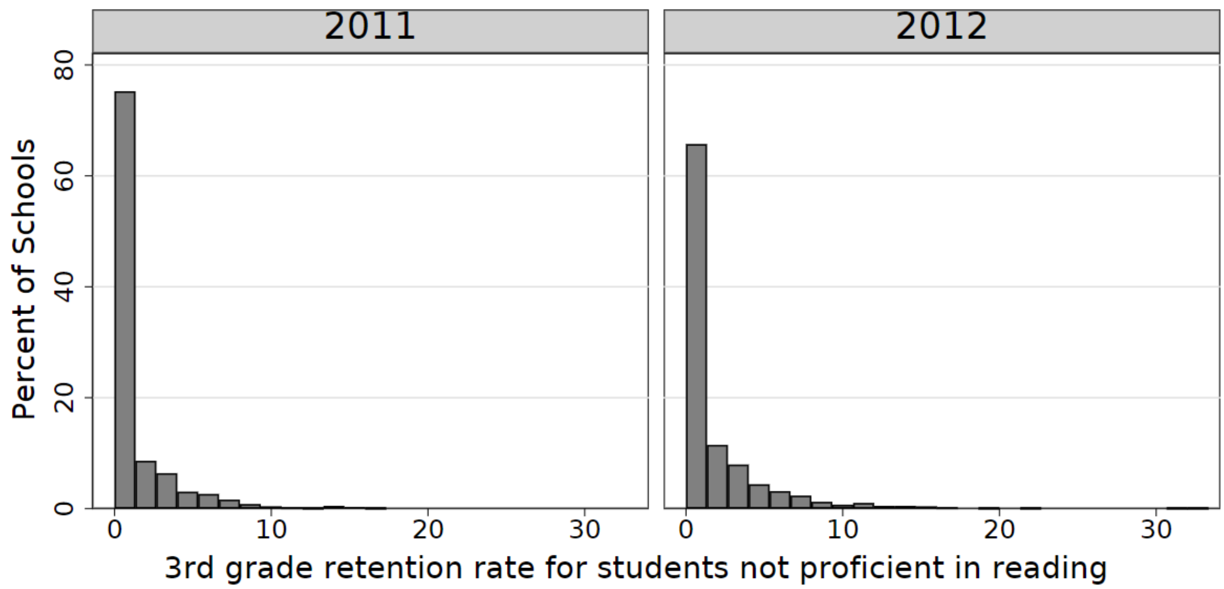
Notes: The dependent variable is the student's standardized reading assessment score in their 4^{th}-8^{th} grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Standard errors are clustered at the school level. Panel A presents results from a basic difference-in-differences specification excluding student covariates and school fixed effects, Panel B presents results from the specification which includes student covariates, and Panel C presents results from the specification which includes both student covariates and school fixed effects. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2.1 Third-grade retention rate by reading proficiency status



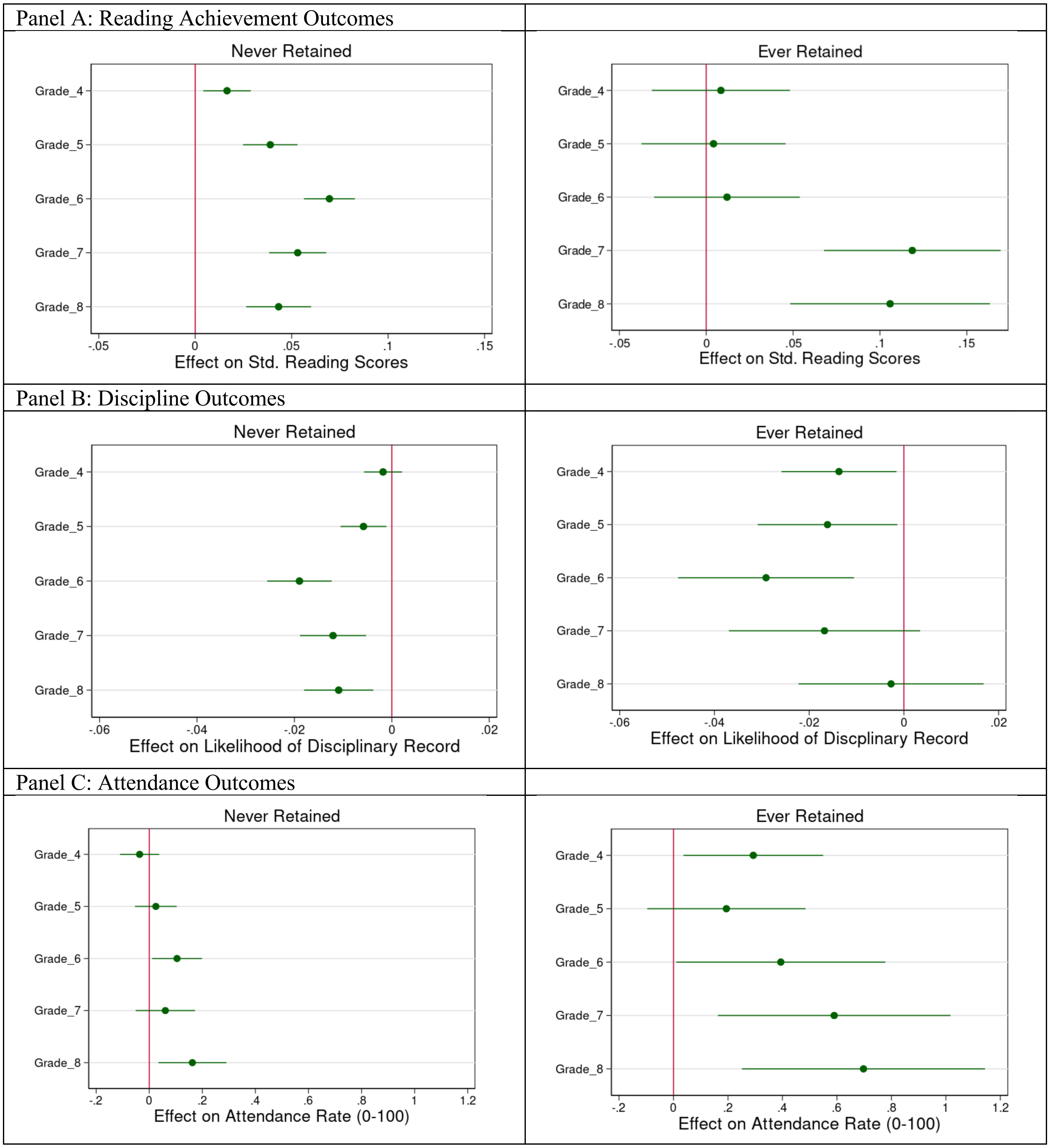
Note: Reading proficiency determined by student scores on third-grade English Language Arts standardized exam (TCAP).

Figure 2.2 Histogram of third-grade retention rates by school



Note: These histograms illustrate the distribution of third-grade retention rates at all public schools in Tennessee for third-grade students who scored below proficient on the third-grade English Language Arts standardized test (TCAP).

Figure 2.3 Heterogeneous Effects by Retention History



Notes: These plots show the coefficients estimated from the main specification for students who never experienced retention before completing third grade (“never retained) and for students who were retained at least once between Kindergarten and third grade (“ever retained”).

2.8 Appendix Tables & Figures

Table 2.6 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Attendance

	Effect on Annual Attendance Rate (0-100)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Economically Disadvantaged	-0.0075	0.0347	0.1777**	0.3562***	0.4958***
	(0.050)	(0.054)	(0.062)	(0.069)	(0.082)
Sample Mean	95.15	95.12	94.68	94.16	93.97
N	198676	193064	188678	184735	178237
B. Non Economically Disadvantaged	0.0335	0.1334*	0.1484*	-0.0199	0.0316
	(0.049)	(0.054)	(0.059)	(0.060)	(0.063)
Sample Mean	96.56	96.58	96.31	96.05	95.98
N	140980	135183	128916	125989	123269
C. Male	0.0367	-0.004	0.1079	0.1942**	0.3288***
	(0.049)	(0.053)	(0.064)	(0.068)	(0.075)
Sample Mean	95.69	95.62	95.2	94.83	94.78
N	171554	165920	160746	157066	152047
D. Female	-0.0802	-0.0151	0.0196	-0.0101	0.1780*
	(0.049)	(0.048)	(0.056)	(0.061)	(0.071)
Sample Mean	95.78	95.82	95.49	95.02	94.8
N	168125	162349	156866	153681	149480
E. Ever Retained	0.2792*	0.2025	0.2718	0.5196*	0.7957***
	(0.132)	(0.159)	(0.174)	(0.214)	(0.235)
Sample Mean	94.72	94.62	94.04	93.35	93.17
N	32992	32578	32079	31440	29769
F. Never Retained	-0.0845*	-0.0619	0.0149	0.0127	0.1777**
	(0.038)	(0.041)	(0.045)	(0.048)	(0.055)
Sample Mean	95.84	95.84	95.49	95.1	94.97
N	306654	295659	285504	279271	271718
G. Special Education	0.1368	0.0562	0.2105	-0.0137	-0.0802
	(0.107)	(0.118)	(0.155)	(0.163)	(0.202)
Sample Mean	95.14	95.09	94.54	94.14	94.02
N	37958	36656	35754	34919	33740
H. Non Special Education	-0.0586	-0.0384	0.0575	0.1115*	0.2908***
	(0.038)	(0.041)	(0.045)	(0.050)	(0.056)
Sample Mean	95.81	95.8	95.44	95.02	94.89
N	301699	291592	281838	275803	267766

Notes: The dependent variable is the student's annual attendance rate (0-100) in their 4^{th}-8^{th} grade year. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e. if a student is labeled as receiving special education services in their first third-grade year, they will be included in the sample for Panel G). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Discipline

	Effect on Likelihood of Having a Disciplinary Record (0-1)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Economically Disadvantaged	-0.0055*	-0.0088**	-0.0349***	-0.0231***	-0.0219***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Sample Mean	0.08	0.12	0.21	0.23	0.22
N	198940	193255	188953	184967	178563
B. Non Economically Disadvantaged	-0.0041*	-0.0094***	-0.0122***	-0.0110**	-0.0061
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)
Sample Mean	0.02	0.03	0.07	0.08	0.08
N	141245	135379	129155	126200	123528
C. Male	-0.005	-0.0080*	-0.0327***	-0.0190***	-0.0186***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Sample Mean	0.08	0.12	0.21	0.23	0.21
N	171839	166100	161012	157305	152346
D. Female	-0.0024	-0.0081**	-0.0161***	-0.0143***	-0.0114***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Sample Mean	0.03	0.05	0.1	0.11	0.11
N	168367	162556	157114	153885	149765
E. Ever Retained	-0.0137*	-0.0175*	-0.0377***	-0.0168	-0.0053
	(0.007)	(0.008)	(0.009)	(0.010)	(0.010)
Sample Mean	0.09	0.14	0.24	0.26	0.24
N	33034	32608	32108	31473	29824
F. Never Retained	-0.0014	-0.0062**	-0.0216***	-0.0147***	-0.0142***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Sample Mean	0.05	0.08	0.14	0.16	0.15
N	307139	296016	285989	279681	272247
G. Special Education	-0.0051	-0.0073	-0.0238**	0.0026	-0.0164
	(0.005)	(0.006)	(0.009)	(0.009)	(0.009)
Sample Mean	0.07	0.11	0.2	0.22	0.21
N	38043	36708	35798	34987	33818
H. Non Special Education	-0.0034	-0.0087***	-0.0255***	-0.0191***	-0.0151***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Sample Mean	0.05	0.08	0.15	0.16	0.16
N	302141	291927	282308	276178	268273

Notes: The dependent variable is the student's annual likelihood of having a disciplinary record in their 4^{th}-8^{th} grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e. if a student is labeled as receiving special education services in their first third-grade year, they will be included in the sample for Panel G). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Standardized Reading Achievement

	Effect on Standardized Reading Score				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Economically Disadvantaged	0.0293***	0.0456***	0.0872***	0.0949***	0.0715***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)
Sample Mean	-0.28	-0.27	-0.25	-0.23	-0.22
N	187092	142538	142263	140170	136174
B. Non Economically Disadvantaged	0.0244**	0.0497***	0.0633***	0.0288**	0.0566***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.011)
Sample Mean	0.44	0.44	0.45	0.45	0.46
N	136385	106196	100790	99215	94914
C. Male	0.0195*	0.0389***	0.0903***	0.1053***	0.0760***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)
Sample Mean	-0.07	-0.08	-0.08	-0.08	-0.07
N	162232	124755	122168	120607	116235
D. Female	0.0203**	0.0540***	0.0577***	0.0366***	0.0433***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)
Sample Mean	0.13	0.15	0.17	0.19	0.2
N	161269	123992	120920	118810	114885
E. Ever Retained	0.0171	0.0221	0.018	0.1283***	0.1055***
	(0.020)	(0.021)	(0.022)	(0.027)	(0.031)
Sample Mean	-0.55	-0.57	-0.55	-0.54	-0.53
N	29300	22938	23395	23134	21766
F. Never Retained	0.0175**	0.0457***	0.0721***	0.0553***	0.0446***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.008)
Sample Mean	0.08	0.1	0.11	0.12	0.12
N	294155	225791	219658	216242	209301
G. Special Education	0.034	-0.0524*	0.0186	0.0765**	0.1521***
	(0.020)	(0.021)	(0.023)	(0.025)	(0.028)
Sample Mean	-0.45	-0.43	-0.49	-0.53	-0.54
N	31187	22970	24521	25758	25580
H. Non Special Education	0.0287***	0.0658***	0.0917***	0.0795***	0.0410***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.008)
Sample Mean	0.08	0.08	0.1	0.12	0.14
N	292288	225764	218539	213628	205512

Notes: The dependent variable is the student's standardized reading assessment score in their 4th-8th grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e. if a student is labeled as receiving special education services in their first third-grade year, they will be included in the sample for Panel G). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.9 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Attendance by Race

	Effect on Annual Attendance Rate (0-100)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. White	0.0279	0.0897*	0.1456**	0.0465	0.1650*
	(0.04)	(0.04)	(0.05)	(0.06)	(0.07)
Sample Mean	95.59	95.56	95.15	94.76	94.58
N	226045	218225	210462	206192	200846
B. Black	0.015	0.0927	0.4503***	0.5383***	0.7153***
	(0.08)	(0.08)	(0.11)	(0.14)	(0.17)
Sample Mean	95.34	95.35	95.01	94.51	94.5
N	78978	76873	75267	73543	70484
C. Hispanic	0.006	0.1273	-0.0231	0.1759	-0.1213
	(0.12)	(0.13)	(0.13)	(0.14)	(0.16)
Sample Mean	96.35	96.38	96.12	95.71	95.56
N	26274	25396	24577	24071	23464
D. Asian/Pacific Islander	0.1929	0.1357	-0.0376	0.0682	-0.4592
	(0.27)	(0.19)	(0.30)	(0.22)	(0.25)
Sample Mean	97.31	97.48	97.43	97.3	97.25
N	7077	6536	6146	5908	5750
E. Native American	-0.1157	-1.0978	-1.3978	-0.6299	1.093
	(1.15)	(1.18)	(1.18)	(1.16)	(1.62)
Sample Mean	94.62	94.55	94.19	93.71	93.51
N	640	566	638	634	592

Notes: The dependent variable is the student's annual attendance rate (0-100) in their 4th-8th grade year. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e., if a student is labeled as White in their first third-grade year, they will be included in the sample for Panel A). Standard errors are in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001

Table 2.10 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Discipline by Race

	Effect on Likelihood of Having a Disciplinary Record (0-1)				
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. White	-0.0060***	-0.0064**	-0.0135***	-0.0029	-0.0009
	(0.0018)	(0.0022)	(0.0031)	(0.0035)	(0.0036)
Sample Mean	0.03	0.05	0.11	0.12	0.12
N	226430	218505	210818	206498	201304
B. Black	-0.0024	-0.0051	-0.0413***	-0.0358***	-0.0321***
	(0.0056)	(0.0071)	(0.0074)	(0.0080)	(0.0077)
Sample Mean	0.14	0.2	0.32	0.34	0.31
N	79079	76947	75385	73650	70587
C. Hispanic	-0.0045	-0.0183*	-0.0266**	-0.0186	-0.0001
	(0.0038)	(0.0083)	(0.0087)	(0.0100)	(0.0101)
Sample Mean	0.03	0.06	0.13	0.14	0.14
N	26304	25418	24607	24095	23479
D. Asian/Pacific Islander	-0.0054	-0.0054	-0.0104	0.0176	0.0142
	(0.0065)	(0.0129)	(0.0147)	(0.0158)	(0.0137)
Sample Mean	0.01	0.03	0.05	0.06	0.05
N	7087	6546	6155	5914	5759
E. Native American	-0.0388	-0.0828	-0.125	-0.0685	-0.0822
	(0.0418)	(0.0635)	(0.0768)	(0.0913)	(0.0698)
Sample Mean	0.06	0.09	0.17	0.18	0.18
N	641	566	638	634	592

Notes: The dependent variable is the student's annual likelihood of having a disciplinary record in their 4th-8th grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e. if a student is labeled as White in their first third-grade year, they will be included in the sample for Panel A). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.11 Heterogeneous Difference in Differences Estimation of Third-grade Retention Law on Subsequent Standardized Reading Achievement by Race

	Effect on Standardized Reading Score				
	(1) 4^{th} Grade	(2) 5^{th} Grade	(3) 6^{th} Grade	(4) 7^{th} Grade	(5) 8^{th} Grade
A. White	0.0219^{**}	0.0347^{***}	0.0631^{***}	0.0647^{***}	0.0648^{***}
	(0.0069)	(0.0074)	(0.0076)	(0.0081)	(0.0087)
Sample Mean	0.18	0.17	0.18	0.17	0.18
N	216393	167277	161943	159872	154299
B. Black	0.012	0.0475^{**}	0.0967^{***}	0.0928^{***}	0.0712^{***}
	(0.0129)	(0.0152)	(0.0140)	(0.0155)	(0.0173)
Sample Mean	-0.46	-0.45	-0.47	-0.44	-0.44
N	74205	57344	56976	55420	53085
C. Hispanic	0.0291	0.0367	0.0480^{*}	0.0690^{**}	0.0660^{**}
	(0.0205)	(0.0225)	(0.0195)	(0.0219)	(0.0253)
Sample Mean	-0.3	-0.25	-0.25	-0.2	-0.17
N	24924	18202	18613	18692	18429
D. Asian/Pacific Islander	0.0279	0.1161^{*}	0.0755	-0.017	0.042
	(0.0405)	(0.0473)	(0.0452)	(0.0459)	(0.0544)
Sample Mean	0.49	0.54	0.56	0.61	0.65
N	6759	4897	4657	4591	4529
E. Native American	0.0394	0.2543	0.1013	0.0637	0.261
	(0.1464)	(0.2089)	(0.1645)	(0.1854)	(0.1630)
Sample Mean	-0.02	-0.04	-0.04	-0.02	0.01
N	581	372	427	455	406

Notes: The dependent variable is the student's standardized reading assessment score in their 4^{th}-8^{th} grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Standard errors are clustered at the school level. This table presents results from the difference-in-differences specification which includes both student covariates and school fixed effects in every panel. Student characteristics are calculated as of a students' third-grade year (i.e. if a student is labeled as White in their first third-grade year, they will be included in the sample for Panel A). Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.12 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Outcomes Restricted to 2011/2012 Cohorts

	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Attendance					
Effect on Annual Attendance Rate (0-100)	0.1068*	0.0152	0.2662***	0.1903**	0.5725***
	(0.0471)	(0.0484)	(0.0552)	(0.0677)	(0.0839)
N	113801	113689	113569	113616	113587
B. Discipline					
Effect on Likelihood of Having a Disciplinary Record (0-1)	-0.0025	-0.0075*	-0.0251***	-0.0160***	-0.0219***
	(0.0028)	(0.0032)	(0.0039)	(0.0041)	(0.0042)
N	113801	113700	113842	113836	113702
C. Reading Achievement					
Effect on Standardized Reading Score	0.0096	0.0357***	0.0674***		
	(0.0087)	(0.0089)	(0.0086)		
N	110520	110727	110655		

Notes: This table presents results from a difference-in-differences specification which includes both student covariates and school fixed effects for the subset of students that were in their first third-grade year in SY2010-11 and SY2011-12. The dependent variable in Panel A is the student's annual attendance rate (0-100) in their 4^{th}-8^{th} grade year. The dependent variable in Panel B is the student's annual likelihood of having a disciplinary record in their 4^{th}-8^{th} grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. The dependent variable in Panel C is the student's standardized reading assessment score in their 4^{th}-8^{th} grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Due to the lack of testing administered in the state of Tennessee in SY2015-16, we are unable to calculate results for this subset of students for 7^{th} and 8^{th} grade outcomes due to the lack of comparison group. Standard errors are clustered at the school level. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.13 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Outcomes Restricted to Students that appear in the sample for all grades

	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Attendance					
Effect on Annual Attendance Rate (0-100)	0.0077	-0.0382	-0.0136	0.0698	0.2342***
	(0.0319)	(0.0337)	(0.0366)	(0.0407)	(0.0536)
N	289750	289614	289792	289841	289433
B. Discipline					
Effect on Likelihood of Having a Disciplinary Record (0-1)	-0.0034	-0.0089***	-0.0208***	-0.0156***	-0.0135***
	(0.0019)	(0.0023)	(0.0027)	(0.0028)	(0.0029)
N	290107	289883	290213	290204	289988
C. Reading Achievement					
Effect on Standardized Reading Score	0.0258***	0.0423***	0.0706***	0.0639***	0.0550***
	(0.0062)	(0.0066)	(0.0061)	(0.0067)	(0.0075)
N	281949	224017	225203	227039	222986

Notes: This table presents results from a difference-in-differences specification which includes both student covariates and school fixed effects for the subset of students that have records in Tennessee schools for all grades 3-8. The dependent variable in Panel A is the student's annual attendance rate (0-100) in their 4^{th}-8^{th} grade year. The dependent variable in Panel B is the student's annual likelihood of having a disciplinary record in their 4^{th}-8^{th} grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. The dependent variable in Panel C is the student's standardized reading assessment score in their 4^{th}-8^{th} grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Standard errors are clustered at the school level. Standard errors are clustered at the school level. Standard errors are in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

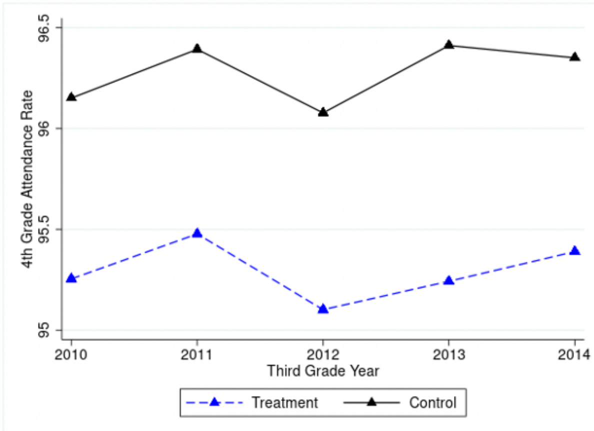
Table 2.14 Difference in Differences Estimation of Third-grade Retention Law on Subsequent Outcomes Restricted to Students that scored just above (Level 3) and just below (Level 2) the proficiency cutoff

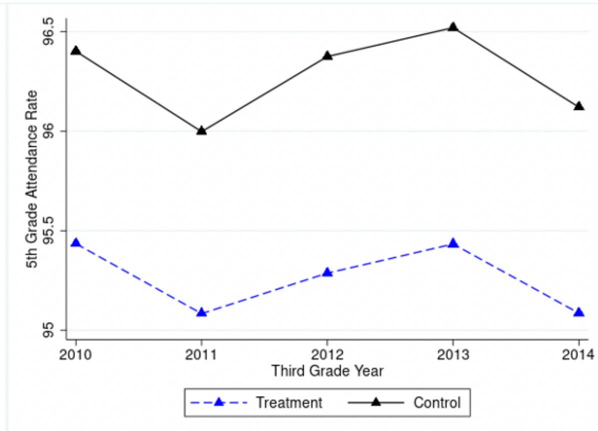
	(1) 4th Grade	(2) 5th Grade	(3) 6th Grade	(4) 7th Grade	(5) 8th Grade
A. Attendance					
Effect on Annual Attendance Rate (0-100)	-0.0101	0.0201	0.0511	0.0415	0.1939***
	(0.0377)	(0.0400)	(0.0459)	(0.0489)	(0.0564)
N	263868	255485	247143	242226	235907
B. Discipline					
Effect on Likelihood of Having a Disciplinary Record (0-1)	-0.0024	-0.0044	-0.0163***	-0.0073*	-0.0066*
	(0.0018)	(0.0023)	(0.0029)	(0.0029)	(0.0029)
N	264290	255799	247565	242580	236360
C. Reading Achievement					
Effect on Standardized Reading Score	-0.0058	0.008	0.0418***	0.0397***	0.0094
	(0.0052)	(0.0062)	(0.0058)	(0.0062)	(0.0075)
N	253721	196305	190726	187354	181958

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

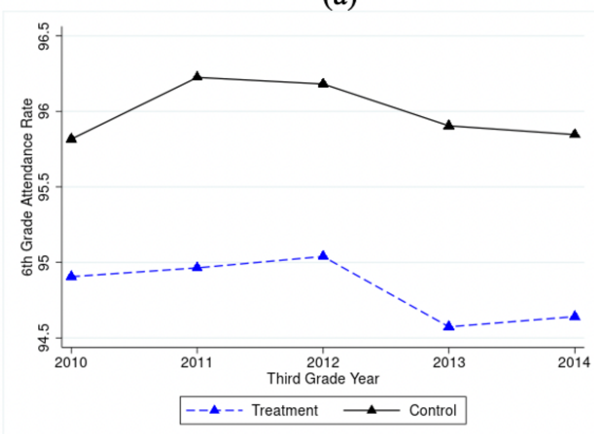
Notes: This table presents results from a difference-in-differences specification which includes both student covariates and school fixed effects for the subset of students that scored just above the proficiency cutoff (Level 3) and just below the proficiency cutoff (Level 2) on their third-grade reading examinations. The dependent variable in Panel A is the student's annual attendance rate (0-100) in their 4^{th}-8^{th} grade year. The dependent variable in Panel B is the student's annual likelihood of having a disciplinary record in their 4^{th}-8^{th} grade year. Students are counted as having a disciplinary record in this setting if they receive at least one in-school or out-of-school suspension in a given school year. The dependent variable in Panel C is the student's standardized reading assessment score in their 4^{th}-8^{th} grade year. This score is standardized within grade-year for the full sample of Tennessee students such that within each grade-year the mean of our reading measure is 0 with a standard deviation of 1. Standard errors are clustered at the school level.

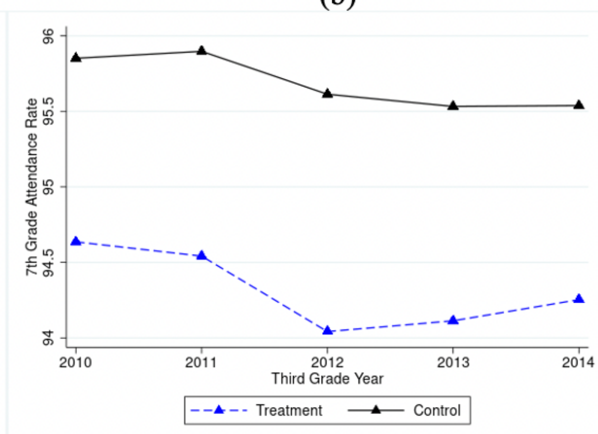
Figure 2.4 Time Trends in Attendance Rates for Third-grade Students scoring below proficient (treatment) compared to Third-grade Students scoring at or above proficient (control)



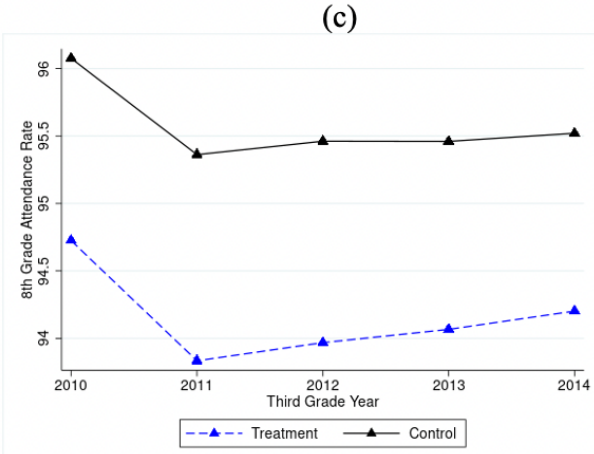


(a) (b)





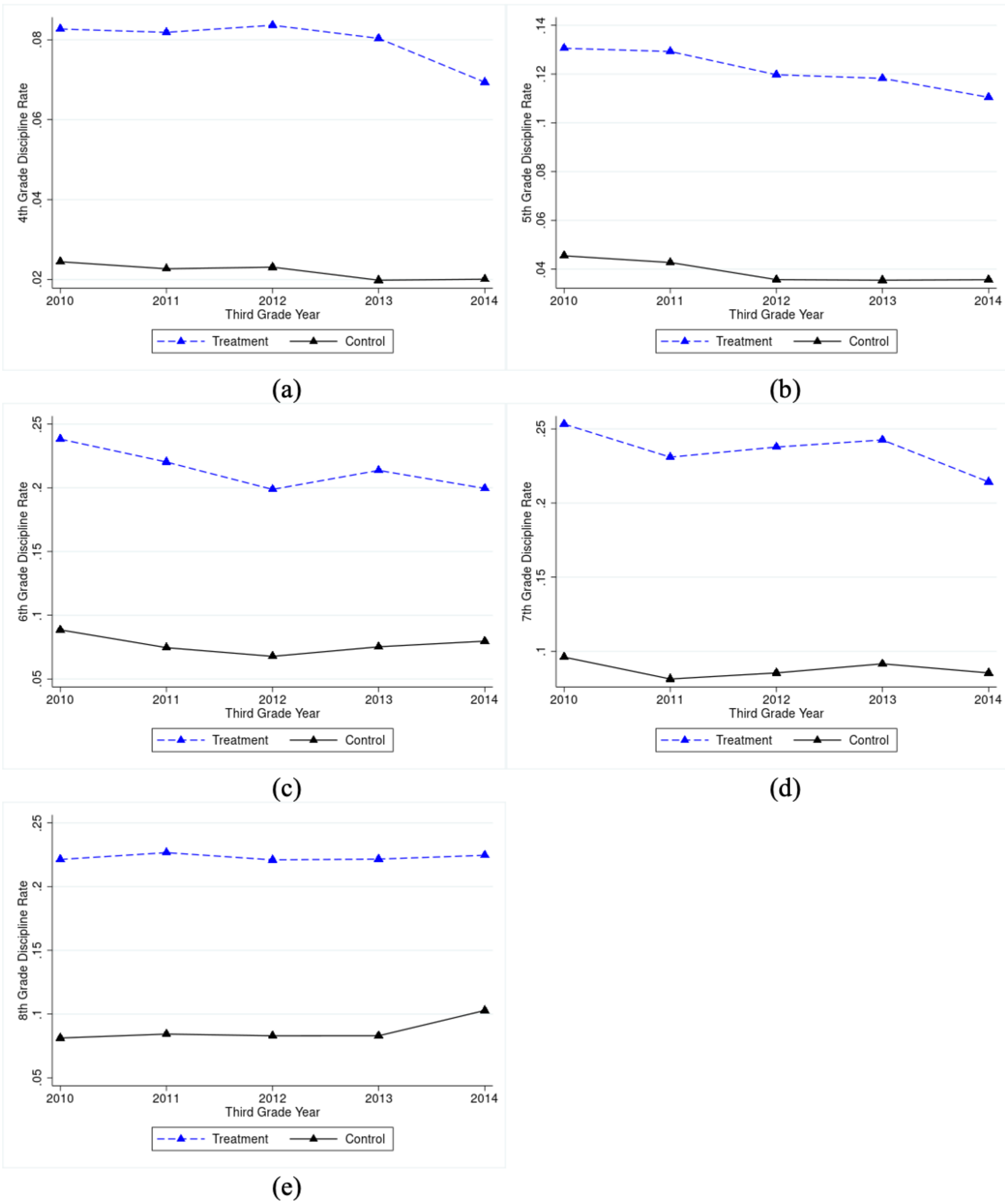
(c) (d)



(e)

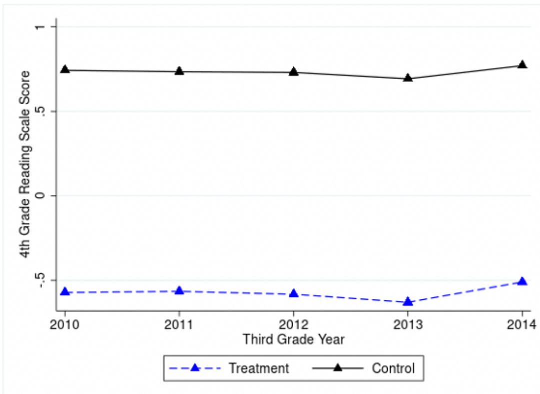
Notes: This figure shows the annual mean attendance rates in 4^{th}-8^{th} grades for students in the treatment and control groups by third-grade cohort year. Panel (a) plots the annual mean attendance rates in 4^{th} grade for students in third-grade from SY2009-10 through SY2013-14, Panels (b)-(e) plot this for grades 5-8 respectively. Students are classified as in the treatment group if their third-grade standardized reading score is below proficient, while students are in the control group if their third-grade standardized reading score is proficient or above.

Figure 2.5 Time Trends in Discipline Rates for Third-grade Students scoring below proficient (treatment) compared to Third-grade Students scoring at or above proficient (control)

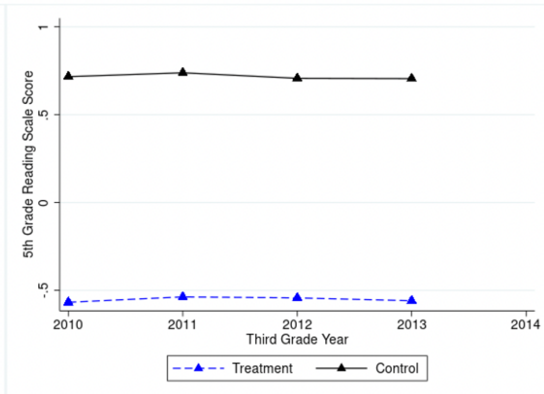


Notes: This figure shows the annual mean discipline rates in 4th-8th grades for students in the treatment and control groups by third-grade cohort year. Panel (a) plots the annual mean discipline rates in 4th grade for students in third-grade from SY2009-10 through SY2013-14, Panels (b)-(e) plot this for grades 5-8 respectively. Students are classified as in the treatment group if their third-grade standardized reading score is below proficient, while students are in the control group if their third-grade standardized reading score is proficient or above.

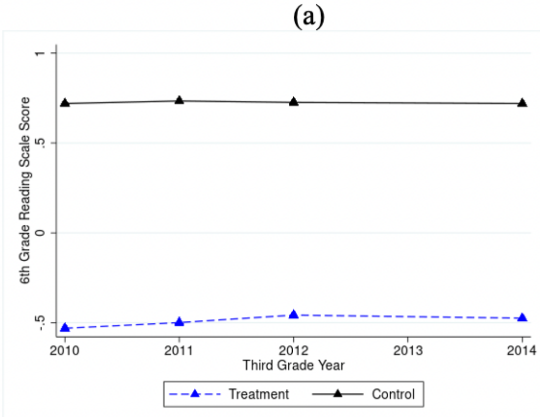
Figure 2.6 Time Trends in Standardized Reading Scores for Third-grade Students scoring below proficient (treatment) compared to Third-grade Students scoring at or above proficient (control)



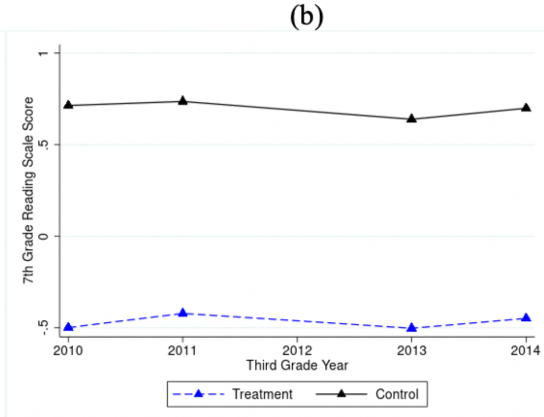
(a)



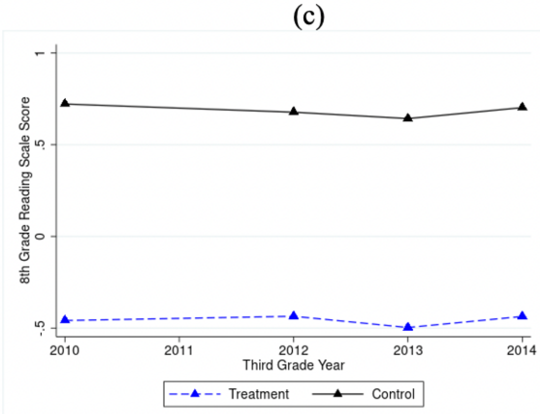
(b)



(c)



(d)



(e)

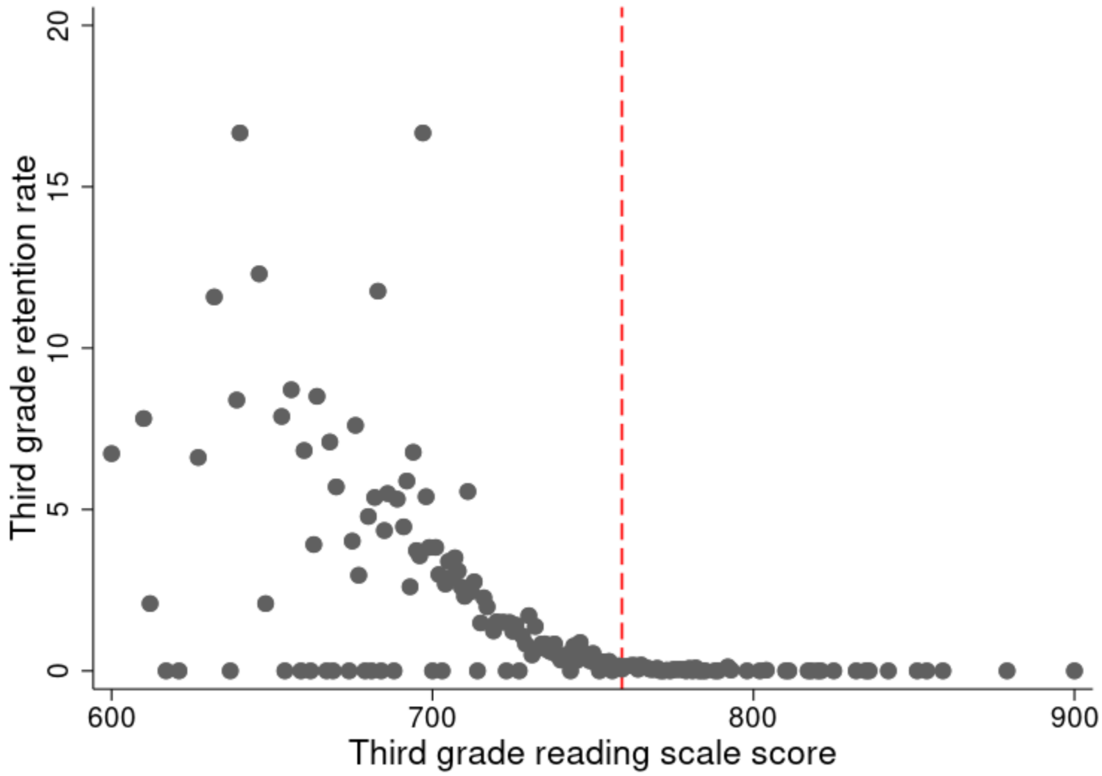
Notes: This figure shows the annual mean standardized reading scores in 4^{th}-8^{th} grades for students in the treatment and control groups by third-grade cohort year. Panel (a) plots the annual mean score in 4^{th} grade for students in third grade from SY2009-10 through SY2013-14, Panels (b)-(e) plot this for grades 5-8 respectively. Students are classified as in the treatment group if their third-grade standardized reading score is below proficient, while students are in the control group if their third-grade standardized reading score is proficient or above. Due to the lack of testing administered in the state of Tennessee in SY2015-16, we are unable to calculate annual mean test scores for 5^{th}-8^{th} grade for certain third-grade cohorts.

Figure 2.7 Students scoring below-proficient on Tennessee Third-Grade Standardized Reading Exams

Year	Percent of 3rd graders below proficient in reading	Number of 3rd graders below proficient in reading
2010	58%	42,269
2011	57%	39,469
2012	54%	37,227
2013	51%	36,351
2014	57%	40,487
2015	58%	43,514
2017	66%	49,465
2018	64%	46,610
2019	64%	45,993

Notes: No comparison data was available for 2016 and 2020 because testing was canceled in those years.

Figure 2.8 Third-Grade Retention Rates by third-grade reading standardized test scale score



Notes: This figure displays third-grade retention rates by third-grade reading scale score pooled across the treatment years (third graders in the academic years 2011-12, 2012-13, 2013-14). During these three years, the proficiency cut score for achieving a passing score was the same. All students scoring a 760 or above were considered proficient.

CHAPTER 3

3 School Lunch Nutrition and Student Outcomes: Analyzing the Impact and Implementation of the Healthy Hunger-Free Kids Act of 2010

3.1 Motivation

A well-documented and widely-accepted literature links childhood nutrition to both cognitive and non-cognitive outcomes. Medical literature on this topic proposes that there are three main channels through which we can observe changes in children's outcomes related to their nutrition: physical development, cognition, and behavior (Sorhaindo and Feinstein 2006). This linkage has prompted further study on the role that schools play in this relationship through the delivery of the National School Lunch Program (NSLP) and School Breakfast Program (SBP), which respectively feed approximately 29.6 million and 14.8 million children each school day. Past research shows that increased access to school meals improves cognitive and behavioral outcomes such as student achievement, attendance, and discipline at the school level (Gordon and Ruffini 2021; Frisvold 2015; Schwartz and Rothbart 2020). However, less is known about the causal impact of the *quality* of these meals provided to students, and how changes in school meal content and quality can impact student outcomes.

In this paper, I provide evidence on the causal impacts of the implementation of the Healthy Hunger-Free Kids Act (HHFKA) of 2010, which dramatically altered nutritional standards for the National School Lunch Program for the first time in recent decades. The HHFKA of 2010 was marketed as a program that would primarily target the growing problem of childhood obesity in the U.S. by improving the nutritional quality of school meals provided largely to low-income and at-risk students. Approximately 80% of school meals served in the U.S. are served as free or reduced-price meals (USDA). I exploit the fact that students eligible for free and reduced-price meals participate in school meals programs at much higher rates than their peers who are not eligible for this assisted pricing to identify the effect of changing school meal nutritional standards on student outcomes.

I identify the effect of changing nutritional standards using a generalized difference-in-differences approach, exploiting variation in the extent to which a school was exposed to improved school meal nutritional quality generated by this new legislation, to examine how changing the quality of school meals impacted student outcomes such as achievement, attendance, discipline, and meal participation. In analyzing school-level outcomes and utilizing a school-level treatment, I can identify the impact of this law on the entire school environment, accounting for differential exposure to spillovers and peer effects induced by an increase in treatment exposure. The main empirical specification in this paper is a dose-response difference-in-differences model which compares changes in outcomes after the policy for schools that vary in their degree of "exposure" to the policy, where "exposure" is measured as the pre-policy share of students eligible for free or reduced-price school meals. The key identifying assumption underlying this strategy is that in the absence of the policy change, trends in outcomes for higher "dose" schools would evolve in parallel to the trends in lower "dose" schools.

I find reductions in disciplinary rates after the introduction of the new meals standards that are similar in magnitude to the disciplinary reductions found in prior research which studies the introduction of a universal free meals program. I find null effects on attendance, middle school mathematics proficiency, and very small albeit statistically significant reductions in elementary reading proficiency. This paper additionally finds associational evidence that students eligible for free or reduced-price meals tended to consume fewer monthly meals on average, possibly due to shifting tastes or preferences regarding the newer meal menus. In doing so, I contribute new evidence on how meal quality can affect student outcomes. This is especially important in light of existing work showing how student outcomes in early grades contribute to life-long human capital formation (Chetty, Friedman and Rockoff 2014; Naven 2019). I also provide new evidence that healthier food initiatives do not deter students from attending school and may improve nutritional intake and behavioral outcomes during the day. For lower-income students that face a calorie crunch outside of school time, such policies could have additional positive spillovers (Kuhn 2018).

There is a fairly extensive economic literature which studies the relationship between improving meal *access* and student outcomes. Gordon and Ruffini (2021) find that the Community Eligibility Provision, an aspect of the HHFKA which allows low-income schools to offer universal free meals to all students, improves disciplinary outcomes in the form of reduced numbers of suspensions. Frisvold (2015) finds that improved access to school breakfast improves student test scores, and Schwartz and Rothbart (2020) find that universal free meals programs improve student achievement. Leos-Urbel et al. (2013) find that universal free meals programs increase student participation in school meals across the board – particularly for students who were previously ineligible for free or reduced price meals. Imberman and Kugler (2014) find that relocating breakfast services from the cafeteria to in-classroom improves student test scores – indicating that both access *and* delivery method can have consequences for student outcomes. In addition to school-provided meals, Bond et al. (2021) link SNAP benefit receipt timing to academic achievement, finding that students who participate in college entrance exams toward the end of the benefit cycle perform worse on their exams and are less likely to attend college.

In addition to the broad literature encompassing the impacts of improved access to meals in general, some prior research more specifically addresses the Healthy Hunger-Free Kids Act's impacts on student outcomes. Vaudrin et al. (2018) found that student participation in school meals programs in 4 New Jersey cities was not significantly changed by the implementation of the Healthy Hunger-Free Kids Act, though they find that meals participation for free and reduced-price lunch (FRPL) eligible students decreased in the first year of implementation. Kenney et al. (2020) found that after the implementation of the HHFKA, obesity rates for children eligible for free and reduced price meals decreased. Bergman et al. (2014) finds that schools were generally compliers with the new regulations, and nutritional content of school meals was significantly improved following the implementation of the HHFKA. This study documented school meal content in four elementary schools before and after implementation of the law by taking digital photographs of school meals purchased and consumed, and found significant improvements in both selected and consumed key nutrients after the HHFKA new meals. These changes included reductions in sodium and the percentage of calories from saturated fat, as well as an increase in fiber and a reduction in calcium. Kinderknecht et al. (2020) also found improved dietary quality amongst NSLP participants after implementation of the HHFKA.

Less is known about the causal impacts of the nutritional quality of school lunches. Anderson, Gallagher, and Ritchie (2018) exploit variation in school meal vendor healthy-eating-indices to estimate the impact of contracting with a "healthy" school meal vendor on student outcomes. The authors find that students at schools which contracted with healthy school meal vendors scored higher on state achievement tests, and that those effects were 50% larger for FRPL-eligible students. The authors find no evidence of changes in attendance and obesity outcomes, or the number of school lunches served. Belot and James (2011) study a healthy meals campaign in the United Kingdom which altered the nutritional content of school lunches in certain UK boroughs. They find that healthier meals were associated with a significant increase in English and Science scores, as well as a decrease in authorized absences – which are most likely linked to illness or health. Figlio and Winicki (2005) find that schools in Virginia facing testing-based accountability sanctions from the state significantly altered their school meal menus on testing dates as an apparent attempt to influence students short-term cognitive function. The districts facing state sanctions modified their menus by increasing caloric content on testing days, and the paper finds associational evidence that this increased the pass rates of 5^{th} graders at the treated schools.

In this paper, I employ a different identification strategy than the quality-related literature above, by utilizing a dose-response specification exploiting the difference in likely exposure to changes in school meals induced by the HHFKA to identify how changes in meal quality influence student outcomes. Additionally, this paper explores a more diverse set of outcomes than has been previously explored in the lunch quality literature, including student achievement, attendance, discipline, and meal participation as well as heterogeneous effects by student characteristics. I utilize school-level administrative data from the state of Texas on student achievement, attendance, discipline, and school meal participation, as well as national school-level achievement and administrative data from the U.S. Department of Education EDFacts and the NCES Common Core of Data. I use this collection of data sources to evaluate the extent to which changing nutritional standards for school lunches impacts student outcomes, and thus contribute to our collective understanding of the overarching influence of school meals and nutritional assistance programs on students.

3.2 The National School Lunch Program

The National School Lunch Program is the second largest nutrition assistance program in the United States, and is administered by the U.S. Department of Agriculture (USDA) Food and Nutrition Service (FNS). The NSLP feeds approximately 29.6 million children each school day, and costs $13.8 billion dollars to operate annually (USDA 2021). The vast majority of traditional public schools participate in the NSLP, and participation in the program is available to all public schools, charter schools, non-profit private schools, and residential care facilities. In 2019, more than 4.8 billion lunches were served nationwide.

The NSLP was established in 1946 under President Truman as part of the National School Lunch Act. At the time, two chief concerns guiding the passage of the National School Lunch Act were childhood malnutrition, and farm surplus. The NSLP was established as a way to simultaneously target these issues, by providing a direct avenue to consume farm surplus as well as a nutritional assistance program for children. Participating schools can purchase foods directly from the USDA to supplement their meals programs, which are generally provided from surplus agricultural stocks. In 1966, the School Breakfast Program was established, and is similarly operated through the USDA at eligible and participating schools across the country. Though school breakfast participation has grown steadily over time, breakfast participation remains consistently lower than lunch participation across the board. While the majority of schools who participate in the NSLP also participate in the SBP, student participation rates in the SBP are only about 50% of the participation rates of the NSLP (USDA 2021).

Schools that choose to participate in the NSLP are reimbursed at a standard federal rate for every NSLP-qualifying meal they serve. For the 2020-21 school year, schools were reimbursed at the following rates for meals served: $0.41 for paid meals, $3.28 for reduced price meals, and $3.68 for free meals[21]. In order for meals to qualify as reimbursable, they must follow all NSLP guidelines regarding health and nutritional quality, meal content, the school must offer free or reduced price meals to all eligible students, and the school must be an official NSLP participant.

Today, 1 in 6 children live in a food insecure household (USDA). Students who attend NSLP-participating schools can purchase lunches during the school day at either a free, reduced, or paid rate, depending on their income-based eligibility. Students are eligible to receive their school meals at a free rate if their household income falls below 130 percent of the federal poverty line, or if their household participates in certain federal assistance programs like SNAP, TANF, or WIC. Students are eligible to purchase school meals at a reduced rate if their household income falls between 130 and 185 percent of the federal poverty line, and students purchasing reduced-rate lunches cannot be charged more than $0.40 for a lunch. Students who do not fall within these categories are able to purchase a school lunch at a federally subsidized paid rate.

3.2.1 Healthy Hunger-Free Kids Act of 2010

In 2010, the Healthy Hunger-Free Kids Act announced sweeping changes to school meals programs as part of a major push by the Obama Administration to tackle childhood obesity and food insecurity. Part of this new legislation included a provision to implement nutritional standards changes to school meals programs; this was the first legislation to make major changes to the nutritional standards for the NSLP in recent decades. The HHFKA new nutritional standards final rule was announced in the Spring of 2012, and all NSLP-participating schools nationwide were required to implement these new lunch standards in the school year starting in the Fall of 2012. Schools participating in School Breakfast Programs were also required to implement nutritional standards changes to their breakfast programs beginning in the Fall of 2013. Alongside changes to the nutritional standards, the HHFKA allotted for an additional $0.06 increase in reimbursement funding per-eligible-meal for schools who participate in school meals programs to account for rising costs of nutritional standards implementation.

In addition to school meal nutritional quality changes, the HHFKA included a provision to allow low-income

[21] Unless the school participates in the Community Eligibility Provision, in which case they are reimbursed differently.

schools and districts to provide universal free meals to all students through the Community Eligibility Provision (CEP). This provision, implemented nationwide in SY 2014-15, allowed for schools or districts with greater than 40% of students in their population labeled categorically eligible[22] for free or reduced-price lunch to offer universal free meals to all students in their school or district. CEP-participating schools are reimbursed for their meals programs using a school-level formula, which differs from schools without CEP. Much of the literature on CEP adoption has shown that offering these universal free meals programs has significant positive impacts on student outcomes, particularly those students that were previously ineligible for free or reduced-price school meals (Gordon and Ruffini 2021; Kho 2018; Gordanier et al. 2020). In the appendix, I include robustness checks which drop ever-participants in CEP from the study population to identify school meal quality effects which are not plausibly driven by later-access to universal free meals programs.

3.3 Data

In this paper, I utilize data from the Texas Education Agency (TEA), the Texas Department of Agriculture (TDA), the NCES Common Core of Data (CCD), and the U.S. Department of Education EDFacts files. In these data, students classified as Economically Disadvantaged are equivalent to students that are eligible for Free & Reduced Price Lunch. Throughout this analysis, the terminology to describe this student group will be used interchangeably.

In all main analyses, I drop non-traditional public schools (charters[23], juvenile institutions, private schools, etc.), and schools that do not participate in the NSLP in the year prior to treatment[24]. Due to the extremely high rates of NSLP participation amongst traditional public schools, if data on pre-treatment NSLP participation is not available, I assume those schools are NSLP-participants and include them in this analysis[25]. All data include school-level administrative characteristics on enrollment demographics, NSLP participation, grade type (elementary, middle, or high school), and operation type (public, charter, etc.).

3.3.1 Exclusionary Discipline

School-level discipline data is from the Texas Education Agency for school years 2007-08 through 2016-17. These data include combined suspension and expulsion rates at the school level. The single school-level measure for discipline consists of total counts of actions which are categorized as in-school-suspensions (ISS), out-of-school suspensions (OSS), and expulsions. Data is reported as action counts (number of actions recorded), student counts (number of students receiving at least one or more actions), and percentages (percent of students receiving at least one or more actions). The main outcome of interest is the percent of students receiving one or more exclusionary

[22] A student is deemed "categorically eligible" if their family receives another form of assistance targeted to low-income families, such as SNAP, TANF, or the Food Distribution Program on Indian Reservations.
[23] All results are robust to including NSLP-participating charter schools.
[24] In TEA discipline data, excluded schools account for N=5,795 observations. In TEA attendance data, excluded schools account for N=7,861 observations. In EDFacts data, excluded schools account for N=36,053 observations.
[25] All results are robust to excluding the schools with missing pre-treatment NSLP participation data.

discipline actions, where exclusionary discipline is defined as a disciplinary action that removes a student from the classroom (in school or out of school suspensions or expulsions).

3.3.2 *Attendance*

The Texas Education Agency provides information on school-level attendance rates for school years 2004-05 through 2016-17. This data further includes attendance rates at the school level for various student subgroups, including students characterized as economically disadvantaged, students with disabilities, and race and gender categories.

3.3.3 *Achievement*

To study effects of the policy on achievement, I use data on standardized test scores from the Texas Education Agency which spans school years 2006-07 through 2016-17 and includes proficiency rates on Texas state assessments for public school students in grades 3-8. In the 2011-12 school year, Texas changed their standardized testing regime from the Texas Assessment of Knowledge and Skills (TAKS) to the State of Texas Assessments of Academic Readiness (STAAR). School-level pass rates changed significantly after the introduction of this new testing regime, which was implemented in the year prior to the treatment studied in this paper. For this reason, I supplement the Texas achievement data with national-level achievement data from the Department of Education EDFacts[26] for school years 2009-10 through 2013-14. The national data includes school-level proficiency rates for standardized testing in grades 3-8, and can be broken down by student subgroups. To supplement this achievement data with school-level characteristics, I utilize the NCES Common Core of Data[27] to include enrollment demographics, school type (elementary, middle, or high), and other school-level information in this national dataset.

3.3.4 *School Meals Participation*

The Texas Department of Agriculture provides data on school meals participation for school years 2011-12 through 2016-17. These data include school-level information on the number of free, reduced, and paid meals served annually, as well as the number of students within each school eligible to purchase meals at each of these pricing rates. The data also include contact information for school food authorities, flags for universal free meals participation, school-level NSLP and SBP participation information, and other administrative information regarding annual school meal operations.

All results utilizing this data source are reported as monthly per-capita calculations. These counts are calculated by dividing the total number of meals served in a school year at a given school by the total number of students enrolled monthly summed across the entire school year. For example, if a school has 50 students enrolled each month, and operated their school lunch program for 8 months, the enrollment quantity for that school in this

[26] These data were accessed through the Urban Institute's Education Data Portal.
[27] These data were accessed through the Urban Institute's Education Data Portal.

data is given as 50 x 8 = 400. Thus, the means and per-capita measures shown in the results tables using this data source should be interpreted as per-capita per-month averages.

3.4 Methods

The main empirical specification is a dose-response difference-in-differences model which compares changes in outcomes after the policy for schools that vary in their degree of "exposure" to the policy. Exposure is measured by the pre-policy share of students eligible for free or reduced-price lunches through the National School Lunch Program. Schools with a higher fraction of students eligible for free and reduced-price lunch overall have higher rates of school meals participation. Figure 3.1 presents the state-level breakdown of the percentage of meals served at each pricing rate in Texas, and indicates that Texas students consume around 80% of school meals at the free or reduced price rate, which is in line with national-level findings from the USDA. Figure 3.2 plots the relationship between monthly per-capita meals served at the school level and the percent of students eligible for free and reduced-price lunch at the school level in the year prior to treatment (SY 2011-12). At all levels and grade types, schools with a higher percentage of students on free and reduced price lunch are also the schools with the greatest rates of monthly per capita school lunch participation. This proportional makeup of pricing of meals served does not seem to dramatically change in the years following the introduction of the HHFKA.

The treatment dose is defined as the average percent of students that are eligible for free or reduced price lunch at the school level in the pre-period, and is a fraction between 0 and 1. The dose remains constant at the school level throughout the entirety of the study period. Post-treatment is equal to 1 in 2013 (SY2012-13) and beyond after required implementation of the HHFKA new nutritional standards. I estimate the equation below:

$$y_{st} = \alpha + \gamma D_S + \beta(P_t x D_s) + \gamma_t + \delta_s + \epsilon_{st} \tag{1}$$

In equation (1), y_{st} is the outcome of interest for school s at time t, D_s is equal to the treatment dose for school s, γ_t are year fixed effects, and δ_s are school-level fixed effects. Figure 3.8 displays the distribution of the treatment dose in each dataset. In my main specification, estimates are weighted by school-level enrollment and standard errors are clustered at the school level, but all main results are robust to utilizing an unweighted specification.

Because all NSLP-participating schools were required to implement new meal standards in the Fall of 2012, there is no variation in treatment timing in this setting. Due to the setting and methodological approach, this analysis will identify the intent-to-treat effect of the HHFKA new nutritional guidelines on student outcomes. The key identifying assumption required for this approach to produce causal effects is that in the absence of the policy change, trends in outcomes for higher "dose" schools would evolve in parallel to the trends in lower "dose" schools.

To the extent that we believe that impacted outcomes for FRPL-eligible students could induce spillover effects on outcomes for non-FRPL eligible students, the dose response specification accounts for these spillovers by assigning a higher dose to schools with a higher number of students likely to be regularly utilizing school lunch as a primary source of nutrition. For example, if we believe that decreases in exclusionary discipline for FRPL students due to a reduction in school fighting incidents could also reduce the exclusionary discipline for a non-FRPL student involved in the same altercation, this specification can best account for these peer effects.

One concern with this approach is that higher-income schools/districts were already serving higher quality meals before the HHFKA was implemented. While I do not observe school lunch nutritional quality directly, as long as the key identifying assumption holds and the differences in nutritional quality between low- and high-poverty schools are not changing over time in the pre-period, this would only result in my approach underestimating the effects of the policy.

3.5 Results

3.5.1 School Meals Participation

In order to further motivate mechanisms behind the policy effect studied in this paper and understand how student-level behaviors and tastes could have changed in response to the newer meal menus, I analyze school meals operations data from the Texas Department of Agriculture. These data span school years 2011-12 through 2016-17, and the new meals standards due to the HHFKA were introduced in SY 2012-13.

Figure 3.3 shows the time trends in monthly per-capita school lunch participation for students in all schools. In all schools, overall meal participation at a per-student level decreased in the post period, with a temporary increase only in high schools in the first year of treatment. Table 3.1 supplements these figures with a regression specification estimating the post-policy impact of the nutritional standards implementation on monthly per-capita meal participation in all schools, and by school grade type. Across the board, meals participation decreases at a statistically significant rate for free and reduced-price meals students. Estimates indicate an overall decrease in lunch participation of approximately 1 meal per month for free and reduced price meals students, who consume approximately 10-12 meals per month on average in this sample. Although estimates for paid meals are statistically insignificant for all schools, results are mixed depending on school type.

Understanding the potential mechanisms that could be influencing these results is important to interpreting any further results on student achievement-related outcomes. Across the board, I find that free and reduced price lunch students decrease their school lunch participation. Prior research has shown that students eligible for free or reduced-price meals participate more frequently in school meals programs, and often utilize school lunches as a source of primary nutrition at much higher rates than students only eligible for paid meals. Despite this fact, I do find associational evidence that students who participate at the free or reduced rates respond to the policy effect by consuming fewer meals on average after the implementation of the HHFKA nutritional standards. While I am not able to directly observe student-level reasoning behind this shift in consumption, it is possible that student preferences changed in response to meal content changes.

3.5.2 *Effects of New Nutritional Standards on Attendance*

School-level attendance data from the Texas Education Agency spans school years 2004-05 through 2016-17. The outcomes are measured as average school-level attendance rates for the listed student subgroup. Figure 3.4 outlines the time trends in attendance rates for schools with a defined treatment dose above the median treatment dose in comparison to schools with a treatment dose below the median. There is little significant change in attendance rates over time, and average attendance rates overall are quite high at 97% for elementary schools, 96% for middle schools, and 93% for high schools. The majority of the estimates from my main specification produce small, albeit sometimes statistically significant effects. To interpret the magnitude of these effects, I consider a 1 standard deviation increase in free lunch eligibility in elementary schools, where the average treatment dose is equal to 0.633 and the standard deviation is 0.267. Given the point estimate from the estimation specification shown in Table 3.2, a one standard deviation increase in the elementary treatment dose corresponds to a 0.0007 percentage point decrease in attendance, or a 0.00069% reduction. Across all schools, at the 5% level these estimates exclude effects greater than a 0.0008 percentage decrease or less than a 0.0005 percentage decrease in attendance. I conclude that these effects, while sometimes statistically significant, are insignificant in magnitude.

Figure 3.5 displays an event study specification where attendance is the primary outcome, which shows only very small magnitude negative effects that are driven by elementary schools. Table 3.2 presents the results for the dose response estimation for schools by grade type. Appendix Table 3.5 breaks down these results by heterogeneous subgroups in all schools. All specifications result in essentially null effects on overall school level attendance as a result of the HHFKA new nutritional standards implementation.

3.5.3 *Effects of New Nutritional Standards on Exclusionary Discipline*

The Texas Education Agency provides data on exclusionary discipline rates at the school-level for school years 2007-08 through 2016-17. The main outcome studied in this analysis is the percentage of students at the school level receiving at least one or more exclusionary discipline actions in a year. Exclusionary discipline actions are defined as actions which remove a student from their normal classroom environment, such as in-school suspensions, out-of-school suspensions, or expulsions.

Figure 3.6 displays the time trends for school-level disciplinary rates for schools with an assigned treatment dose above the median dose as compared to schools with a treatment dose below the median broken down by school grade type. Figure 3.7 displays the dose response event studies for the same outcome broken down by school grade type which illustrates the similarity in pre-trends between high dose and low dose schools and the divergence in these trends in the post period. Negative pre-trends for elementary schools indicate a potential for bias, and therefore these results should be interpreted with caution.

Table 3.3 displays the results from the main specification which analyzes the effect of the implementation of new nutritional standards on student discipline rates. On average, 6% of elementary schoolers, 23% of middle schoolers, and 22% of high schoolers have at least one exclusionary discipline infraction per year over the sample study period. Panel A of Figure 3.8 displays the distribution of the treatment dosage at the school level for the discipline data analyzed here. The average treatment dose is 0.569 with a standard deviation of 0.243 for middle schools, and an average of 0.504 with a standard deviation of 0.226 for high schools. Thus, the estimates in Table 3.3 for middle schools suggest that a one standard deviation increase in treatment exposure to the new nutritional standards leads to a 1.38 percentage point decrease in students receiving a disciplinary action, which corresponds to a 5.9% decrease. In high schools, I find a one standard deviation increase in exposure leads to a 2.08 percentage point decrease in students receiving a disciplinary action, which corresponds to a 9.3% decrease.

Overall, my estimates suggest that fewer students were subject to exclusionary discipline after the new school meal standards are implemented. Gordon and Ruffini (2021) find that state-level CEP adoption decreased middle school suspension rates by 6%, while Kho (2018) finds that CEP adoption in Tennessee decreased the rate of students disciplined in schools by 10%, and finds the largest reductions in high schools. These estimates suggest that improving nutritional quality of the school meals, in addition to increasing overall access, can have implications for behavioral outcomes for students. A broad medical literature has found that improved nutritional intake can positively influence behavioral outcomes, which supports the potential mechanisms influencing these results to be improved nutritional intake of the students consuming these new school lunches in the post-period. However, due to data limitations I am unable to directly attribute these discipline reductions to improved nutritional intake. Other potential mechanisms for discipline reductions include decreased meal participation leading to less-crowded cafeteria rooms, improved peer effects resulting in reduced suspension rates in the entire school environment, or changes in teacher participation in school meals. Though I cannot observe teacher participation in school meals, if teachers also consume these school meals regularly and are additionally impacted by nutritional quality of these meals, this could influence disciplinary practices from the top down. While my estimates are in line with findings in prior literature on the effects of school nutrition on student discipline outcomes, they remain significantly smaller in magnitude than prior research on school-level interventions that specifically target reductions in discipline, such as Restorative Justice programs (Davison, Penner, and Penner 2022).

3.5.4 Effects of New Nutritional Standards on Achievement

In my main achievement analysis, I utilize both Texas state-level data on student achievement, and supplement this analysis with a national-level dataset from the Common Core of Data and EDFacts[28].

The Texas Education Agency provides school-level proficiency rates on standardized test scores for school years 2005-06 through 2015-16. From school years 2005-06 through 2010-11, the standardized test administered in the state of Texas was the Texas Assessment of Knowledge and Skills (TAKS). In school year 2011-12, Texas began a new testing regime, switching their standardized testing to the State of Texas Assessments of Academic Readiness (STAAR) exam.

Figure 3.9 shows average proficiency rates in Math and Reading exams over time in the state of Texas. Due

[28] These data were accessed through the Urban Institute's Education Data Portal.

to this change in testing regime, Texas proficiency rates in math and reading assessments dropped significantly, and the Texas Education Agency reports that these testing scores are not comparable over regime changes. In analyzing this data, there is also evidence (outlined in Appendix C) that this testing regime change differentially impacted economically disadvantaged students as compared to their non-economically disadvantaged peers. I formally test for whether this testing regime change differentially impacted the treatment group in Appendix Table 3.12, and find evidence that this is the case. Due to this disparity in testing proficiency rates induced by the testing regime change, I will be unable to draw credibly causal conclusions about the impact of new nutritional standards in Texas on achievement scores due to the nature of treatment as defined in this model.

For this reason, I supplement my state-level analysis with a national-level achievement analysis. To properly account for changing testing regimes and their potentially differential effect on economically disadvantaged students as demonstrated in the Texas data, I drop any states that experience a testing regime change within the study period[29]. Additionally, I exclude states which conduct their testing in the Fall semester[30]. Within the EDFacts proficiency data, school-by-grade-level proficiency rates are reported as a range with a high, low, and midpoint proficiency rate. The main outcome of interest utilized in this paper are the midpoint proficiency rates[31].

Figure 3.10 displays time trends in average proficiency rates for all students in the national level data. Figure 3.11 further decomposes these trends for elementary and middle school test scores for schools with a treatment dose above the median versus schools with a treatment dose below the median. Differential pre-trends in middle school reading and elementary school mathematics as outlined in the event studies displayed in Figure 3.12 indicate a potential for bias in the estimates, and therefore I am unable to interpret those results as causal.

Table 3.4 outlines the results of the main specification on reading and mathematics proficiency rates by school type. Due to the pre-trends outlined above, I am unable to draw credible conclusions from the elementary school mathematics or middle school reading results outlined in this table. However, I do find evidence of reductions in elementary school reading proficiency rates and null effects on middle school mathematics scores. To interpret the magnitude of these results, I consider a 1 standard deviation increase in the treatment dose, 0.265, which corresponds to a 0.14 percentage point decrease in the elementary school reading proficiency rate, or a 0.002% decrease. Thus, we can conclude that changes in school lunch nutritional standards resulted in very small, albeit statistically significant, decreases in elementary reading proficiency rates and null effects on middle school mathematics proficiency rates.

3.5.5 Robustness Checks

To supplement my main analyses, I conduct various robustness checks which are outlined in Appendix B. My main results for attendance, discipline, and achievement are robust to various checks such as including charter

[29] The excluded states due to testing regime changes are: Alabama, Florida, Illinois, Michigan, Minnesota, Mississippi, New York, North Carolina, Oklahoma, Oregon, Pennsylvania, Texas, Utah, Virginia, Wyoming.
[30] The excluded states due to Fall testing are: Maine, Michigan, Nebraska, New Hampshire, North Dakota, Rhode Island, Vermont, Wisconsin.
[31] All results are robust to using the high end or low end of proficiency ranges as a primary outcome of interest.

schools, excluding weighting in the regression specification, including only schools which participate in both the NSLP and SBP, and dropping ever-participants in the CEP program. I additionally estimate placebo checks (outlined in Appendix Table 3.11) where I place grade-level enrollment percentages on the left hand side of the estimating equation, and find null results of the dose response treatment on the percentages of enrolled students in each grade at the school level, an outcome that we should not expect to be impacted by this treatment.

3.6 Discussion & Conclusion

There is a general consensus within the prior literature that improved access to nutrition in the form of school meals has a positive impact on student outcomes. Less is known about the causal impact of nutritional quality of these meals, but recent literature has concluded mostly positive or null effects of improving the health quality of school meals provided to students. This paper adds to this literature by providing some of the first causal evidence regarding the impacts of the most recent and most dramatic change to the nutritional standards in school meals across the country induced by the Healthy Hunger-Free Kids Act of 2010. The results indicate that students eligible for free or reduced-price meals tended to consume fewer monthly meals on average, possibly due to shifting tastes or preferences regarding the newer meal menus. I further find reductions in disciplinary rates after the introduction of the new meals standards, indicating positive impacts on short-term behavioral outcomes for students as a result of the changes in nutritional content of meals. I find largely null effects on attendance overall, null effects on middle school mathematics achievement, and very small reductions in elementary reading achievement.

Overall, the body of work on school nutrition programs indicates that food served in schools can be greatly influential in students' short and long-term outcomes. This second-largest nutrition program in the U.S. reaches almost 30 million students each day, and therefore research on its impacts carries particular policy relevance. The findings in this paper coupled with prior literature suggest that *quality* of these meals, in addition to access, can have important implications for student outcomes in the form of reduced incidence of exclusionary discipline while not dramatically impacting overall student standardized test achievement or deterring students from attending school. More research is needed to determine the exact mechanisms through which improved school meal nutritional quality impacts student outcomes to better inform policy decisions around this important topic.

3.7 References

[1] Akin, John S., David K. Guilkey, Barry M. Popkin, and James H. Wyckoff. 1983. "The Demand for School Lunches: An Analysis of Individual Participation in the School Lunch Program." The Journal of Human Resources 18 (2): 213–30. https://doi.org/10.2307/145483.

[2] Alderman, Harold, Johannes Hoogeveen, and M. Rossi. 2009. "Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania." Economic Development and Cultural Change 57 (February): 239–60. https://doi.org/10.1086/592875.

[3] Anderson, Michael, Justin Gallagher, and Elizabeth Ramirez Ritchie. 2017. "School Lunch Quality and Academic Performance." w23218. Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w23218.pdf.

[4] Anderson, Patricia M., and Kristin F. Butcher. 2006. "Reading, Writing, and Refreshments: Are School Finances Contributing to Children's Obesity?" Journal of Human Resources 41 (3): 467–94.

[5] Belot, Michèle, and Jonathan James. 2011. "Healthy School Meals and Educational Outcomes." Journal of Health Economics 30 (3): 489–504. https://doi.org/10.1016/j.jhealeco.2011.02.003.

[6] Bergman, Ethan A, Tim Englund, Katie Weigt Taylor, Tracee Watkins, Stephen Schepman, and Keith Rushing. 2014. "School Lunch Before and After Implementation of the Healthy Hunger-Free Kids Act." The Journal of Child Nutrition and Management 38 (2): 12.

[7] Bond, Timothy N., Jillian B. Carr, Analisa Packham, and Jonathan Smith. 2022. "Hungry for Success? SNAP Timing, High-Stakes Exam Performance, and College Attendance." American Economic Journal: Economic Policy 14 (4): 51–79. https://doi.org/10.1257/pol.20210026.

[8] Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." American Economic Review 104 (9): 2633–79. https://doi.org/10.1257/aer.104.9.2633.

[9] "Common Core of Data Directory, Education Data Portal (Version 0.14.0)." Urban Institute. Accessed March 5, 2022. https://educationdata.urban.org/documentation/.

[10] Confessore, Nicholas. 2014. "How School Lunch Became the Latest Political Battleground: The Food Issue." New York Times (Online). New York, United States: New York Times Company. October 7, 2014. https://www.proquest.com/docview/2212956331/abstract/CACF5665FB214457PQ/1.

[11] Cornish, Disa, Natoshia Askelson, and Elizabeth Golembiewski. 2016. "'Reforms Looked Really Good on Paper': Rural Food Service Responses to the Healthy, Hunger-Free Kids Act of 2010." Journal of School Health 86 (2): 113–20. https://doi.org/10.1111/josh.12356.

[12] Datar, Ashlesha, and Nancy Nicosia. 2012. "Junk Food in Schools and Childhood Obesity." Journal of Policy Analysis and Management : [The Journal of the Association for Public Policy Analysis and Management] 31 (2):

312–37. https://doi.org/10.1002/pam.21602.

[13] Davison, Miles, Andrew M. Penner, and Emily K. Penner. 2022. “Restorative for All? Racial Disproportionality and School Discipline Under Restorative Justice.” American Educational Research Journal 59 (4): 687–718. https://doi.org/10.3102/00028312211062613.

[14] “EDFacts, Education Data Portal (Version 0.14.0).” Urban Institute. Accessed March 5, 2022. https://educationdata.urban.org/documentation/.

[15] Figlio, David N., and Joshua Winicki. 2005. “Food for Thought: The Effects of School Accountability Plans on School Nutrition.” Journal of Public Economics 89 (2–3): 381–94. https://doi.org/10.1016/j.jpubeco.2003.10.007.

[16] “Final Rule: Nutrition Standards in the National School Lunch and School Breakfast Programs | Food and Nutrition Service.” Accessed February 13, 2024. https://www.fns.usda.gov/cn/fr-012612.

[17] Frisvold, David E. 2015. “Nutrition and Cognitive Achievement: An Evaluation of the School Breakfast Program.” Journal of Public Economics 124 (April): 91–104. https://doi.org/10.1016/j.jpubeco.2014.12.003.

[18] Golley, R, E Baines, P Bassett, L Wood, J Pearce, and M Nelson. 2010. “School Lunch and Learning Behaviour in Primary Schools: An Intervention Study.” European Journal of Clinical Nutrition 64 (11): 1280–88. https://doi.org/10.1038/ejcn.2010.150.

[19] Gordanier, John, Orgul Ozturk, Breyon Williams, and Crystal Zhan. 2020. “Free Lunch for All! The Effect of the Community Eligibility Provision on Academic Outcomes.” Economics of Education Review 77 (August): 101999. https://doi.org/10.1016/j.econedurev.2020.101999.

[20] Gordon, Nora, and Krista Ruffini. 2021. “Schoolwide Free Meals and Student Discipline: Effects of the Community Eligibility Provision.” Education Finance and Policy 16 (3): 418–42. https://doi.org/10.1162/edfp_a_00307.

[21] Hinrichs, Peter. 2010. “The Effects of the National School Lunch Program on Education and Health.” Journal of Policy Analysis and Management 29 (3): 479–505. https://doi.org/10.1002/pam.20506.

[22] Imberman, Scott A., and Adriana D. Kugler. 2014. “The Effect of Providing Breakfast in Class on Student Performance.” Journal of Policy Analysis and Management 33 (3): 669–99. https://doi.org/10.1002/pam.21759.

[23] Johnson, Donna B., Mary Podrabsky, Anita Rocha, and Jennifer J. Otten. 2016. “Effect of the Healthy Hunger-Free Kids Act on the Nutritional Quality of Meals Selected by Students and School Lunch Participation Rates.” JAMA Pediatrics 170 (1): e153918. https://doi.org/10.1001/jamapediatrics.2015.3918.

[24] Kenney, Erica L., Jessica L. Barrett, Sara N. Bleich, Zachary J. Ward, Angie L. Cradock, and Steven L. Gortmaker. 2020. “Impact Of The Healthy, Hunger-Free Kids Act On Obesity Trends: Study Examines Impact of the Healthy, Hunger-Free Kids Act of 2010 on Childhood Obesity Trends.” Health Affairs 39 (7): 1122–29. https://doi.org/10.1377/hlthaff.2020.00133.

[25] Kho, Adam. 2018. “Three Essays on School Reform.” Ph.D., United States -- Tennessee: Vanderbilt University.

https://www.proquest.com/docview/2184637416/abstract/6112097CA8BE42A8PQ/1.

[26] Kinderknecht, Kelsey, Cristen Harris, and Jessica Jones-Smith. 2020. "Association of the Healthy, Hunger-Free Kids Act With Dietary Quality Among Children in the US National School Lunch Program." JAMA 324 (4): 359. https://doi.org/10.1001/jama.2020.9517.

[27] Kuhn, Michael A. 2018. "Who Feels the Calorie Crunch and When? The Impact of School Meals on Cyclical Food Insecurity." Journal of Public Economics 166 (October): 27–38. https://doi.org/10.1016/j.jpubeco.2018.08.001.

[28] Leos-Urbel, Jacob, Amy Ellen Schwartz, Meryle Weinstein, and Sean Corcoran. 2013. "Not Just for Poor Kids: The Impact of Universal Free School Breakfast on Meal Participation and Student Outcomes." Economics of Education Review 36 (October): 88–107. https://doi.org/10.1016/j.econedurev.2013.06.007.

[29] "National School Lunch Program | Food and Nutrition Service." Accessed February 13, 2024. https://www.fns.usda.gov/nslp.

[30] Naven, Matthew. 2019. "Human-Capital Formation During Childhood and Adolescence: Evidence from School Quality and Postsecondary Success in California." Munich Personal RePEc Archive, December.

[31] Ohri-Vachaspati, Punam, Lindsey Turner, Marc A. Adams, Meg Bruening, and Frank J. Chaloupka. 2016. "School Resources and Engagement in Technical Assistance Programs Is Associated with Higher Prevalence of Salad Bars in Elementary School Lunches in the United States." Journal of the Academy of Nutrition and Dietetics 116 (3): 417–26. https://doi.org/10.1016/j.jand.2015.10.023.

[32] Ollinger, Michael, Joanne F. Guthrie, and Audrey Peo. 2017. "Which School Food Authorities Pay More for Food Used in USDA School Meal Programs?" Journal of Nutrition Education and Behavior 49 (7): S108. https://doi.org/10.1016/j.jneb.2017.05.133.

[33] Schanzenbach, Diane Whitmore. 2009. "Do School Lunches Contribute to Childhood Obesity?" The Journal of Human Resources 44 (3): 684–709.

[34] Schanzenbach, Diane Whitmore, and Mary Zaki. 2014. "Expanding the School Breakfast Program: Impacts on Children's Consumption, Nutrition and Health." w20308. Cambridge, MA: National Bureau of Economic Research. http://www.nber.org/papers/w20308.pdf.

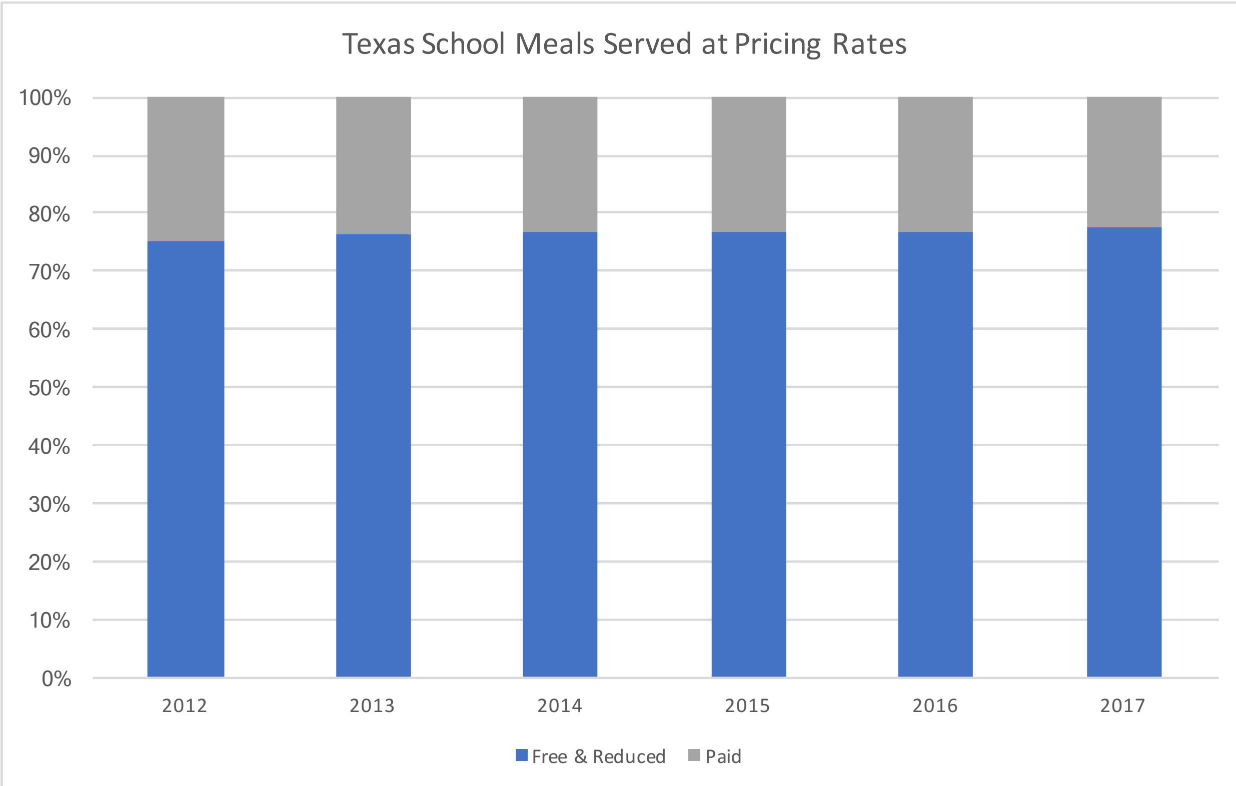
[35] Schwartz, Amy Ellen, and Michah W. Rothbart. 2020. "Let Them Eat Lunch: The Impact of Universal Free Meals on Student Performance." Journal of Policy Analysis and Management 39 (2): 376–410. https://doi.org/10.1002/pam.22175.

[36] Schwartz, Colin, and Margo G. Wootan. 2019. "How a Public Health Goal Became a National Law." Nutrition Today 54 (2): 67–77. https://doi.org/10.1097/NT.0000000000000318.

[37] Sorhaindo, Annik, and Leon Feinstein. 2006. "What Is the Relationship between Child Nutrition and School Outcomes?" Centre for Research on the Wider Benefits of Learning Institute of Education, June.

[38] Vaudrin, Nicole, Kristen Lloyd, Michael J. Yedidia, Michael Todd, and Punam Ohri-Vachaspati. 2018. "Impact of the 2010 US Healthy, Hunger-Free Kids Act on School Breakfast and Lunch Participation Rates Between 2008 and 2015." American Journal of Public Health 108 (1): 84–86. https://doi.org/10.2105/AJPH.2017.304102.

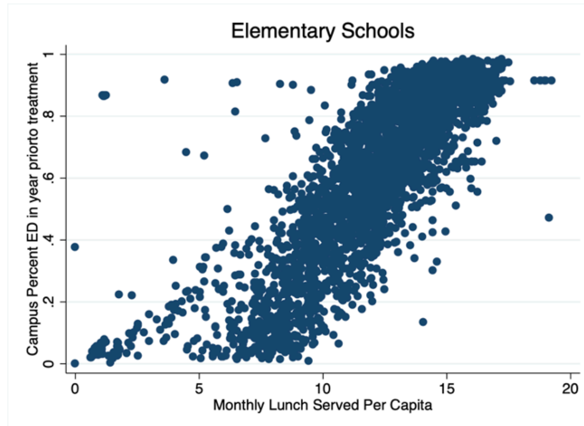
Figure 3.1 School Meals Served in Texas at Free & Reduced and Paid Rates as a proportion of All Meals Served



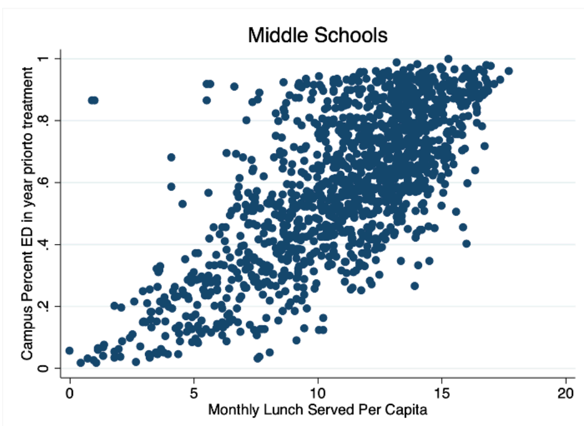
Data Source: Texas Department of Agriculture

Figure 3.2 Scatterplots of Campus-Level Percent ED Students vs. Campus-Level Lunch Served Per-Capita

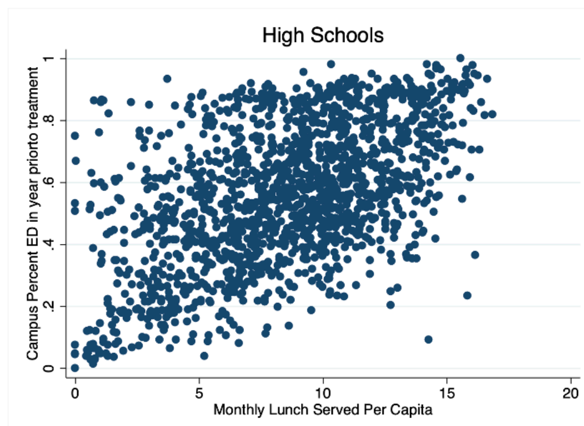
(a)



(b)



(c)

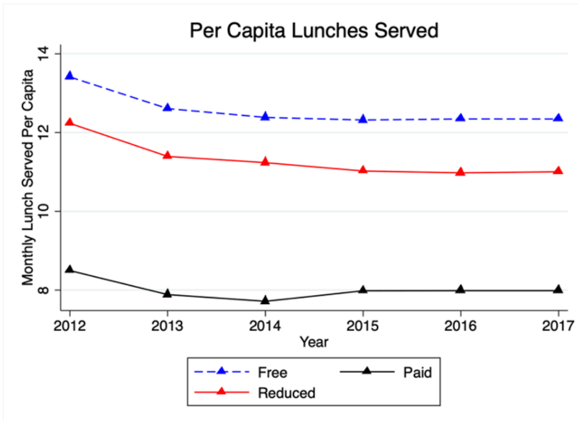


(d)

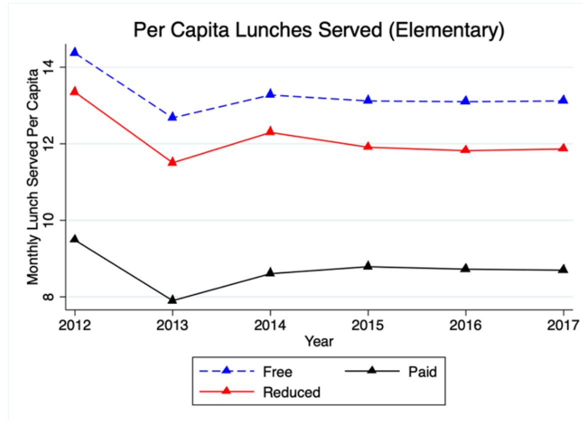
Data Source: Texas Department of Agriculture

Notes: Panel (a) shows the relationship between campus-level percent economically disadvantaged (ED) students in SY 2011-12 and campus-level monthly per capita meals served in 2011-12. Panel (b) shows this relationship in elementary schools, panel (c) for middle schools, and panel (d) for high schools. Panel (a) has a correlation coefficient of 0.7087, panel (b) has a correlation coefficient of 0.8365, panel (c) 0.7302, and panel (d) 0.5380.

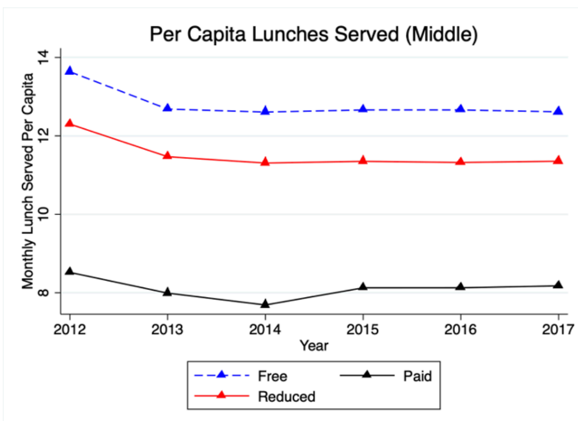
Figure 3.3 Time Trends in per capita lunches served at each payment rate



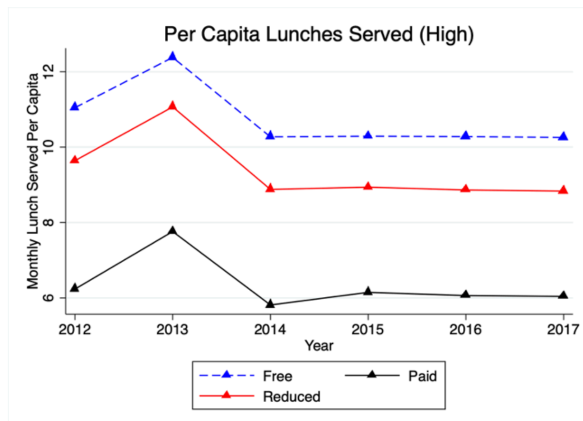
(a)



(b)



(c)



(d)

Data Source: Texas Department of Agriculture

Notes: Panel (a) plots the time trends in monthly per-capita lunches served for all schools. Panel (b) plots this for elementary schools, panel (c) for middle schools, and panel (d) for high schools. This includes only traditional public schools and NSLP participating schools.

Table 3.1 Estimating the impact of the policy change on per-capita school meal participation

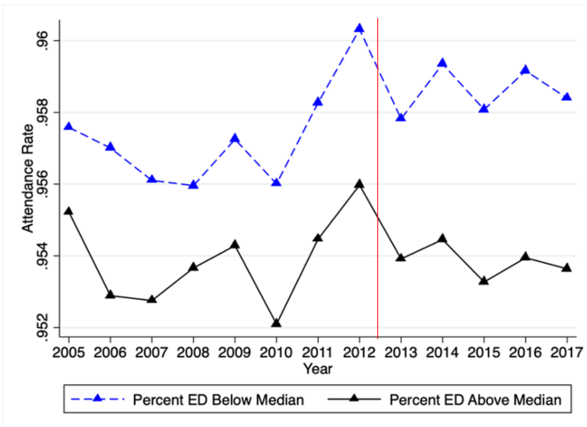
	(1) Paid Meals	(2) Free Meals	(3) Reduced Price Meals
PANEL A: ALL SCHOOLS			
Post 2012	0.02	-1.04***	-0.87***
	(0.05)	(0.02)	(0.03)
Mean	8.01	12.57	11.34
Observations	37744	38850	34865
PANEL B: ELEMENTARY SCHOOLS			
Post 2012	-0.23**	-1.18***	-1.06***
	(0.08)	(0.02)	(0.03)
Mean	8.70	13.28	12.16
Observations	18324	18954	16796
PANEL C: MIDDLE SCHOOLS			
Post 2012	0.02	-1.02***	-0.81***
	(0.10)	(0.05)	(0.06)
Mean	8.11	12.82	11.54
Observations	6910	7112	6362
PANEL D: HIGH SCHOOLS			
Post 2012	0.35***	-0.96***	-0.70***
	(0.10)	(0.05)	(0.06)
Mean	6.34	10.74	9.40
Observations	7980	8194	7454

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

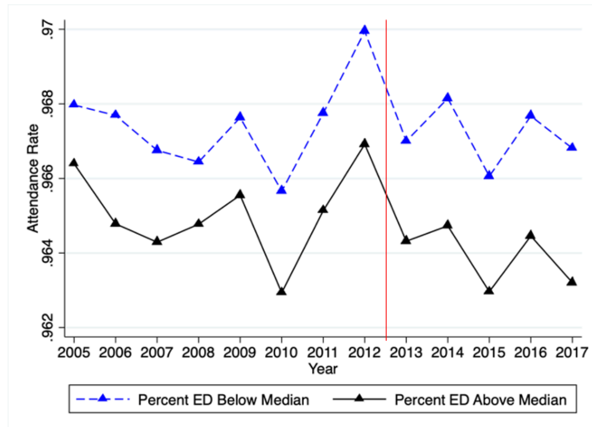
Data Source: Texas Department of Agriculture

Notes: This table estimates the equation $Y_{st} = \alpha + \beta Post2012 + X_s$ where Y_{st} is equal to monthly per-capita meals served and Post2012 is equal to 1 in SY 2012-13 and beyond. These estimates are weighted by enrollment quantity and the specification includes campus fixed effects, with standard errors clustered at the campus level.

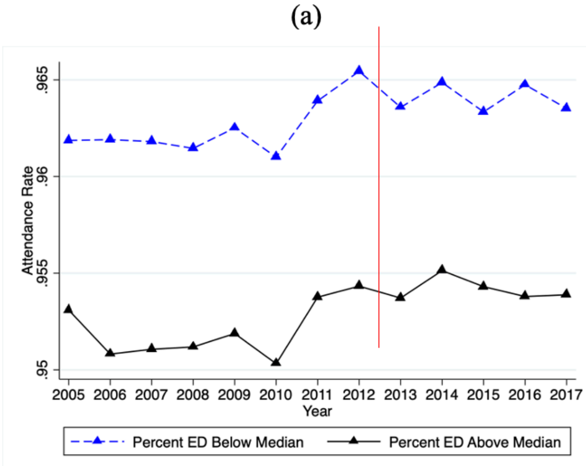
Figure 3.4 Time-trends in attendance rates for schools with a dose above the median vs. schools with a dose below the median



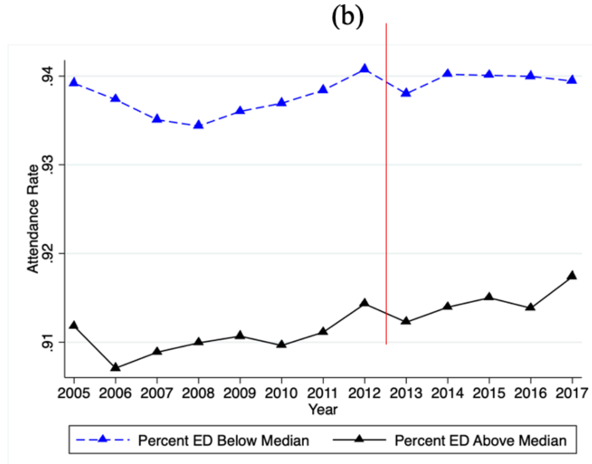
(a)



(b)



(c)

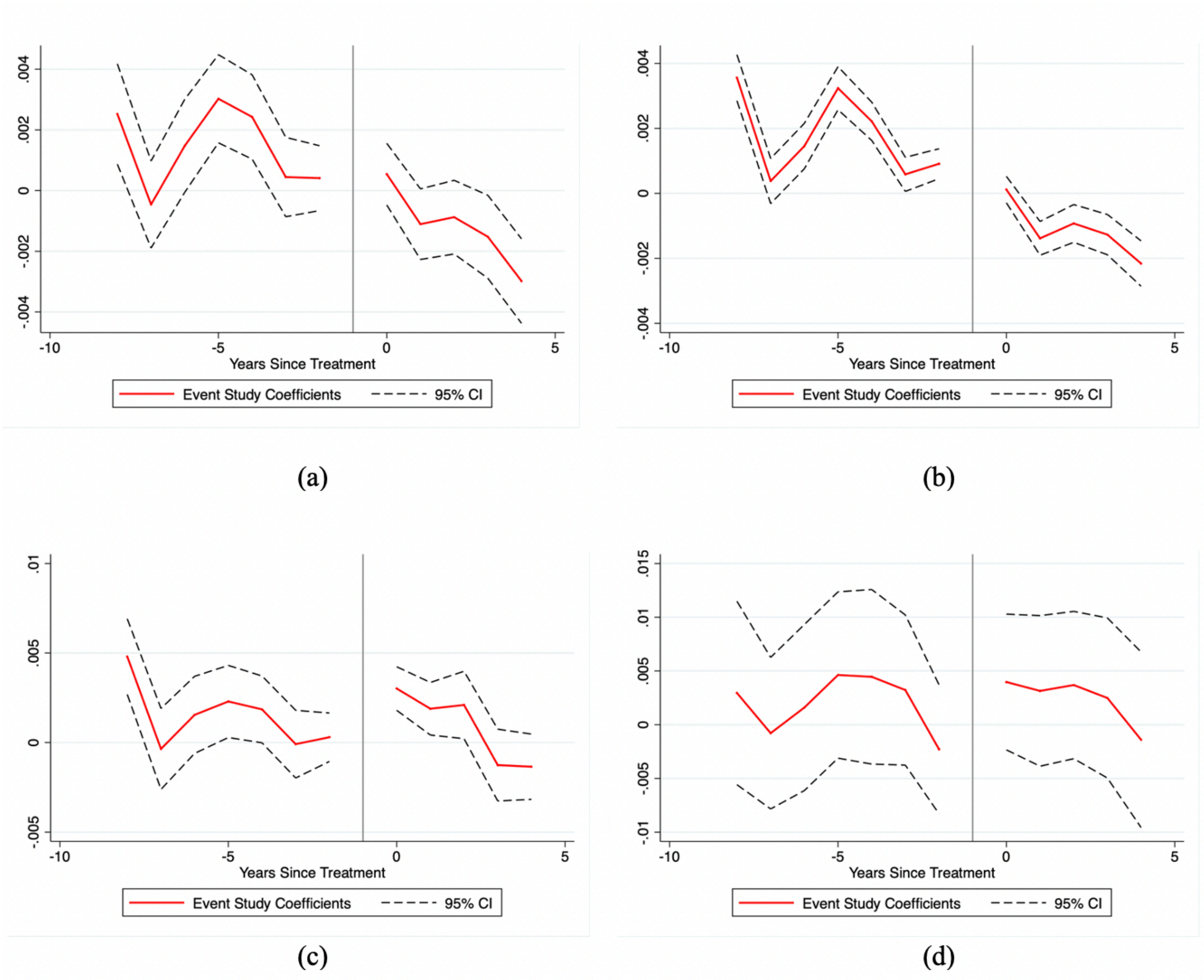


(d)

Data Source: Texas Education Agency

Notes: This figure shows the annual mean attendance rates at the school level in the state of Texas from SY 2004-05 through SY 2016-17. Panel (a) plots the annual mean attendance rates in all schools, panel (b) plots the annual mean attendance rate in elementary schools, panel (c) plots this for middle schools, and panel (d) plots this for high schools. The "dose" is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2004-05 – SY 2011-12). Schools with a baseline average percentage of FRPL eligible students above the statewide median are classified in this figure as being "above median" and are represented by the solid black line. Schools with a baseline percentage of FRPL eligible students below the statewide median are classified as being "below median" and are represented by the dashed blue line.

Figure 3.5 Dose Response Event Study for School-Level Attendance Rates



Data Source: Texas Education Agency

Notes: This figure shows the event studies for overall attendance rates in the pre and post-policy periods. Panel (a) plots the event study for attendance rates in all schools, panel (b) plots the event study for attendance rate in elementary schools, panel (c) plots this for middle schools, and panel (d) plots this for high schools. The “dose” is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2004-05 – SY 2011-12).

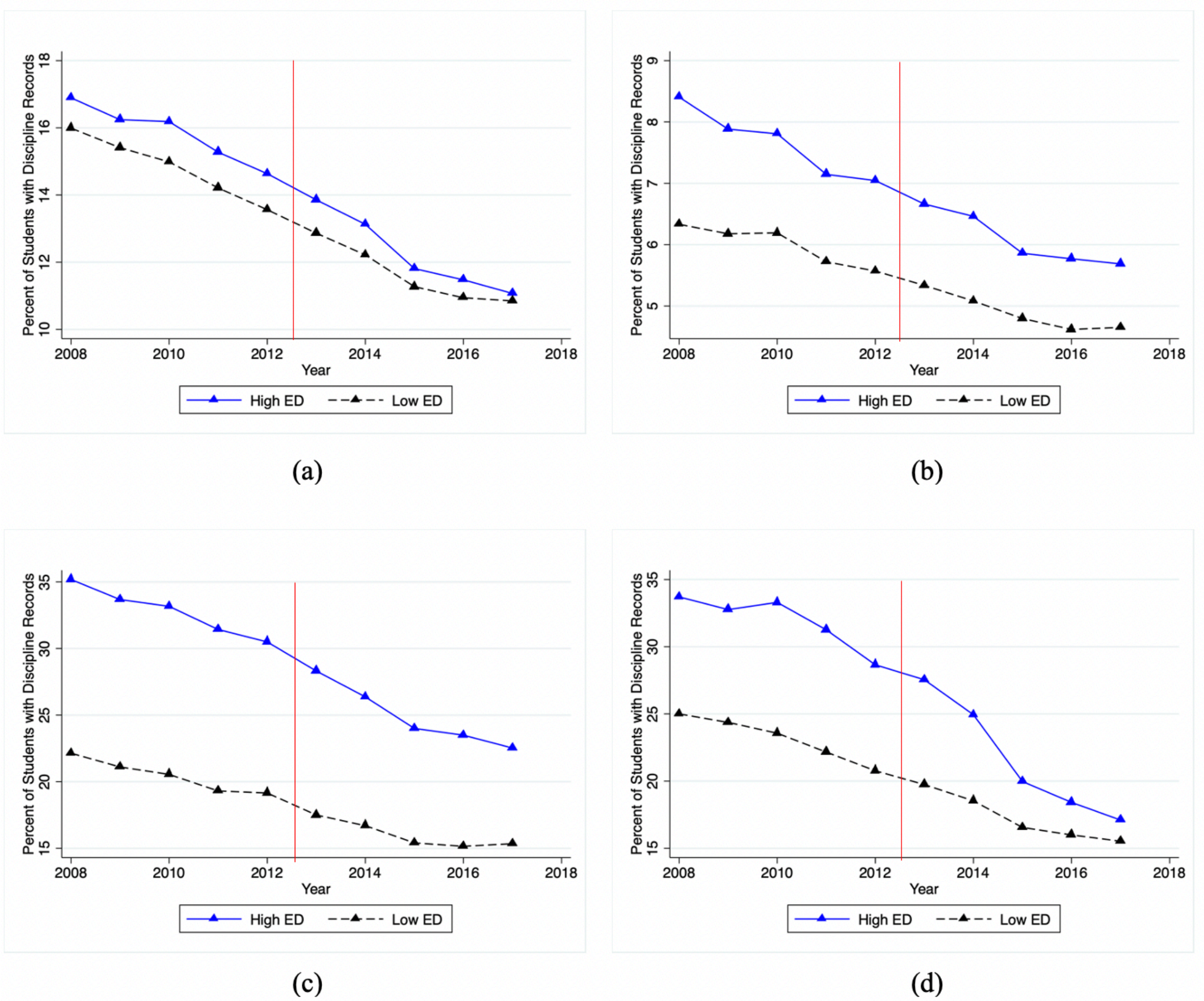
Table 3.2 Dose Response Estimation of School Lunch Nutritional Changes on Attendance Rates

	(1) All	(2) Elementary	(3) Middle	(4) High
Dose x Post	-0.002604***	-0.002533***	-0.000695	0.000065
	(0.000371)	(0.000226)	(0.000580)	(0.001791)
Mean	0.96	0.97	0.96	0.93
Observations	94581	52787	19881	18020

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency

Notes: This table shows the dose response estimation of school lunch nutritional changes on attendance rates at the school level. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2004-05 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools that do not participate in the National School Lunch Program in the year prior to treatment are excluded from this analysis. All specifications include campus and year fixed effects.

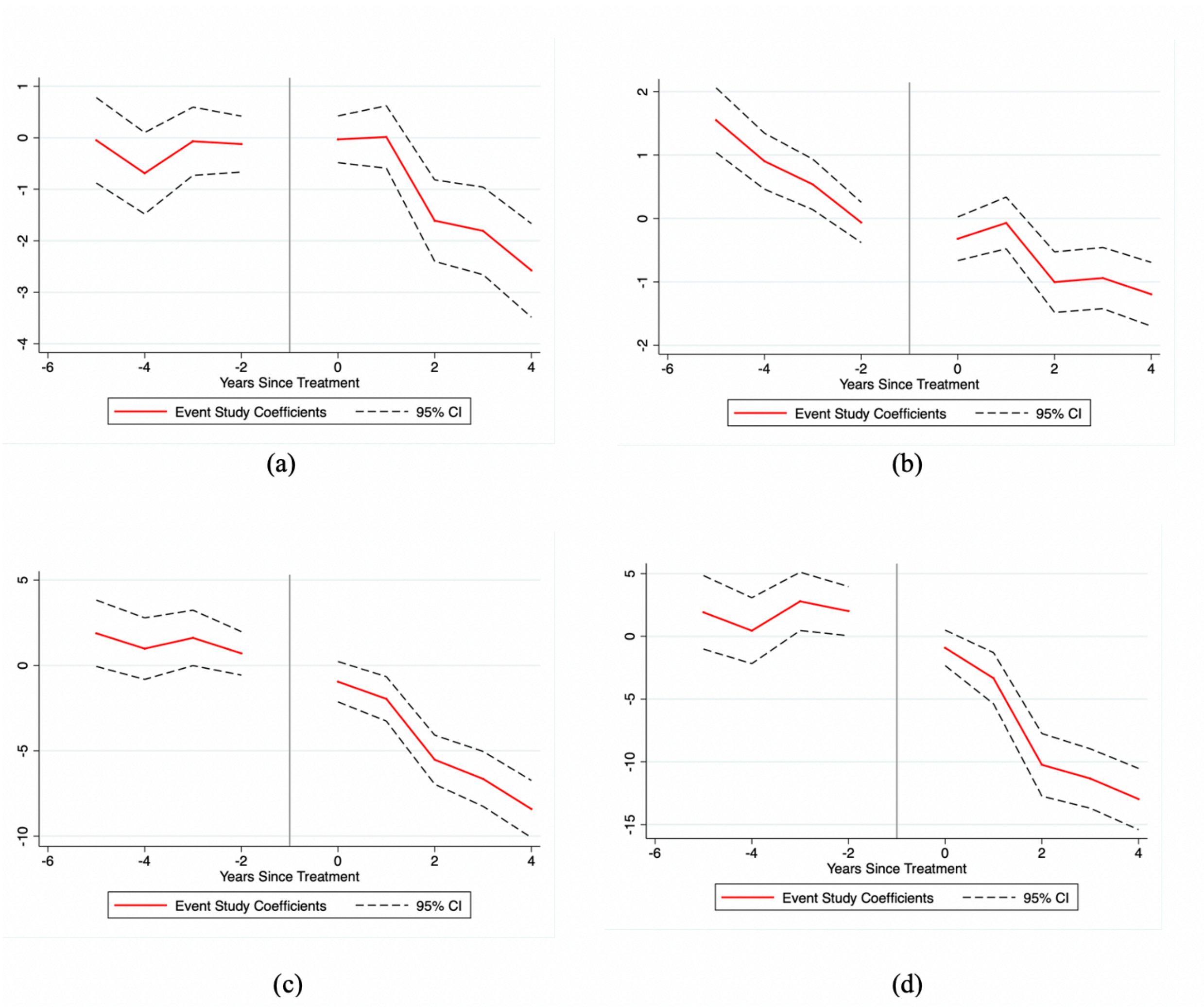
Figure 3.6 Time-trends in disciplinary rates for schools with a dose above the median vs. districts with a dose below the median



Data Source: Texas Education Agency

Notes: This figure shows the annual mean disciplinary rates at the school level in the state of Texas from SY 2007-08 through SY 2016-17 for schools with a treatment dose above the median as compared to schools with a treatment dose below the median. Panel (a) plots the annual mean disciplinary rates in all schools, panel (b) plots the annual mean disciplinary rate in elementary schools, panel (c) plots this for middle schools, and panel (d) plots this for high schools. The "dose" is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2007-08 – SY 2011-12). Schools with a baseline average percentage of FRPL eligible students above the statewide median are classified as being "above median" and are represented by the solid black line. Schools with a baseline percentage of FRPL eligible students below the statewide median are classified as being "below median" and are represented by the blue line.

Figure 3.7 Dose Response Event Study for School-Level Discipline Rates



Data Source: Texas Education Agency

Notes: This figure shows the event studies for overall disciplinary rates in the pre and post-policy periods. Panel (a) plots the event study for disciplinary rates in all schools, panel (b) plots the event study for disciplinary rate in elementary schools, panel (c) plots this for middle schools, and panel (d) plots this for high schools. The "dose" is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2007-08 – SY 2011-12).

Table 3.3 Dose Response Estimation of School Lunch Nutritional Changes on Percent of Students with Exclusionary Discipline Records

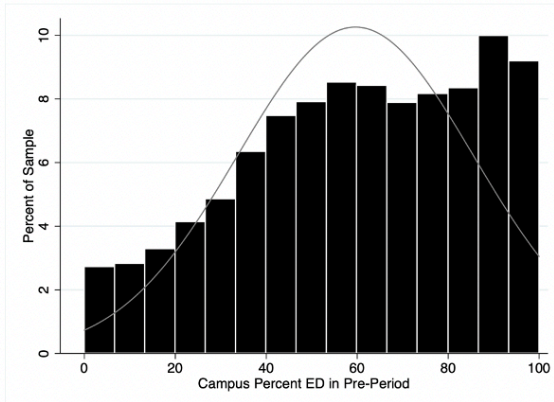
	(1) All	(2) Elementary	(3) Middle	(4) High
Dose x Post	-0.9856*	-1.1152***	-5.6680***	-9.2211***
	(0.3840)	(0.1837)	(0.6737)	(1.1622)
Mean	13.68	6.36	23.09	22.21
Observations	59780	32034	14498	11624

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency

Notes: This table shows the dose response estimation of school lunch nutritional changes on discipline rates in schools by grade type. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2007-08 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools that do not participate in the National School Lunch Program in the year prior to treatment are excluded from this analysis. All specifications include campus and year fixed effects.

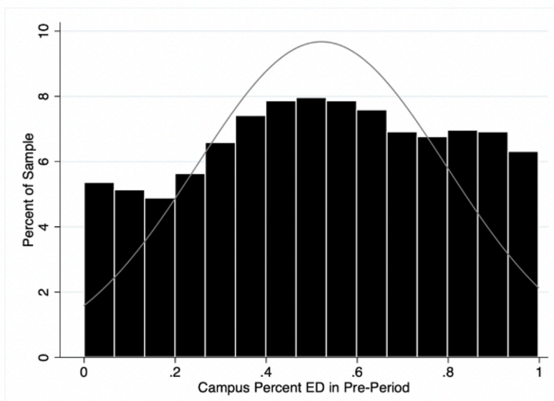
Figure 3.8 Distribution of the Treatment Dose

a. Distribution of the Treatment Dose in the School Level Discipline Data



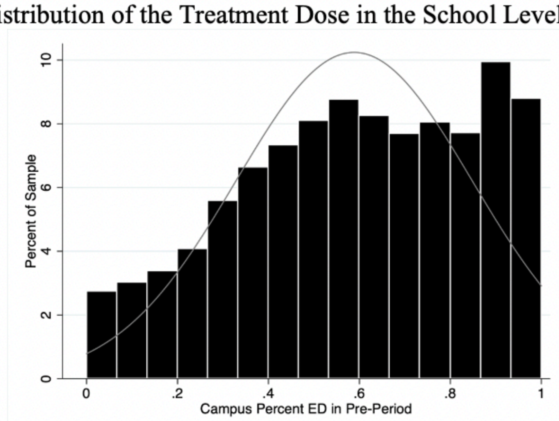
Data Source: Texas Education Agency

b. Distribution of the Treatment Dose in the School Level Achievement Data



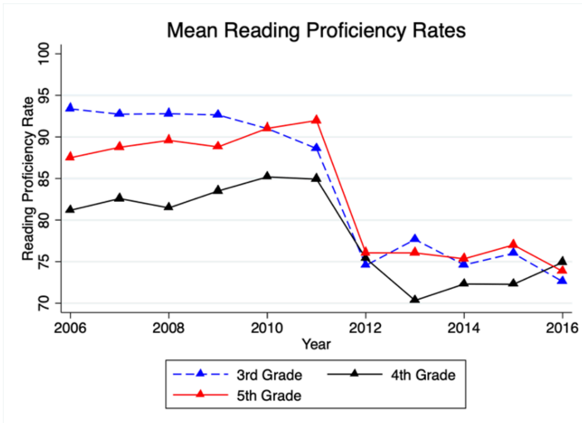
Data Source: NCES Common Core of Data

c. Distribution of the Treatment Dose in the School Level Attendance Data

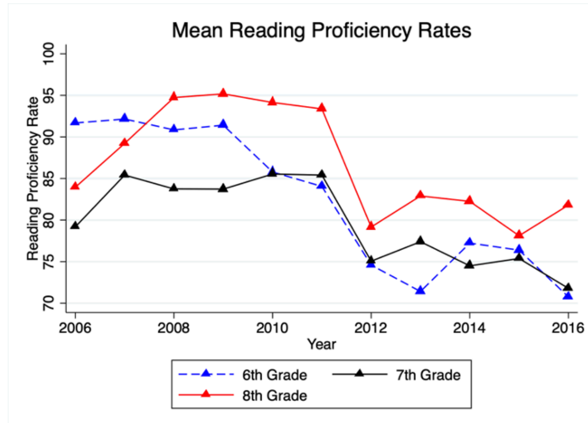


Data Source: Texas Education Agency

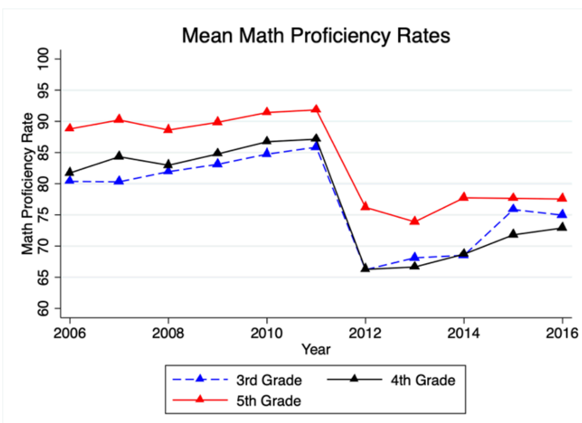
Figure 3.9 Time Trends in Proficiency Rates for All Students in Texas



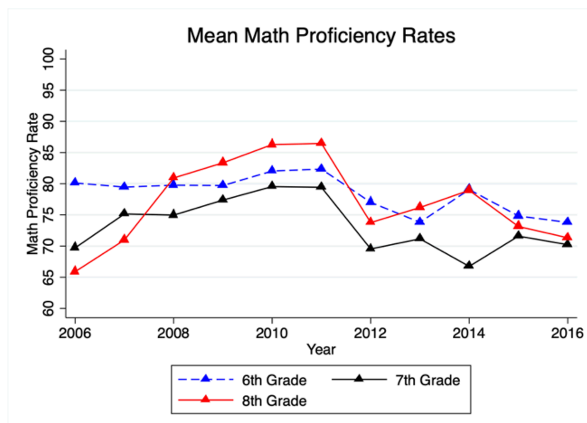
(a)



(b)



(c)

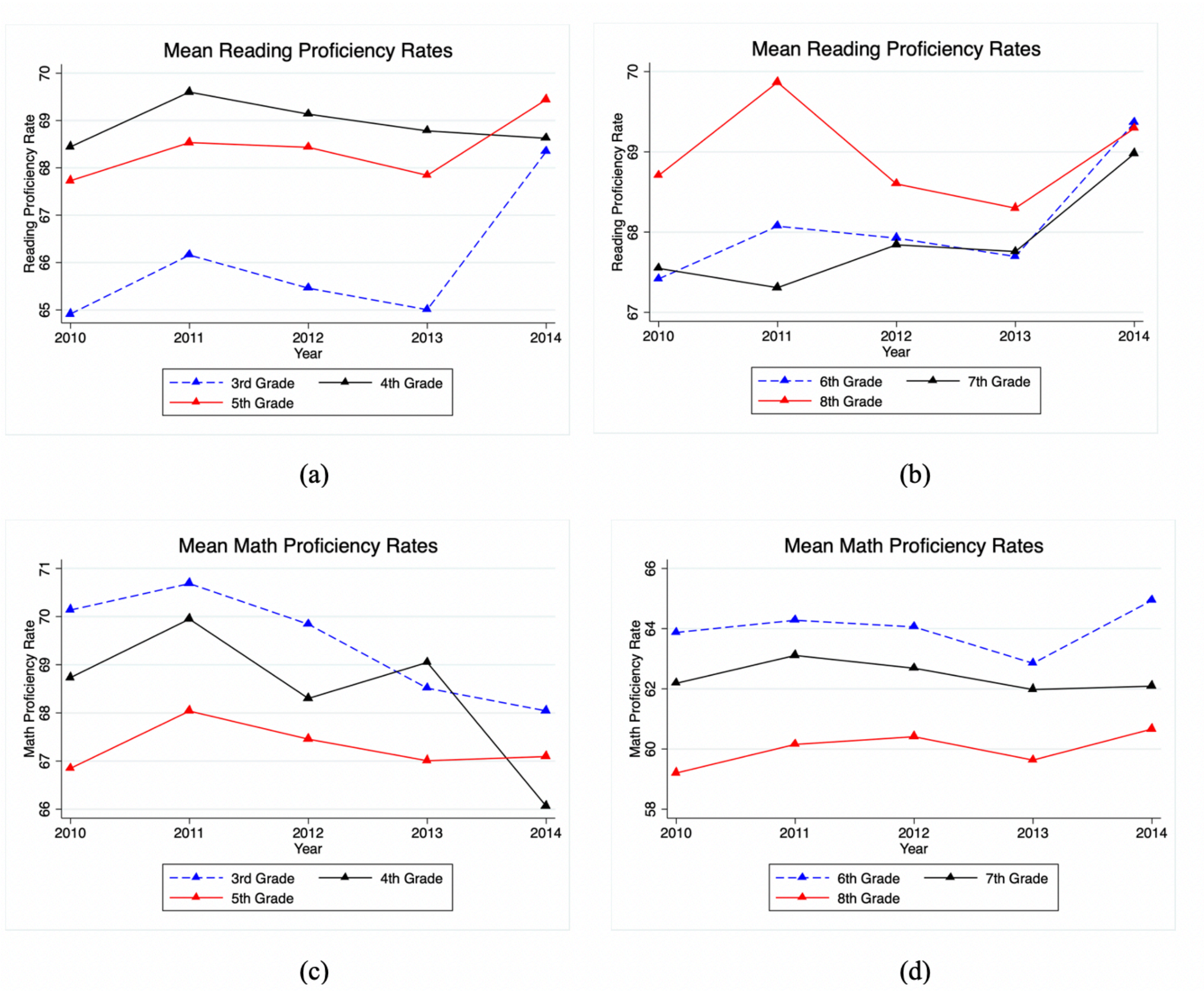


(d)

Data Source: Texas Education Agency

Notes: This figure shows the annual mean proficiency rates at the school level in the state of Texas from SY 2005-06 through SY 2015-16 for reading and math standardized test scores. Panel (a) plots the annual mean proficiency rates on reading exams in elementary school grades, panel (b) plots the annual mean proficiency rates on reading exams in middle school grades, panel (c) plots annual mean proficiency rates on math exams in elementary grades, and panel (d) plots this middle school grades.

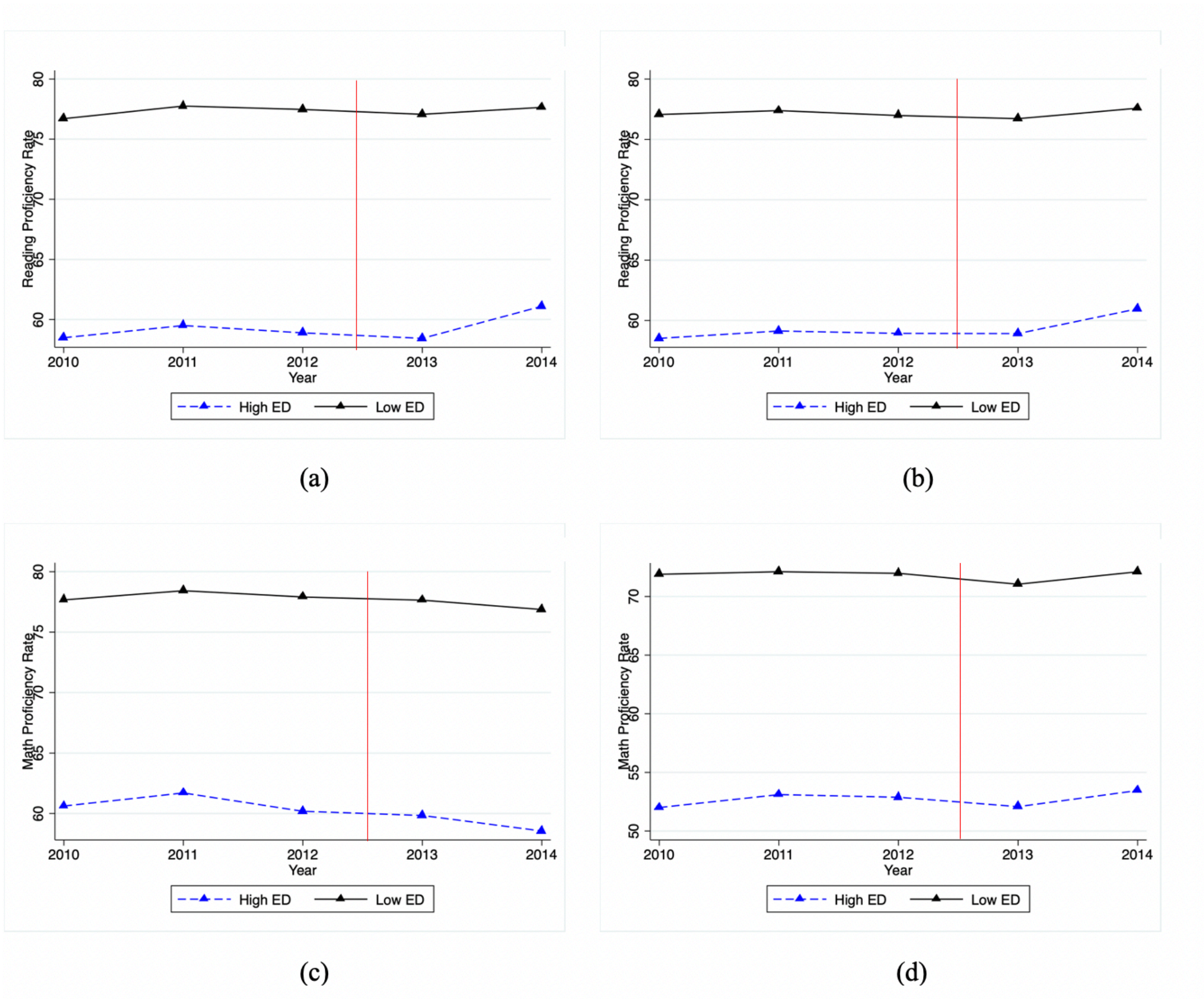
Figure 3.10 Time Trends in Proficiency Rates for All Students in EDFacts National Data



Data Source: U.S. Department of Education EDFacts

Notes: This figure shows the annual mean proficiency rates at the school level in the included states from SY 2009-10 through SY 2013-14 for reading and math standardized test scores. Panel (a) plots the annual mean proficiency rates on reading exams in elementary school grades, panel (b) plots the annual mean proficiency rates on reading exams in middle school grades, panel (c) plots annual mean proficiency rates on math exams in elementary grades, and panel (d) plots this middle school grades.

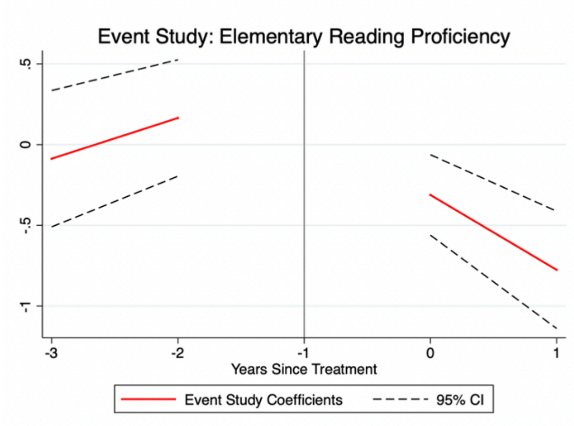
Figure 3.11 Weighted Time Trends in Proficiency Rates for schools with a dose above vs. below the median in EDFacts National Data



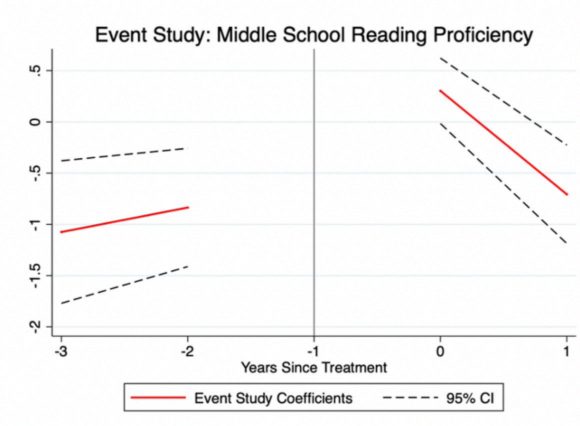
Data Source: U.S. Department of Education EDFacts and NCES Common Core of Data

Notes: This figure shows the annual mean standardized test proficiency rates at the school level in the study states from SY 2009-10 through SY 2013-14 for schools with a treatment dose above the median as compared to schools with a treatment dose below the median. Panel (a) plots the annual mean reading proficiency rates in elementary schools, panel (b) plots the annual mean reading proficiency rates in middle schools, panel (c) plots the annual mean mathematics proficiency rates in elementary schools, and panel (d) plots this for middle schools. The "dose" is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2009-10 – SY 2011-12). Schools with a baseline average percentage of FRPL eligible students above the statewide median are classified as being "above median" and are represented by the solid black line. Schools with a baseline percentage of FRPL eligible students below the statewide median are classified as being "below median" and are represented by the dashed blue line.

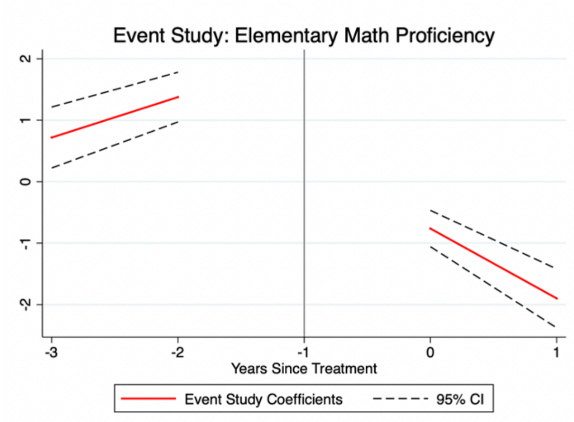
Figure 3.12 Dose Response Event Studies for Standardized Testing Proficiency in EDFacts National Data



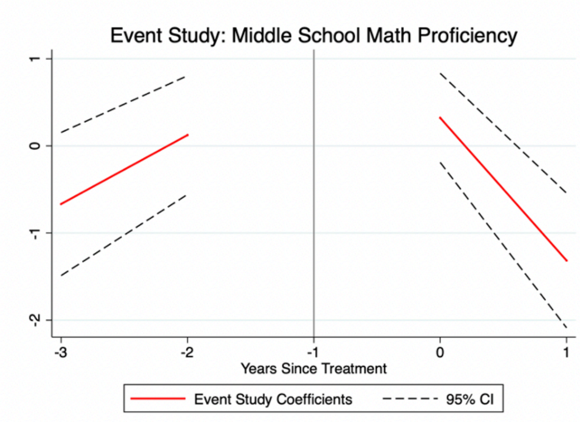
(a)



(b)



(c)



(d)

Data Source: U.S. Department of Education EDFacts and NCES Common Core of Data

Notes: This figure shows the event studies for overall reading and mathematics proficiency rates in the pre and post-policy periods. Panel (a) plots the event study for reading proficiency rates in elementary schools, panel (b) plots the event study for reading proficiency rates in middle schools, panel (c) plots this for mathematics proficiency rates in elementary schools, and panel (d) plots this for middle schools. The “dose” is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2009-10 – SY 2011-12).

Table 3.4 Dose Response Estimation of School Lunch Nutritional Changes on Standardized Testing Proficiency Rates

	(1) Reading	(2) Math
	PANEL A: ELEMENTARY SCHOOLS	
Dose x Post	-0.51**	-1.88***
	(0.16)	(0.19)
Mean	67.68	68.49
Observations	335788	335782
	PANEL B: MIDDLE SCHOOLS	
Dose x Post	0.55*	-0.11
	(0.23)	(0.35)
Mean	68.23	62.29
Observations	177385	177422

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Data Source: U.S. Department of Education EDFacts and NCES Common Core of Data

Notes: This table shows the dose response estimation of school lunch nutritional changes on proficiency rates in schools by grade type. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2009-10 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students tested at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools in states which experienced a testing regime change over the study period are excluded from this analysis. All specifications include campus, year, and state fixed effects.

3.8 Appendix A: Heterogeneous Results

Table 3.5 Dose Response Estimation of School Lunch Nutritional Changes on Attendance Rates for Heterogeneous Groups

	(1) Black	(2) Hispanic	(3) White	(4) NA	(5) ED	(6) IEP
Dose x Post	-0.004508***	-0.005129***	-0.001837***	0.003859	-0.006315***	-0.002352***
	(0.000795)	(0.000544)	(0.000377)	(0.003587)	(0.000515)	(0.000641)
Mean	0.96	0.96	0.95	0.95	0.95	0.95
Observations	68518	92561	84306	9297	93715	91098

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Data Source: Texas Education Agency

Notes: This table shows the dose response estimation of school lunch nutritional changes on attendance rates at the school level for various student subgroups. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2004-05 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students in each student subgroup at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools that do not participate in the National School Lunch Program in the year prior to treatment are excluded from this analysis. All specifications include campus and year fixed effects. NA stands for Native American, ED stands for Economically Disadvantaged, and IEP stands for Individualized Education Plan or students with a documented disability.

Table 3.6 Dose Response Estimation of School Lunch Nutritional Changes on Reading Proficiency Rates by Grade

	(1) Reading	(2) Math
	PANEL A: 3rd Grade	
Dose x Post	-1.21***	-2.96***
	(0.23)	(0.26)
Mean	65.82	69.54
Observations	113229	113214
	PANEL B: 4th Grade	
Dose x Post	-0.23	-1.11***
	(0.23)	(0.27)
Mean	68.94	68.58
Observations	112418	112410
	PANEL A: 5th Grade	
Dose x Post	-0.05	-1.41***
	(0.21)	(0.28)
Mean	68.32	67.30
Observations	108243	108254
	PANEL B: 6th Grade	
Dose x Post	0.39	-0.53
	(0.28)	(0.41)
Mean	68.00	63.93
Observations	70251	70295
	PANEL A: 7th Grade	
Dose x Post	1.16***	-1.24**
	(0.30)	(0.43)
Mean	67.83	62.42
Observations	52929	52924
	PANEL B: 8th Grade	
Dose x Post	0.14	1.77***
	(0.32)	(0.48)
Mean	68.93	59.98
Observations	52706	52684

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: U.S. Department of Education EDFacts and NCES Common Core of Data

Notes: This table shows the dose response estimation of school lunch nutritional changes on proficiency rates in schools by grade level. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2009-10 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students tested at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools in states which experienced a testing regime change over the study period are excluded from this analysis. All specifications include campus, year, and state fixed effects.

Table 3.7 Dose Response Estimation of School Lunch Nutritional Changes on Discipline Rates by School Meal Vendor Status in Texas

	(1) All Schools	(2) Elementary	(3) Middle	(4) High
	PANEL A: School Meals Prepared In-House			
Post 2012	-0.74	-0.8265***	-5.6950***	-9.2556***
	(0.46)	(0.20)	(0.77)	(1.38)
Mean	13.70	6.33	23.10	22.65
Observations	47715	25667	11564	9184
	PANEL B: School Meals Prepared by Outside Vendor			
Post 2012	-1.20	-2.7963***	-5.7368***	-6.5154*
	(0.77)	(0.46)	(1.56)	(2.58)
Mean	14.26	6.50	23.66	24.44
Observations	10394	5757	2604	1912

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency and Texas Department of Agriculture

Notes: This table shows the dose response estimation of school lunch nutritional changes on discipline rates in Texas schools by school meal vendor status. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2009-10 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students tested at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) are excluded from this analysis. All specifications include campus and year fixed effects.

3.9 Appendix B: Robustness Checks

Table 3.8 Robustness of Dose Response Estimation of School Lunch Nutrition on Exclusionary Discipline Rates

	(1) All Schools	(2) Elementary	(3) Middle	(4) High
	PANEL A: Unweighted Estimates			
Post 2012	-0.8006**	-1.1152***	-5.6702***	-8.5405***
	(0.28)	(0.18)	(0.67)	(1.08)
Mean	13.68	6.36	23.09	22.21
Observations	59780	32034	14498	11624
	PANEL B: Including Charter Schools			
Post 2012	-0.7623*	-0.9715***	-5.4906***	-8.8999***
	(0.38)	(0.18)	(0.67)	(1.14)
Mean	13.32	6.29	22.79	21.39
Observations	62787	33127	14779	12323
	PANEL C: Excluding Non-SBP Participating Schools			
Post 2012	-0.8895*	-1.0880***	-5.6346***	-9.1970***
	(0.39)	(0.19)	(0.69)	(1.17)
Mean	13.72	6.39	23.20	22.23
Observations	59520	31879	14416	11601
	PANEL D: Only Confirmed NSLP Participants in Pre-Period			
Post 2012	-0.8533*	-1.2234***	-5.6837***	-8.9091***
	(0.39)	(0.18)	(0.69)	(1.21)
Mean	13.80	6.36	23.20	22.96
Observations	58109	31424	14168	11096
	PANEL E: Excluding Ever-Participants in CEP			
Post 2012	-1.2352**	-1.9077***	-6.2820***	-8.1829***
	(0.42)	(0.22)	(0.72)	(1.29)
Mean	13.42	6.42	21.87	21.28
Observations	48496	25272	11909	9932

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table shows the dose response estimation of school lunch nutritional changes on discipline rates in schools by grade type. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2007-08 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and in Panels B-E estimates are weighted by the total number of students at the school-level. All specifications include campus and year fixed effects.

Table 3.9 Robustness of Dose Response Estimation of School Lunch Nutrition on Attendance

	(1) All Schools	(2) Elementary	(3) Middle	(4) High
		PANEL A: Unweighted Estimates		
Post 2012	-0.002371***	-0.002606***	-0.000304	0.000777
	(0.00)	(0.00)	(0.00)	(0.00)
Mean	0.96	0.97	0.96	0.93
Observations	94581	52787	19881	18020
		PANEL B: Including Charter Schools		
Post 2012	-0.002504***	-0.002441***	-0.000747	0.000170
	(0.00)	(0.00)	(0.00)	(0.00)
Mean	0.96	0.97	0.96	0.93
Observations	99170	54543	20327	19094
		PANEL C: Excluding Non-SBP Participating Schools		
Post 2012	-0.002667***	-0.002569***	-0.000742	0.000123
	(0.00)	(0.00)	(0.00)	(0.00)
Mean	0.96	0.97	0.96	0.93
Observations	93871	52246	19751	17981
		PANEL D: Only Confirmed NSLP Participants in Pre-Period		
Post 2012	-0.002664***	-0.002535***	-0.000658	-0.000170
	(0.00)	(0.00)	(0.00)	(0.00)
Mean	0.96	0.97	0.96	0.93
Observations	90640	51519	19234	16643
		PANEL E: Excluding Ever-Participants in CEP		
Post 2012	-0.001825***	-0.001822***	0.000248	-0.001326
	(0.00)	(0.00)	(0.00)	(0.00)
Mean	0.96	0.97	0.96	0.93
Observations	77500	42295	16374	15433

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency

Notes: This table shows the dose response estimation of school lunch nutritional changes on attendance rates at the school level. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2004-05 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and in Panels B-E estimates are weighted by the total number of students at the school-level. All specifications include campus and year fixed effects.

Table 3.10 Robustness of Dose Response Estimation of School Lunch Nutrition on Reading Achievement

	Panel 1: Reading Achievement		Panel 2: Math Achievement	
	(1) Elementary	(2) Middle	(1) Elementary	(2) Middle
	Panel 1A: Unweighted Estimates		Panel 2A: Unweighted Estimates	
Post 2012	-0.69***	0.43	-2.08***	0.10
	(0.18)	(0.24)	(0.21)	(0.32)
Mean	67.68	68.23	68.49	62.29
Observations	335833	177408	335827	177445
	Panel 1B: Including Charter Schools		Panel 2B: Including Charter Schools	
Post 2012	-0.32*	0.56**	-1.72***	0.00
	(0.15)	(0.22)	(0.19)	(0.33)
Mean	67.56	68.11	68.14	61.56
Observations	354977	196374	354982	196397

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: U.S. Department of Education EDFacts and NCES Common Core of Data

Notes: This table shows the dose response estimation of school lunch nutritional changes on proficiency rates in schools by grade type. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2009-10 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates in Panels 1B and 2B are weighted by the total number of students tested at the school-level. All specifications include campus, year, and state fixed effects.

Table 3.11 Dose Response Estimation of School Lunch Nutrition on Grade Level Enrollment Percentages

	(1) Grade Enrollment
	3rd Grade
Dose x Post	0.17
	(0.12)
Mean	14.95
Observations	42708
	4th Grade
Dose x Post	-0.06
	(0.14)
Mean	14.94
Observations	42708
	5th Grade
Dose x Post	-0.03
	(0.16)
Mean	12.72
Observations	42708
	6th Grade
Dose x Post	1.01
	(0.65)
Mean	29.31
Observations	15370
	7th Grade
Dose x Post	-0.36
	(0.37)
Mean	32.24
Observations	15370
	8th Grade
Dose x Post	-0.57
	(0.39)
Mean	32.24
Observations	15370

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency

Notes: This table shows the dose response estimation of school lunch nutritional changes on enrollment rates in schools by grade level. The dose in this estimation is defined as the average school-level percent of students eligible for Free & Reduced Price Lunch (FRPL) in the pre-period (SY 2007-08 – SY2011-12). Post-treatment is equal to 1 in SY2012-13 and beyond. Standard errors are clustered at the campus level, and estimates are weighted by the total number of students at the school-level. Non-traditional public schools (including charters, non-profit private, juvenile institutions, etc.) and schools that do not participate in the National School Lunch Program in the year prior to treatment are excluded from this analysis. All specifications include campus and year fixed effects.

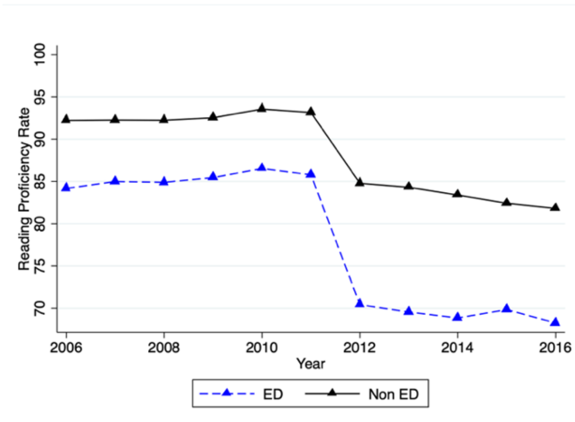
3.10 Appendix C: Texas Achievement Data Analysis

In both Texas Education Agency standardized testing regimes, elementary aged students have the option to take their exam in either English or Spanish, based on their language proficiency and what their school determines is the best fit exam delivery for their language preferences. In all reported results, the proficiency rates represent proficiency for *all* students taking the Texas state exams, in both language formats.

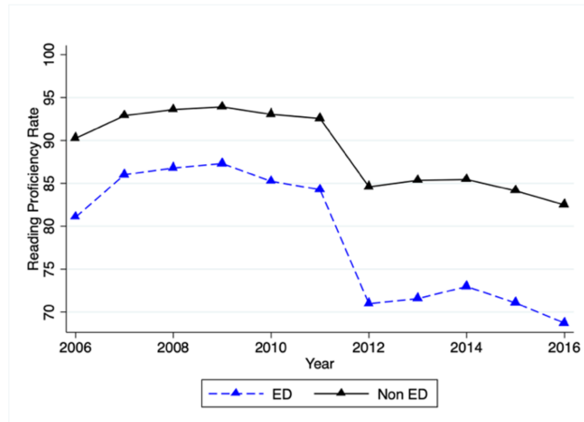
In the year of the testing regime change from STAAR to TAKS, SY 2011-12, proficiency rates drop and remain below the trend average during the TAKS regime. Figure 3.13 plots these time trends for elementary and middle school reading and math scores for students who are classified as economically disadvantaged as compared to students who are classified as non-economically disadvantaged. Figure 3.14 further plots these same time trends for schools with a treatment dose above the median as compared to schools with a treatment dose below the median. Figure 3.13 and 3.14 similarly display a sharp decrease in proficiency rates for all students under the new STAAR testing regime, but also indicate that these magnitudes appear stronger for economically disadvantaged students as compared to non-economically disadvantaged students, and for schools with a higher treatment dose percentage of economically disadvantaged students.

To formally test whether the change from TAKS to STAAR testing regimes disproportionately impacted the proficiency rates of the treatment group, I run a basic difference-in-differences specification where treatment is equal to 1 for the student subgroup "economically disadvantaged" and treatment is equal to 0 for the student subgroup "non-economically disadvantaged". I utilize data from school years 2005-06 through 2011-12, where 2011-12 is the first year within the new STAAR testing regime. Post-treatment is equal to 1 only in the year 2011-12, and zero otherwise. I weight by number of students tested, and include school fixed effects. Table 3.12 shows the regression results from this specification, which indicate that across the board in all grades and in all subjects, economically disadvantaged students' proficiency scores dropped statistically significantly more than the scores of their non-economically disadvantaged peers.

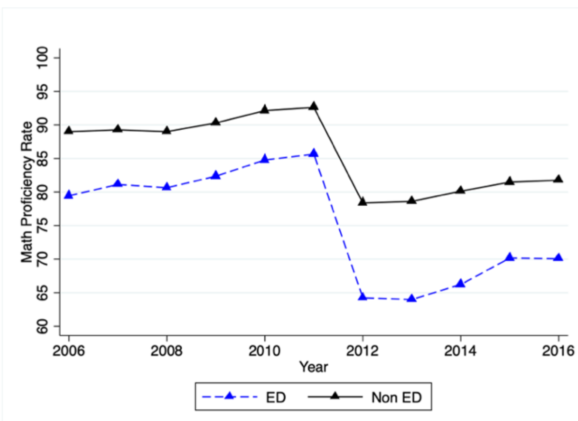
Figure 3.13 Weighted Time Trends in Proficiency Rates for Economically Disadvantaged Students vs Non-Economically Disadvantaged Students in Texas



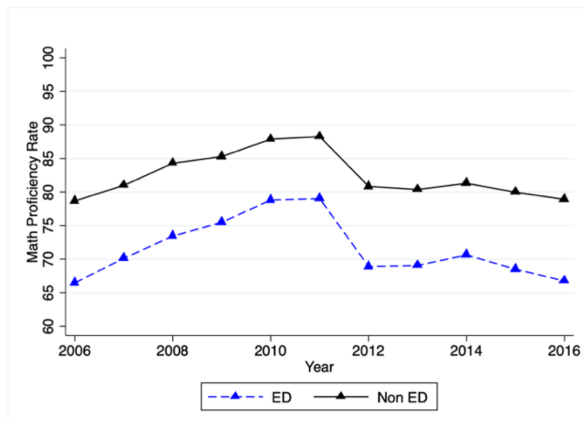
(a)



(b)



(c)

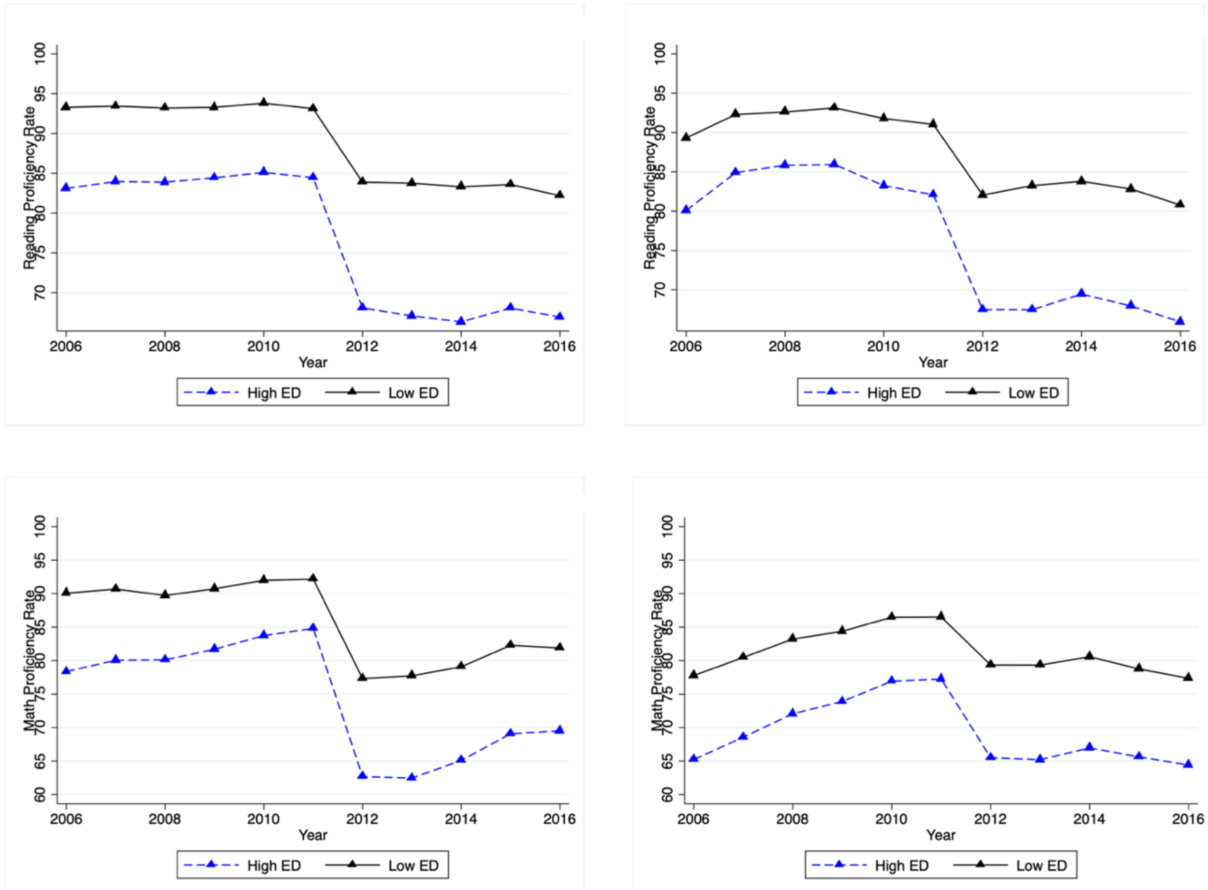


(d)

Data Source: Texas Education Agency

Notes: This figure shows the annual mean proficiency rates at the school level in the state of Texas from SY 2007-08 through SY 2015-16 for reading and math standardized test scores. Panel (a) plots the annual mean reading proficiency rates in elementary schools for economically disadvantaged students and non-economically disadvantaged students, panel (b) plots this for middle schools, panel (c) plots the annual mean math proficiency rates in elementary schools for economically disadvantaged students and non-economically disadvantaged students, and panel (d) plots this for middle schools.

Figure 3.14 Weighted Time Trends in Proficiency Rates for schools with a dose above vs. below the median in Texas



Data Source: Texas Education Agency

Notes: This figure shows the annual mean proficiency rates at the school level in the state of Texas from SY 2007-08 through SY 2015-16 for reading and math standardized test scores. Panel (a) plots the annual mean reading proficiency rates in elementary schools with an above median dose vs. schools with a below median dose, panel (b) plots this for middle schools, panel (c) plots the annual mean math proficiency rates in elementary schools with an above median dose vs. schools with a below median dose, and panel (d) plots this for middle schools. The "dose" is defined as the average percent of a campus that qualifies for Free or Reduced Price Lunch (FRPL) in the years prior to treatment (SY 2005-06 – SY 2011-12). Schools with a baseline average percentage of FRPL eligible students above the statewide median are classified in this figure as being "above median" and are represented by the solid black line. Schools with a baseline percentage of FRPL eligible students below the statewide median are classified in this figure as being "below median" and are represented by the blue dashed line.

Table 3.12 Difference in Differences Estimation of TAKS to STAAR testing regime change on Economically Disadvantaged student proficiency rates

	(1) Reading	(2) Math
	PANEL A: 3rd Grade	
Dose x Post	-0.12***	-0.08***
	0.00	0.00
Mean	0.89	0.81
Observations	50948	51010
	PANEL B: 4th Grade	
Dose x Post	-0.06***	-0.09***
	0.00	0.00
Mean	0.83	0.82
Observations	50493	50578
	PANEL A: 5th Grade	
Dose x Post	-0.09***	-0.08***
	0.00	0.00
Mean	0.88	0.88
Observations	47082	47092
	PANEL B: 6th Grade	
Dose x Post	-0.11***	-0.02***
	0.00	0.00
Mean	0.88	0.80
Observations	28524	28511
	PANEL A: 7th Grade	
Dose x Post	-0.05***	-0.04***
	0.00	0.00
Mean	0.83	0.75
Observations	24131	24103
	PANEL B: 8th Grade	
Dose x Post	-0.09***	-0.02***
	0.00	0.00
Mean	0.90	0.78
Observations	24290	24171

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Data Source: Texas Education Agency

Notes: This table estimates the equation $Y_{st} = \alpha + \beta ED * Post2011 + \gamma ED + \delta Year + X_{st}$ where Y_{st} is equal to proficiency rates for student group *s* at time *t* and Post2011 is equal to 1 in SY 2011-12 and zero otherwise. *ED* is equal to 1 if the student group is "economically disadvantaged" and 0 if the student group is "non-economically disadvantaged". The estimates are weighted by enrollment quantity and the specification includes campus fixed effects, with standard errors clustered at the campus level. These specifications include data spanning school years 2005-06 through 2011-12 only.