



2016

School Discipline, Law Enforcement, And Student Outcomes

Amanda Scanlon

University of Pennsylvania, scull@gse.upenn.edu

Follow this and additional works at: <https://repository.upenn.edu/edissertations>



Part of the [Education Policy Commons](#)

Recommended Citation

Scanlon, Amanda, "School Discipline, Law Enforcement, And Student Outcomes" (2016). *Publicly Accessible Penn Dissertations*. 2570.
<https://repository.upenn.edu/edissertations/2570>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/2570>
For more information, please contact repository@pobox.upenn.edu.

School Discipline, Law Enforcement, And Student Outcomes

Abstract

Zero tolerance discipline policies have lost favor in recent years due to concerns that they reduce offending students' classroom time, beget further misconduct, and decrease student engagement. In the same vein, school police programs often associated with zero tolerance policies are frequently charged with increasing student involvement in the criminal justice system. Despite much discourse on the topic of school discipline, few studies have rigorously examined causal links between zero tolerance, school-based law enforcement, and student outcomes. This dissertation examines the effects of dismantling zero tolerance and reducing police officer presence in Philadelphia schools on school-level rates of student misconduct, administrative responses, and academic achievement. Quasi-experimental methods applied to the data include propensity score matching, generalized difference-in-differences analysis, comparative interrupted time series analysis, and fuzzy regression discontinuity analysis. Results suggest that dismantling zero tolerance did not affect school arrests rates or the rates of incidents involving law enforcement; and while transfers and expulsions decreased two years after the policy change, truancy increased. Limiting school police officer staff positions may have led to declines in the rates of incidents involving law enforcement and arrests, but the evidence is weak due to low statistical modeling power. Areas for future work are discussed.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Education

First Advisor

Andrew C. Porter

Subject Categories

Education Policy

SCHOOL DISCIPLINE, LAW ENFORCEMENT, AND STUDENT OUTCOMES

Amanda J. Scanlon

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

Andrew C. Porter, Professor Emeritus of Education

Graduate Group Chairperson

J. Matthew Hartley, Professor of Education

Dissertation Committee:

Andrew C. Porter, Professor Emeritus of Education

Paul A. McDermott, Professor of Education

Robert F. Boruch, Professor of Education

SCHOOL DISCIPLINE, LAW ENFORCEMENT, AND STUDENT OUTCOMES

COPYRIGHT

2016

Amanda J. Scanlon

ACKNOWLEDGEMENT

I extend my heartfelt appreciation to my advisor and mentor Andy Porter for his invaluable wisdom, guidance, and encouragement throughout my career at Penn. Thank you for inspiring me to carve my own path through graduate school while providing me insightful counsel and advocating for me along the way. I am also wholeheartedly grateful for the unwavering support of Robert Boruch, Paul McDermott, and Laura Desimone. Each of you offered me a rich learning environment, an astute soundboard, and, most often, a welcome smile.

This dissertation would not have been possible without the love and support of my family and friends. Thank you for your curiosity; thank you for your advice; thank you for your time. And little Ellie—thank you for sharpening my blurry-eyed focus by providing me a new purpose in this and all my endeavors.

ABSTRACT

SCHOOL DISCIPLINE, LAW ENFORCEMENT, AND STUDENT OUTCOMES

Amanda J. Scanlon

Andrew C. Porter

Zero tolerance discipline policies have lost favor in recent years due to concerns that they reduce offending students' classroom time, beget further misconduct, and decrease student engagement. In the same vein, school police programs often associated with zero tolerance policies are frequently charged with increasing student involvement in the criminal justice system. Despite much discourse on the topic of school discipline, few studies have rigorously examined causal links between zero tolerance, school-based law enforcement, and student outcomes. This dissertation examines the effects of dismantling zero tolerance and reducing police officer presence in Philadelphia schools on school-level rates of student misconduct, administrative responses, and academic achievement. Quasi-experimental methods applied to the data include propensity score matching, generalized difference-in-differences analyses, comparative interrupted time series analyses, and fuzzy regression discontinuity analyses. Results suggest that dismantling zero tolerance did not affect school arrests rates or the rate of incidents involving law enforcement; and while transfers and expulsions decreased two years after the policy change, truancy increased. Limiting school police officer staff positions may have led to declines in the rates of incidents involving law enforcement and arrests, but the evidence is weak due to low statistical modeling power. Areas for future work are discussed.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iii
ABSTRACT.....	iv
LIST OF TABLES	viii
LIST OF ILLUSTRATIONS	ix
CHAPTER 1: INTRODUCTION	1
The School District of Philadelphia	3
The Current Study	7
Organization of the Dissertation.....	9
CHAPTER 2: REVIEW OF THE LITERATURE	11
Theoretical Perspectives on Misbehavior and Crime.....	12
Youth Behavior and Zero Tolerance.....	17
Path from Zero Tolerance to Criminalization	21
Research Linking School Discipline and Student Outcomes.....	24
Research on Zero Tolerance	25
Research on School-based Law Enforcement	30
CHAPTER 3: DISCIPLINE POLICY IN PHILADELPHIA.....	33
Recent History of the SDP	34
Zero Tolerance in the SDP	36
School-based Law Enforcement	42
CHAPTER 4: METHODOLOGY	46

Primary Research Questions	46
Boundaries of the Research Questions	50
Data	55
Reliability and Validity	59
Methods	70
Design	71
Sample	77
Analyses.....	81
CHAPTER 5: PENNSYLVANIA SCHOOL DEMOGRAPHICS AND BEHAVIOR .	102
The Pennsylvania Context: Trends and Demographics	103
Research Question (1) School Groups	108
Demographic Baseline.....	108
Behavioral Trends.....	110
Research Question (2) School Groups	112
Demographic Baseline.....	112
Behavioral Trends.....	114
Propensity Score Matched Samples	116
Evaluation of Matched RQ1 Groups	116
Evaluation of Matched RQ2 Group.....	123
CHAPTER 6: REMOVAL OF ZERO TOLERANCE AND BEHAVIOR.....	125
Preliminary Model Selection.....	126
Results	130

Discussion	139
CHAPTER 7: SCHOOL POLICE OFFICER STAFFING AND BEHAVIOR	142
Preliminary Model Selection.....	143
DD Model Type Selection.....	143
FRD Assignment Variable and Cutoff Value.....	145
FRD Framework, Group Differences, and Models	148
Results	154
Discussion	158
CHAPTER 8: CONCLUSION	161
Limitations	163
Future Work	165
APPENDIX A: ADDITIONAL TABLES.....	167
APPENDIX B: ADDITIONAL FIGURES	204
APPENDIX C: SDP DISCIPLINE POLICY, 2002-2003 CODE OF CONDUCT.....	215
APPENDIX D: SDP DISCIPLINE POLICY, 2012-2013 CODE OF CONDUCT	223
APPENDIX E: CODEBOOK AND VARIABLE DEFINITIONS	235
BIBLIOGRAPHY.....	256

LIST OF TABLES

Table 1. <i>Raw Data Error Rates by Variable and Year</i>	65
Table 2. <i>Design Elements, Sampling, and Analytic Methods</i>	74
Table 3. <i>Pennsylvania School Demographics, 2013-2014 School Year</i>	104
Table 4. <i>RQ1 Purposive Sample Baseline Demographics</i>	109
Table 5. <i>RQ2 Purposive Sample Baseline Demographics</i>	113
Table 6. <i>Balance of RQ1 Propensity Score Matches</i>	117
Table 7. <i>Balance of RQ2 Propensity Score Matches</i>	122
Table 8. <i>AIC and BIC Values by Sample and Model Type</i>	129
Table 9. <i>RQ1 DD Treatment Estimates, by Sample</i>	131
Table 10. <i>RQ1 GDD Treatment Estimates, by Sample</i>	134
Table 11. <i>RQ1 CITS Treatment Estimates, by Sample</i>	136
Table 12. <i>Mean Differences in RQ2 ITT Samples at Cutoff</i>	149
Table 13. <i>RQ2 Treatment Estimates, by Analysis Type and Sample</i>	155
Table A1. <i>DD Regression Estimates, Purposive Phil. Charter School Control</i>	168
Table A2. <i>DD Regression Estimates, PSM Phil. Charter School Control</i>	170
Table A3. <i>DD Regression Estimates, Purposive Pennsylvania School Control</i>	172
Table A4. <i>DD Regression Estimates, PSM Pennsylvania School Control</i>	174
Table A5. <i>GDD Regression Estimates, Purposive Phil. Charter School Control</i>	176
Table A6. <i>GDD Regression Estimates, PSM Phil. Charter School Control</i>	178
Table A7. <i>GDD Regression Estimates, Purposive Pennsylvania School Control</i>	180
Table A8. <i>GDD Regression Estimates, PSM Pennsylvania School Control</i>	182
Table A9. <i>CITS Regression Estimates, Purposive Phil. Charter School Control</i>	184
Table A10. <i>CITS Regression Estimates, PSM Phil. Charter School Control</i>	187
Table A11. <i>CITS Regression Estimates, Purposive Pennsylvania School Control</i>	190
Table A12. <i>CITS Regression Estimates, PSM Pennsylvania School Control</i>	193
Table A13. <i>DD Regression Estimates, Purposive SDP School Control</i>	196
Table A14. <i>DD Regression Estimates, PSM SDP School Control</i>	198
Table A15. <i>FRD Intent-to-Treat Estimates</i>	200
Table A16. <i>FRD Treatment-on-the-Treated Estimates</i>	202

LIST OF ILLUSTRATIONS

<i>Figure 1.</i> Graphical Portrayal of DD Analysis	83
<i>Figure 2.</i> Graphical Portrayal of GDD Analysis	87
<i>Figure 3.</i> Graphical Display of DD, GDD, and CITS Analyses with Identical Dataset ..	93
<i>Figure 4.</i> Graphical Representation of Sharp Regression Discontinuity and FRD	98
<i>Figure 5.</i> Pennsylvania Trends in Student Outcomes, 2001-2002 to 2013-2014.....	106
<i>Figure 6.</i> Arrests per Incident Type.....	107
<i>Figure 7.</i> School-level Mean Student Outcomes by RQ1 Group and Year.....	111
<i>Figure 8.</i> School-level Mean Student Outcomes by RQ2 Group and Year.....	115
<i>Figure 9.</i> Balance Statistics of SDP and Philadelphia Charters, RQ1	118
<i>Figure 10.</i> Balance Statistics of SDP and Pittsburgh Schools, RQ1	119
<i>Figure 11.</i> Balance Statistics of SDP and Pennsylvania Schools, RQ1	120
<i>Figure 12.</i> Balance Statistics of Affected and Non-Affected Schools, RQ2.....	121
<i>Figure 13.</i> Outcome Variable Distributions, All PA Schools in 2012-2013	127
<i>Figure 14.</i> Density of Treatment and Control Groups by Assignment Variable, RQ2 ..	147
<i>Figure 15.</i> FRD Assignment Variable versus Baseline Outcome Measures, RQ2	150
<i>Figure 16.</i> Treatment Discontinuity at the Assignment Cutoff, RQ2.....	152
<i>Figure B1.</i> Predicted and Residual Values of RQ1 DD Analysis, Purposive Philadelphia Charter Control	205
<i>Figure B2.</i> Predicted and Residual Values of RQ1 DD Analysis, PSM Philadelphia Charter Control	206
<i>Figure B3.</i> Predicted and Residual Values of RQ1 DD Analysis, Purposive Pennsylvania Schools Control.....	207
<i>Figure B4.</i> Predicted and Residual Values of RQ1 DD Analysis, PSM Pennsylvania Schools Control.....	208
<i>Figure B5.</i> Cook's D Estimates of RQ1 DD Analysis, Purposive Philadelphia Charter Control	209
<i>Figure B6.</i> Cook's D Estimates of RQ1 DD Analysis, PSM Philadelphia Charter Control	210
<i>Figure B7.</i> Cook's D Estimates of RQ1 DD Analysis, Purposive Pennsylvania Schools Control	211
<i>Figure B8.</i> Cook's D Estimates of RQ1 DD Analysis, PSM Pennsylvania Schools Control	212
<i>Figure B9.</i> Predicted and Residual Values of RQ2 DD Analysis.....	213
<i>Figure B10.</i> Cook's D Estimates of RQ2 DD Analysis.....	214

CHAPTER 1: INTRODUCTION

Youth crime and misconduct have captured the attention of policymakers at all levels for the last thirty years. Since a rising “epidemic of youth violence” beginning in the 1980s (Office of the Surgeon General, National Center for Injury Prevention and Control, National Institute of Mental Health, & Center for Mental Health Services, 2001), law enforcement and school officials have sought to understand the root causes of youth misbehavior and establish policies to dissuade acts of violence and delinquency. Never before, however, has public attention focused so raptly on the virtues of such policies and their long-term effects on outcomes such as educational attainment and employment. Policies established in the wake of rising crime—calling for the mandatory punishment of offenders committing specific violent and non-violent crimes, such as drug trafficking—have come under increasing scrutiny for fostering high rates of incarceration and doing little to deter further crime and misconduct. President Obama has made history for the number of federal inmates incarcerated for minor crimes that he has pardoned throughout his tenure (Koren, 2016). Congress has taken steps to reform federal crime policy, introducing legislation to reduce minimum mandatory sentences for nonviolent crimes, permit judges more power to determine penalty on a case-by-case basis, and expand opportunities for rehabilitation (New York Times Editorial Board, 2016). This rollback of strict delineated consequences for minor offenses has seen bipartisan support, as many policymakers have come to believe that the cons of such policies outweigh the intended benefits.

Shifting attitudes regarding crime and punishment are not limited to the criminal justice realm. The education sector has seen parallel movement to and from such policies over the last two decades. Zero tolerance policies first entered school buildings in the 1990s as part of the Gun Free Schools Act, requiring schools receiving federal funding to expel any student bringing a weapon to school (*Gun-Free Schools Act*, 1994). Many states expanded upon the act, establishing policies that delineated specific consequences for specific acts of misconduct. In the years since, many schools and districts have relied on out-of-school suspensions, expulsions, and transfers to disciplinary schools both to address individual misconduct and to remove disruptive and violent students from the classroom. Strict school responses to misbehavior have led critics to argue that such policies ultimately do more harm than good, placing misbehaving students on a fast track toward school dropout and incarceration. This ‘criminalization’ of students—often referred to as the ‘school-to-prison pipeline’—is viewed both as a direct result of zero tolerance policies, and as a result of the increase in the presence of school police officers that has occurred in tandem with the enactment of zero tolerance policies. In light of such critiques, many states and districts have recently begun to dismantle the zero tolerance policies once favored by administrators.

Despite continued activity in the realm of school discipline policies, scant research has examined the effects of zero tolerance policies on student behavior and outcomes, or on the effects of dismantling zero tolerance policies once in place. This dissertation addresses this gap in the literature. Two particular policy changes are examined here, both emerging from the School District of Philadelphia (SDP). First, the

SDP decided in August 2012 to replace its decade-old zero tolerance policy with one permitting more subjectivity in administrator response to misconduct and encouraging the use of in-school behavioral supports and resolutions. Second, the SDP opted not to fill some school police officer vacancies in the 2013-2014 school year, requiring 28 schools to share police officers part-time. The SDP operates its own police department of nearly 400 officers, a number which rose concurrently with the implementation of its zero tolerance policy throughout the 2000s. The decision not to fill the vacancies came as a result of budgetary concerns.

The two policy changes—one a matter of the district’s code of conduct, the other an administrative fiscal decision—provide the opportunity to examine the effects of dismantling a zero tolerance policy in a large, urban district. Because the policy changes occurred at different times, they also provide the opportunity to examine their effects in isolation, i.e., removing the zero tolerance disciplinary policy while maintaining the level of employment of school-based law enforcement, and limiting the employment of school-based law enforcement while holding constant student discipline policy. As districts and states across the country look to enact disciplinary policies that address in-school misconduct while supporting the achievement and well-being of all students, this research can provide insights as to the impact of two aspects of disciplinary policies that are widely critiqued in current discourse.

The School District of Philadelphia

As of December 2015, the SDP operates 218 schools enrolling 136,000 students and employing 8400 teachers. Of the 218 schools, 149 operate as elementary schools

enrolling students generally between kindergarten and the sixth grade; 16 operate as middle schools enrolling students generally between the sixth and eighth grades; and 53 operate as high schools enrolling students generally between the ninth and twelfth grades. The vast majority of the students are those enrolled in traditional educational programs; just 2,500 students are enrolled in alternative education programs, and 300 students in virtual programs. Just over half of all students, or 51 percent, identify as black/African American; Hispanic/Latino students make up 20 percent of the student body, while white students comprise 14 percent and Asian students comprise 8 percent. Overall, 14 percent of students are identified as students with disabilities, and 10 percent of students are English language learners (School District of Philadelphia, 2015).

The SDP ranks among the 20 largest school districts in the country. The district's unique demographics, however, render the school district among the country's most disadvantaged. According to the most recent available year of federal data (the 2013-2014 school year), the SDP serves the third largest proportion of students receiving free or reduced-price lunch among the 20 largest districts in the United States. The district ranks first among its peers in its proportion of students with disabilities, fifth in its proportion of English language learners, and seventh in its overall proportion of minority students, measured as the proportion of black and Hispanic students combined (National Center for Education Statistics, 2015). The district also struggles academically. Among the approximately 500 other districts in Pennsylvania, the SDP reports the twelfth lowest rate of students achieving grade-level proficiency in reading in the 2014-2015 school year and the eighteenth lowest rate of students achieving grade-level proficiency in

mathematics in the 2014-2015 school year (author's calculations, based on Pennsylvania Department of Education, 2015c). Further, the SDP sees the fifth lowest daily attendance rate among its high school students and the twenty-ninth lowest cohort graduation rate as compared to other Pennsylvania districts (Pennsylvania Department of Education, 2015a).

Student and parent perceptions, however, reveal a more positive picture of the SDP. Surveys conducted by the district show that students and parents are mostly content with the instructional rigor and education provided by Philadelphia schools. Over half (58 percent) of students report that they “learn a lot” in their classes “most or all of the time,” while just 13 percent report that they do so “never” or “rarely” (School District of Philadelphia & University of Pennsylvania Graduate School of Education, 2016b). An overwhelming 91 percent of parents agree that “my child’s school meets the specific academic needs of my child;” and again, 91 percent of parents affirm that they are “pleased with the quality of education my child’s school is providing for my child” (School District of Philadelphia & University of Pennsylvania Graduate School of Education, 2016a). Educational surveys have consistently shown, however, that parents and students are apt to rate their own schools highly, even as they demonstrate less confidence in the education system nationally: a full 70 percent of public school parents give their own child’s school the grade of “A” or “B,” while just 19 percent give a grade of “A” or “B” to public schools “in the nation as a whole” (Richardson, 2015). These national trends suggest that the favorable perspectives of Philadelphians towards their local schools are perhaps rosier than warranted, ignoring endemic inadequacies apparent

when viewed objectively. Further, the response rates of the SDP student and parent surveys reach only about 10 percent of potential participants, and thus may not accurately represent the general perceptions of the population. What is clear is that while some parents and students are content with their children's education in the SDP, district performance metrics suggest good cause for discontent among those same populations.

While academic measures are the most widely available metrics for comparing district and school performance across the state and country, other educational aspects—such as school safety and student misconduct—are also important in weighing the quality of a school or school district. For now, no distinction is made between rates of student misbehavior and rates of punishment; this topic is discussed further in Chapter 4.

Assuming some correlation among the two, however, the SDP sees higher rates of severe consequences for misconduct than any other district in the state of Pennsylvania. The out-of-school suspension rate in the SDP was three times higher than that in the rest of the state in the 2008-2009 school year. Further, students in the SDP were arrested at a rate 3.5 times greater than the average across the state in the 2008-2009 school year, while the number of school police, security, and resource officers per 1,000 students in the SDP was ten times higher than in the rest of the state (Youth United for Change, Advancement Project, & The Education Law Center, 2011). School misbehavior, crime, and punishment has plagued the district for years, as described in a seven-part 2011 expose published by the *Philadelphia Inquirer*, which documented serious acts of violence and aggression on the part of students, and acts of both oversight and overreaction on the part of administrators (Sullivan, Snyder, Graham, & Purcell, 2011).

Beyond student behavior on school grounds, the district operates in a community facing widespread crime and unemployment, both of which affect and are affected by students' educational experiences. Only 58 percent of high school graduates in Philadelphia were employed in 2006, while just 38 percent of high school dropouts were employed. Approximately 1 percent of adults 18 to 60 were institutionalized in 2006. This figure rises as levels of education decline: about 2 percent of those with only a high school diploma or GED were in jail, compared with 5 percent of high school dropouts (Fogg, Harrington, & Khatiwada, 2009). When surveyed as to why they decided to drop out of high school, those reporting a discipline-related motive gave reasons such as multiple suspensions (32 percent), expulsion or transfer (30 percent), and punishment for fighting (25 percent) (Youth United for Change, 2011).

Given concerns for student safety and well-being in the SDP, and the link between student behavior in schools and in the community, research on the topic of discipline policy is crucial information for policymakers and administrators, particularly as they work to find a balance between containing misconduct and supporting those students exhibiting behavioral problems. Of key importance is not just the effects of such policies on student misconduct, but also on the nature and severity of the punishments put in place by administrators in response to misconduct, and on the academic performance of all students, whether misbehaving or not. These issues are explored in this dissertation.

The Current Study

This paper provides evidence on issues of school discipline policy in a large, urban school district by addressing the following research questions:

- (1) To what extent did the dismantling of the SDP's zero tolerance discipline policy affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?; and
- (2) To what extent did the reduction of police officer presence in particular schools in the SDP affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?

Severe student misconduct as defined in this dissertation represents those student actions which harm other individuals or place them in danger; those which entail possession of controlled substances or weapons; or those which could potentially be considered criminal acts. For example, acts of violence would fall into this category, whereas classroom disruption would not. Administrative reactions as defined in this dissertation are the consequences outlined by the district as potential responses to such behaviors, including out-of-school suspensions, disciplinary transfers, expulsion, and arrest.

These questions are explored with the use of quasi-experimental methods. Given that each of these policy decisions took place in a natural, nonexperimental setting, they raise complex issues of isolating the effects of the two policy 'treatments' and identifying statistically equivalent control groups against which SDP schools can be compared.

To address Research Question (1), the paper employs propensity score matching to sample non-treated students, paired with methodological strategies of difference-in-differences analyses and comparative interrupted time series analyses. To address

Research Question (2), the paper employs propensity score matching paired with difference-in-difference analyses, as well as a fuzzy regression discontinuity approach. Propensity score matching is used to approximate baseline equivalence of treated schools and comparison schools; difference-in-differences, comparative interrupted time series, and fuzzy regression discontinuity are used to eliminate alternative confounding effects that could bias the estimate of the treatment. The strategies are used collectively to provide a multifaceted picture of potential treatment effects.

Organization of the Dissertation

This study examines the history and research on zero tolerance policies and school-based law enforcement, the context of discipline policy in the SDP in terms of its past and present history, and the effects of two discipline policy changes within the SDP. Chapter 1 introduces the study and its significance. Chapter 2 explores theoretical perspectives on misbehavior and crime in order to illuminate how shifting trends in youth behavior led to the emergence of zero tolerance policies. The discussion then turns to critiques of zero tolerance, namely, the purported link between such policies and the ‘criminalization’ of students, before exploring the current body of research on the effects of zero tolerance and school-based law enforcement. Chapter 3 explores the SDP’s history of discipline and zero tolerance policy since the early 2000s, with particular attention to changing leadership and district management. Chapter 4 describes the methodology driving the quantitative research of the study, outlining the research questions, the data used in the study and their reliability and validity, and the quasi-experimental design, sample, and analyses applied to address the research questions.

Chapter 5 presents a descriptive analysis of student and administrator behavior across Pennsylvania and within the SDP since the early 2000s. Chapter 6 presents results of analyses examining the effects of the SDP dismantling its zero tolerance policy in 2012. Chapter 7 present results of analyses examining the effects of the SDP reducing police officer staff at a number of its schools in the 2013-2014 school year. Chapter 8 concludes with a review of the research questions, a summary and discussion of the findings, and recommendations for future study.

CHAPTER 2: REVIEW OF THE LITERATURE

Zero tolerance policies, while examined here in the educational sector, first gained traction outside of the educational context amid a complex interplay of shifting national demographics, increasing levels of crime associated with drug and weapon use, and evolving thinking around behavioral theory and legislative best practices at the federal, state, and local levels. Debates over zero tolerance and its legacy in schools run parallel to those currently contested in the larger domestic political realm in regards to the vestiges of the “get tough” approaches of the 1990s and the high rates of incarceration for minor offenses seen today (see, for example, Brooks, 2015). Understanding zero tolerance policies necessitates delving into the historical context in which they rose to popularity, as well as the theories of behavior which both support and undermine the policies’ objectives.

This chapter begins with a discussion of select theories of behavior and crime which have shaped how policymakers—and school administrators, parents, and advocacy organizations—think about how students might respond to various types of discipline policy. The chapter then turns to a brief history of trends in youth crime and in-school behavior which prompted school systems to adopt policies under the umbrella of zero tolerance and boost the number of school-based law enforcement in the mid-1990s. This leads into a narrative of the more recent critiques of zero tolerance policies, primarily the line of reasoning that these policies have led to undue incarceration and exclusion of students. Finally, the chapter examines the current research base on zero tolerance

policies and school-based law enforcement, and the effects of these policies on student behavior and outcomes.

Theoretical Perspectives on Misbehavior and Crime

As with nearly all human behaviors, student misbehavior within school walls—ranging from minor acts of delinquency to serious acts of violent crime—occur within social and environmental contexts that both drive and influence individual decision-making. Whether acting consciously or unconsciously, students act in certain ways based on the unique value that each assigns to a given act; this value is a weighting of potential profits against potential consequences. Thus at the heart of all school discipline policies is the attempt to craft a policy under which students perceive a greater benefit resulting from following rules of good behavior than from misbehaving. Discipline and behavior within the schoolhouse context is particularly complex, however, given the overlapping frameworks of hierarchical school authority; social milieu of students peers, families, and communities; and individual responsibility that form the scaffolding of the school system. While a number of theories have emerged in crime literature to explain the conscious and unconscious thought processes that may lead an individual to act out, some are more or less relevant to the behavior of youth in schools. More likely, many of these theories interact to different degrees based on the place and time unique to each student.

The theories of delinquency and criminal behavior that are particularly salient in the context of schooling and discipline are *risk aversion*, *social learning*, *labeling*, and *strain*. Theories of risk aversion and social learning may lead school or district authorities to implement strict discipline policies delineating clear and potentially severe

consequences for defined misbehaviors; theories of labeling and strain may lead others to implement discipline policies that aim to support and encourage positive behaviors while engendering atmospheres of trust and respect. These two approaches are not necessarily mutually exclusive, nor are the different theories of behavior. Still, they help illuminate why certain students might misbehave and why certain schools and districts choose to respond—and attempt to preempt such behaviors—in certain ways. The four theories are not exhaustive of all theories of behavior which might illuminate student behavior in schools and school administrators’ disciplinary responses, but they are perhaps most salient; each is broadly introduced below.

The basic notion underlying the theory of risk aversion is that individuals act in ways that they perceive to maximize their expected benefits and minimize their expected costs (Becker, 1968). Costs and benefits here refer not (necessarily) to monetary profits, but to any number of potential outcomes such as reputation, liberation, punishment, etc. Punishment is often weighed as both the likelihood of punishment and the severity of punishment (Eide, 2000). Risk aversion theories fall under the umbrella of larger behavioral theories that view rational choice as the driver of individual behavior. A student in a classroom, for example, can choose to sit quietly or interrupt the teacher. His or her choice will depend on a rational weighting of the extent to which the potential benefits—peer reputation, empowerment, delay of instruction, etc.—outweigh the potential consequences—time out, embarrassment, a note home, etc. Not all individuals perceive the same rationality of each choice; individual characteristics, such as low self-control, may lead some actors to different choices than others (Nagin & Paternoster,

1993). For example, research has shown that certain populations, such as convicted criminals, are more concerned with likelihood of punishment than with severity of punishment, while the general population is attuned to both (Block & Gerety, 1995). Still, from this perspective, a clear delineation of potential consequences of a given action is a reasonable approach to prevention of misbehavior and crime. In fact, most deterrence policies, such as zero tolerance policies, are premised on this theory of behavior (Akers, 1990). The expectation of student risk aversion is thus salient in the establishment of many school discipline policies.

While the theory of risk aversion posits that an individual's expectation of future benefits and costs shapes his or her decisions, the theory of social learning posits that past and present rewards and punishments influence an individual's proclivity to misbehave (Akers, 1973; Akers & Jensen, 2011). In other words, stimuli from one's environment—such as positive reinforcement from one's peers, negative reinforcement from authority figures or family members, or vice versa, etc.—condition an individual to behave in a certain manner. This is not entirely separate from the theory of risk aversion, and indeed some experts subsume one theory under the other (Akers, 1990). But beyond more simplistic concepts of trial and error, social learning helps explain why certain individuals might act in certain ways in some situations, such as the home, but not others, such as the classroom (Conger, 1976). Social norms and perceived definitions of good and bad behavior play a significant role in the theory, as these are reinforced over time through interaction with one's environment (Akers, Krohn, Lanza-Kaduce, & Radosevich, 1979). Social learning theory is supported by research suggesting that social

norms of one's peers, family, and authority figures are closely related to one's likelihood to misbehave or commit crime (Akers et al., 1979; Baron, 2003). Theories of social learning shape discipline policies that seek to take command of schoolhouse discourse on good and bad behavior; resulting policies often aim to establish clear norms surrounding defined behaviors, positively reinforcing good behavior through reward and supports, and negatively reinforcing bad behavior through punishment and loss of reward.

On the other hand, the theory of labeling maintains that rulemaking and discipline enforcement itself can engender misbehavior and crime. This theory imagines a scaffolding of social groups in various positions of power; those at the top define behavior and impose rewards and sanctions, and in the process, those at the bottom are more likely to be identified as rule-breakers and delinquents. At the same time, such labeling encourages those at the bottom to perceive themselves as rule-breakers and delinquents, perpetuating a counter-productive cycle (Paternoster & Iovanni, 1989). This transformation of personal identity occurs in particular once an individual person is stigmatized for having misbehaved or committed a crime; their likelihood to then continue misbehaving or committing crimes is greater (Bernburg, Krohn, & Rivera, 2006). Labeling can also operate to disrupt traditional developmental paths: a student sent to juvenile court is not only labeled a criminal in name, but consequently may be transferred to an alternative program or expelled altogether. This negative educational impact may affect future career attainment and earnings income and may lead the once-student to continue down a path toward crime (Bernburg & Krohn, 2003). Many school discipline programs that have emerged in the wake of zero tolerance policies explicitly

aim to avoid such ‘criminalization’ of students, i.e., the purportedly routine introduction of students to the criminal justice system and the permanent labeling of these students as criminals and delinquents. Instead, they aim to avoid such labeling and address behavioral issues through establishing norms and expectations with explicit student input and by employing within-school interventions—such as in-school suspensions, as opposed to out-of-school suspensions—when misconduct occurs.

Finally, strain theory suggests that individuals are compelled toward crime and misbehavior because of negative relationships and their characteristics—such as anger—and loss of opportunity. Strain occurs in situations in which failure is perceived or anticipated, positive reinforcement such as reward for good behavior is all but absent, and negative reinforcement such as punishment—or reward for bad behavior—is prevalent (Agnew, 1992). Negative relationships causing strain can include pressure from deviant peers as well as situations of abuse, victimization, or poverty (Baron, 2004). Even in the same scenarios, individual conceptions of strain can vary, as can the individual emotions and coping mechanisms generated by strain (Broidy, 2001; Elliott, Ageton, & Canter, 1979). This strain theory itself might not explain all variance in an individual’s proclivity to commit crime, but strain is a trigger without which an individual might not turn to misbehavior at all. Strain theory is particularly relevant in the context of a school district such as Philadelphia, which enrolls many low-income, minority students living in crime-ridden neighborhoods. The theory suggests that policies that aim to reduce misbehavior and crime must also recognize strains on the student body and work to alleviate them. Again, many discipline policies that have emerged in recent years emphasize school-wide

supports for students and behavioral interventions as opposed to policies perceived as increasing strain by stigmatizing and severely punishing students.

None of these theories perfectly explains why individuals, and students in schools in particular, might misbehave or commit crime, but they are useful in understanding the extent to which discipline policies might vary. Given that no single theory has emerged as the final word on misbehavior, the wide range in discipline policies across the country—and critiques leveled against them—should come as no surprise. Policymakers simply make decisions based on which theory, or combination of theories, seems most plausible and which overlap with policies that are popular and easily implementable. Further, as described below, the dearth of empirical research on successful versus unsuccessful school discipline policies exacerbates this issue. The shift from zero tolerance policies to policies of school-wide supports signals a larger shift in public discourse from frameworks of risk aversion and social learning to those of labeling and strain. Evidence presented below suggests why this shift occurred and how advocates of each theory can point to literature supporting his or her position.

Youth Behavior and Zero Tolerance

General approaches to school discipline have evolved over the last half century in response to growing student bodies and changing societal norms. Understanding the rise and fall of zero tolerance policies requires understanding the context in which zero tolerance policies came to prominence. Rapidly changing student bodies in the wake of the civil rights movement and subsequent desegregation efforts led school administrators in the 1960s and early 1970s to increase their use of out-of-school suspensions and

expulsions. These exclusion policies served as a quick fix to addressing individual misbehavior and removing such disruptions from the classroom. A number of lawsuits in the 1970s, however, placed new due process requirements on such exclusionary practices and prompted a shift away from such policies. In-school suspensions replaced out-of-school suspensions as common and accepted responses to student misbehavior (Adams, 2000).

School discipline practices shifted again following changing patterns of youth crime and delinquency. In a 2001 report, the Office of the Surgeon General wrote of an “epidemic of youth violence” beginning in the 1980s (Office of the Surgeon General et al., 2001). Both official administrative crime records and self-reports of victimization indicate a rise in crime throughout the 1980s: between 1983 and 1993, the report finds a 160 percent increase in youth homicide/manslaughter arrests and a 70 percent increase in youth violent crime arrests; paired with 40-50 percent increases in self-report incidents of assault, robbery, and violent crimes (Office of the Surgeon General et al., 2001). These increases, in many locales associated with rising gang and drug-related activity (Thompkins, 2000), led to the enactment of strict laws related to drug and weapon possession in the late 1980s and early 1990s (Skiba & Peterson, 1999).

School systems soon followed suit, establishing stern rules and enforcement policies surrounding issues of drug and weapon possession, but also extending into other areas of misbehavior as well, such as cigarette use and classroom disruption (American Psychological Association Zero Tolerance Task Force, 2006). This atmosphere of strict oversight was enshrined in federal law in the passing of the national Gun Free Schools

Act of 1994, signed by President Clinton and part of the larger Elementary and Secondary Education Act reauthorization of 1994, requiring states to impose one-year expulsion for student possession of a weapon on school grounds (*Gun-Free Schools Act*, 1994). Pennsylvania, for example, subsequently passed Act 26 in 1995 outlining automatic expulsion for weapons possession; many Pennsylvania school districts then built on this policy requirement and included additional behaviors and delineated consequences in their disciplinary codes (American Civil Liberties Union of Pennsylvania, 2015). These new zero tolerance policies allowed policymakers to sidestep requirements related to due process when excluding students from schools (Adams, 2000). Federal policies also encouraged the expansion of law enforcement in schools. Federal subsidies from the justice department helped pay for schools across the country to staff uniformed police officers for the first time in the 1990s (Cook, Gottfredson, & Na, 2009; Petteruti, 2011).

Evidence on student behavior in schools does not necessarily parallel the nationwide increases in youth crime in this period. The percentage of students reporting criminal victimization at school remained stable at 15 percent between 1989 and 1995; also stable were the proportions of students reporting injury or threatened injury at school, from as far back as 1976 through 1996 (Kaufman et al., 1998). Of course, these numbers may mask variation among different student groups. For African American students, the reported proportion of students injured with a weapon at school increased from 5 percent to 10 percent between 1983 and 1990 (Office of the Surgeon General et al., 2001). Whether or not actual rates crime and misbehavior were changing for different

groups, what is clear is that perception of declining school safety certainly grew in this time period. The proportion of students who feared of attack at school jumped by 50 percent between 1989 and 1995, while the proportion reporting avoiding one or more places at school jumped by 80 percent in those years (Kaufman et al., 1998).

In the wake of these perceived threats of heightened school violence, school districts imposed increasingly strict, punitive school discipline policies. Policymakers in part built on the theories of criminal behavior of risk aversion and social learning: the new discipline codes modeled on zero tolerance aimed to use deterrence to mitigate issues of misbehavior in school—increasing students’ perceived risk of misbehavior by outlining severe punishment—as well as to establish social environments of clear rules and regulations in which reprisal for misbehavior is known, expected, and accepted (Ewing, 2000). As zero tolerance policies spread throughout the mid to late 1990s and through the early 2000s, self-reports of school crime and misbehavior decreased by many metrics. The percentage of students reporting criminal victimization at school declined to 5 percent by 2003 and to 3 percent by 2013, though the percentage of students reporting injury or threat of injury with a weapon remained stable (Robers, Zhang, Morgan, & Musu-Gillette, 2015). The percentage of students who reported being in a physical fight at school decreased from 16 percent in 1993 to just 8 percent in 2013; in the same time period, the percentage of students who reported carrying a weapon at school decreased from 12 percent to 5 percent (Center for Disease Control and Prevention, 2014). These trends, however, were reflected in changing patterns of youth behavior outside of school as well. Between 1999 and 2008, juvenile arrest rates decreased by 10 to 30 percent

across the board for offenses such as aggravated assault, property crime, burglary, and theft; violent crimes and murders decreased by 9 percent, while drug abuse violations decreased by 7 percent (Federal Bureau of Investigation, 2009).

Whether or not these changes in youth crime and misbehavior both inside and outside of school were the result of or related to the introduction of zero tolerance policies is not clear; they may have simply occurred over the same timeframe that the new policies were enacted. The link between school discipline policies and crime—or crime rates—is often assumed in discourse, however, that draws ties between strict discipline, school law enforcement presence, and student criminalization. This topic is explored further below.

Path from Zero Tolerance to Criminalization

Like many school districts, Philadelphia moved away from its zero tolerance policy amid an atmosphere in which policy organizations and student advocacy groups increasingly pointed to a direct link between zero tolerance and criminalization, or the increased likelihood that students are labeled as criminals or delinquents and introduced into the criminal justice system. In many locales, the transition to more delineated, punitive discipline policies went hand-in-hand with a heightened presence of police in schools. Police entered school buildings as part of a wide variety of policies and programs, some subsidized by grants from the federal government and other organizations, some more locally based. A handful of districts including Philadelphia established their own police departments in the mid-1990s, while others partnered with local police departments to employ officers (Vera Institute of Justice, 1999). Many

districts established school resource officer programs that placed a greater emphasis on community building, mentoring, education, and conflict resolution than on crime containment and consequences for misbehavior. Such school resource officers may or may not operate in plain clothes and in some cases are not armed; still, they retain the authority to arrest or at minimum detain students until local law enforcement arrives (Petteruti, 2011; Shaw, 2004; Thureau & Wald, 2010).

The distinction between school police officers and school resource officers can be difficult to discern, particularly in aggregate administrative records tallying officer employment in schools. The National Center on Education Statistics reports that the percentage of schools using security guards and/or police officers as school security measures increased from 54 percent of schools in 1999 to 70 percent of schools in 2013 (Roberts et al., 2015). The Bureau of Justice Statistics reports that the percentage of local police departments deploying school resource officers grew from 30 percent in 1999 to 43 percent in 2003, and then declined to 38 percent in 2007 (Bureau of Justice Statistics, 2001; 2006; 2010). Together, these figures show that while the overall use of police officers in schools has increased since the late 1990s through the 2000s, the specific use of school resource officers has stalled somewhat. This may suggest that schools are ending their use of school resource officers as mentors and educators but continuing to employ school police for security reasons.

Regardless of the capacity in which schools are employing police officers, the rhetoric of advocacy organizations, student groups, and even researchers frequently links strict school discipline policies with increased presence of law enforcement on school

grounds and ultimately with criminalization of students. This linkage is often termed the ‘school-to-prison pipeline’, a term increasing in usage since the early 2000s (Advancement Project, 2010; Christle, Jolivette, & Nelson, 2005; Fowler, 2011; Gonzalez, 2012; Kim, Losen, & Hewitt, 2010; Wald & Losen, 2003). The term has become so common that a 2011 press release headline from the Department of Justice announced a new initiative to “respond to the school-to-prison pipeline” through programming aimed at introducing less punitive discipline policies (United States Department of Justice, 2011). The pipeline represents the belief that, through both direct and indirect paths, zero tolerance policies increase the likelihood of student contact with the criminal justice system: severe punishments for minor offenses push students out of school and increase the likelihood of student detainment or incarceration, thereby creating barriers to reentry into the traditional education system, encouraging further delinquency, and limiting future career prospects (Harper, 2011; Kim et al., 2010). The increased likelihood of detainment or incarceration is in some cases a direct result of increased police presence in schools (Klehr, 2009). In others cases, even if students are not introduced to the criminal justice system by the school itself, their exclusion from school through suspensions or expulsions limits their career potential and ultimately increases crime and incarceration rates (Darling-Hammond, 2006).

Much of the criticism of zero tolerance policies follows from theories of behavior such as labeling and strain: once students are identified as delinquents through harsh consequences of zero tolerance such as removal from the classroom, they carry that label both internally—seeing themselves as rule breakers—and externally—as others

categorize them as rule breakers and continue to exclude them based on this label. At the same time, these negative experiences and relationships with authority as a result of severe punishments encourage further disengagement and delinquency. A key aspect of this logic is the assumed negative consequences associated with exclusion from the classroom through suspensions, expulsions, and disciplinary transfers. For example, reports and policy papers on zero tolerance repeatedly point to research linking disciplinary practices such as suspensions and expulsions with negative student outcomes such as low student achievement and higher rates of incarceration (Advancement Project, 2010; American Civil Liberties Union of Pennsylvania, 2015; American Psychological Association Zero Tolerance Task Force, 2006; Youth United for Change et al., 2011). Much of the research literature in this area, however, relies on descriptive and correlational approaches rather than experimental or quasi-experimental analyses. The question remains whether these causal relationships play out in reality. This question is examined in the next section.

Research Linking School Discipline and Student Outcomes

Research on school discipline policy and student behavior—ranging from academic achievement to school crime—is limited (Astor, Guerra, & Acker, 2010; Cook et al., 2009). While researchers have begun to address the complex relationships between student behavior, achievement, and discipline policies, most research only addresses one or two connecting threads of a much larger web (Cornell & Mayer, 2010). One notable exception is Chen (2008), who articulates an intricate model of school crime and estimates associations among variables using structural equations modeling. Much of this

research, including the study by Chen (2008), does not utilize experimental and quasi-experimental methods to analyze variable relationships. In doing so, the current literature often fails to take context into account, that is, to eliminate the threat of omitted variables and properly isolate and identify valid causal effects. While much has been written on the topic of discipline policy and student behavior, this paper considers only those studies employing quasi-experimental methods and experimental methods as rigorous research which could conclusively demonstrate causal effects and form the foundation of a sound literature base on the topic. Of course, experiments testing the effects of various disciplinary consequences on students would likely be unethical; nevertheless, research in the absence of such experiments should, to the greatest extent possible, approximate experimental conditions through design and modeling of analyses. Below, this paper examines depth and quality of existing research on the many linkages among zero tolerance policies, school-based law enforcement, and student behavior.

Research on Zero Tolerance

Despite the swelling outcry against zero tolerance policies in the last decade, research conclusively demonstrating negative effects—or any effects—of zero tolerance policies is scarce. A 2006 task force on zero tolerance policies found “no reliable studies on the impact of zero tolerance on student behavior or school climate...conducted under any kind of controlled experimental conditions” (American Psychological Association Zero Tolerance Task Force, 2006). Other influential reports—most condemning zero tolerance, such as a 2010 report by the Advancement Project, a civil rights policy organization—conclude that associations alone between the presence of a zero tolerance

policy, higher rates of suspensions, and lower academic achievement are sufficient to denounce zero tolerance policies in their entirety (Advancement Project, 2010). At least the Advancement Project is conscientious in claiming only “associations” among these variables; other research on zero tolerance and negative student behaviors claims causality on the basis of descriptive and correlational research alone (see, for example, Fowler, 2007).

Where research on zero tolerance falls short of fashioning valid causal linkages between zero tolerance policy and student behavior, it relies on evidence linking suspensions and expulsions—often seen as cornerstones of zero tolerance policies—to negative student outcomes such as incarceration or lower achievement in order to provide evidence of the detrimental effects of such policies. Policy reports on zero tolerance often draw on research suggesting that suspensions predict a number of detrimental outcomes, including dropout (Arcia, 2006; Costenbader & Markson, 1998; Kim et al., 2010); lower academic achievement (Arcia, 2006; Brown, 2007); lower rates of on-time graduation (Mendez, 2003); negative socio-/emotional wellbeing (Brown, 2007); increased likelihood of contact with the criminal justice system (Costenbader & Markson, 1998; Fabelo et al., 2011; Klehr, 2009); and higher rates of future suspensions (Mendez, 2003). In the same vein, zero tolerance policies are also condemned on the assumption that they are linked to greater rates of absenteeism and contact with the criminal justice system; research has linked absenteeism to low academic achievement (Gottfried, 2009) and dropout (Balfanz, Herzog, & Mac Iver, 2007); and court appearance to lower graduation rates (Sweeten, 2006; Vermeire, Merluzzi, & Ridolfi, 2013). Note that not one of the

studies cited here used experimental or appropriate quasi-experimental methods, instead relying primarily on correlation and simple regression analyses to support their conclusions that one variable predicts another or is linked to another. While such non-experimental research certainly suggests relationships among variables and helps direct future lines of inquiry, it cannot answer questions as to the effects of student exclusion from the classroom—whether through suspensions, absenteeism, or incarceration—much less answer questions as to the effects of zero tolerance policies.

Another area of research often summoned in discussions of the efficacy of zero tolerance policies is the relationship between school climate and student behavior. School climate here refers not to the organizational climate of a school—a combination of the material resources, size, and background characteristics of individuals that help define a school—but to the social climate of a school, including the rules of operation and the norms, belief systems, and relationships of individuals within the school (Anderson, 1982). Zero tolerance explicitly identifies and condemns poor behavior, and then defines and broadcasts specific consequences for such behavior. Critics of zero tolerance, following the labeling and strain theories of behavior and crime, argue that such policies strain the relationships within schools and create a climate of mistrust and fear, potentially engendering more misbehavior. At minimum, researchers argue that school climate should be a central consideration in literature on the motivations of student behavior as it represents the milieu within which behavior is conceived and enacted (Astor et al., 2010).

School climate research has grown out of studies suggesting that school-level factors are closely related to individual behavior. For example, Christle et al. (2005) find significant relationships between individual delinquency and school-level factors such as suspension rates, average academics, and dropout rates. Research has suggested that just one disruptive peer in a class of 25 students can reduce future earnings potential of an individual student by an average 3 to 4 percent (Carrell, Hoekstra, & Kuka, 2016). Related lines of inquiry move further outside the schoolhouse and examine larger constructs such as neighborhood context and social strata norms, also finding strong associations between these and individual behavior such as crime (Sampson, Morenoff, & Raudenbush, 2005). The strongest research in this field has exploited student assignment to schools through randomized lotteries to form valid causal estimates. Cullen, Jacob, and Levitt (2006) find that assignment to schools of higher average achievement and attainment reduces the individual likelihood of misbehavior; in the same vein, DeAngelis and Wolf (2016) find that enrollment in private school through a voucher system leads to lower rates of criminal activity.

Research specifically focusing on school-level disciplinary practices and their relationship with student behavior is less robust, at least in the ability to draw causal conclusions. Some studies address factors more intimately related to zero tolerance: Skiba et al. (2014) find that schools' whose administration is more favorable to consequences entailing school exclusion see more misbehavior than those more favorable to prevention strategies. Others take a broader approach that could support arguments both for and against zero tolerance. For example, research shows negative associations

between perceived respect for students and fairness of rules with misconduct (Gottfredson, Gottfredson, Payne, & Gottfredson, 2005; Welsh, 2000). Critics could argue that zero tolerance policies do not communicate respect for students or fairness of rules, while advocates could argue that such norms are exactly what zero tolerance intends to communicate.

Much recent research and policy discussions have focused on school-wide behavioral support interventions, often viewed as an alternative to zero tolerance policies. These programs vary widely, but generally denounce the punitive aspects of zero tolerance and instead aim to build positive, inclusionary responses to misconduct that do not remove students from the classroom or at least the school (Gonzalez, 2012; Klehr, 2009). Meta-analyses of studies examining the impacts of school-wide positive behavioral support programs find promising results: these programs on average decrease not just disruptive behavior but also violent and aggressive behavior (Hahn et al., 2007; Wilson & Lipsey, 2007). Of course, this research does not de facto demonstrate that zero tolerance policies have the opposite effect, or that zero tolerance policies do not promote the same types of behavioral supports. Research on zero tolerance is simply less prevalent and broader as a result of the larger-scale implementation of district-level policies versus school-level programs. Overall, much of the research supporting arguments for and against zero tolerance policies is tangential to the enactment or removal of such policies themselves. Rigorous research on the effects of zero tolerance is largely non-existent, and related research can only begin to suggest patterns and relationships among related student behaviors.

Research on School-based Law Enforcement

More targeted research exists on the topic of school-based law enforcement, but much of the current literature focuses on specific programs that introduce police into schools rather than the district-wide increase in police presence in response to a discipline policy directive. Further, these studies tend to use descriptive and correlational analyses rather than experimental and quasi-experimental approaches to examine relationships among law enforcement presence and student behavior. A 2012 systematic review of research on police-based programs in schools found only a single study employing quasi-experimental methods that would earn the accolade of “acceptable with reservations” from the What Works Clearinghouse operated by the U.S. Department of Education (Petrosino, Guckenburg, & Fronius, 2012). The study by Theriot (2009) used a non-equivalent statistically controlled comparison group design to examine whether the addition of school resource officers into a school led to higher rates of arrests, among other outcomes; it found that schools with SROs saw higher numbers of arrest for disorderly conduct but not for other types of arrest (Theriot, 2009).

Examining the relationship between police presence in schools and the likelihood of student crime—whether measured by rates of arrest, incarceration, calls for service, or student-reported criminal acts—is a common line of inquiry in the literature. Counter to Theriot (2009), research employing comparison groups generally suggests a null relationship of police presence on student crime (Brady, Balmer, & Phenix, 2007; Katz, Schaefer, & Uchida, 2002; Maguire, Solomon, & Uchida, 2003; Schreck, Miller, & Gibson, 2003), while those studies relying on correlational analyses generally suggest a

positive relationship of police presence on student crime (Gonzalez, 2012; Mayer & Leone, 1999; Na & Gottfredson, 2013; Vermeire et al., 2013). Of course, such correlational analyses may simply work to highlight that schools with more minority students and more socio-economically disadvantaged students, presumably more urban schools with higher rates of crime to begin, see higher rates of police presence (Verdugo, 2002). Other research examines the impact of police presence in schools on the rates of student suspensions, absenteeism, and perceptions. Studies generally show a positive relationship between the presence of police and rates of suspensions and expulsions (Brady et al., 2007; Vermeire et al., 2013) or a null relationship (Maguire et al., 2003). One exception is Johnson (1999b), which finds a decrease in suspensions for misbehavior, though the study analysis relies on descriptive aggregation of survey data. Bowles, Reyes, and Pradiptyo (2005) find a decrease in truancy associated with an increase in police presence. Finally, McKay, Covell, and McNeil (2006) find that a partnership to promote school officers as role models was associated with a positive belief in the efficacy of the program, though it had no association with subsequent student behavior. Bhabra, Hill, and Ghate (2004), on the other hand, find that a school-based police program was not associated with student attitudes toward and perceptions of police officers.

Beyond the non-experimental nature of these studies on school-based law enforcement, research in this area is hampered by limited data and potential unreliability of the scant data that are collected (Thurau & Wald, 2010). The validity of the data is also in question; for example, which is a more valid measure of student misbehavior—student

self-report or school suspension rates? Or is neither truly representative of the breadth and depth of student misbehavior? These questions are particularly pertinent to crime versus arrest rates—do arrest rates accurately represent criminal activity, or do they over/underestimate actual criminal behavior? Any researcher must work within the limitations of the data available to be generated and collected; still, these topics are explored in the Methodology section of this paper.

• • •

While zero tolerance policies have seen both ardent support and vitriolic critique since their zenith in the 1990s and 2000s, the literature base provides little overall evidence on the effects and legacy of such policies in schools. Current research also falls short of providing strong evidence on the effects of school-based law enforcement. Much of the limitation in the current literature is not due to a scarcity of writing on the topic, but a lack of studies employing appropriate, rigorous methods that would result in valid estimates of the effects of these policies. This paper addresses this shortfall by using quasi-experimental methods to approximate equivalent control groups and isolate other potential confounding effects. In doing so, this paper aims to fill important gaps in the literature, at a time in which districts across the country are rethinking their disciplinary strategies and looking for sound evidence of the successes and shortfalls of zero tolerance policies.

CHAPTER 3: DISCIPLINE POLICY IN PHILADELPHIA

Philadelphia presents a unique setting in which to study the effects of zero tolerance policy and school-based law enforcement, particularly given the school district's shift toward and away from zero tolerance since the early 2000s. This paper considers the impact of such policies on student behavior and academics specifically in the School District of Philadelphia (SDP). While Philadelphia also houses a large charter school sector, the charter school sector operates outside of the policies of the SDP and is used as an analytic comparison group in this study where possible (see Chapter 4: Methodology). While much national attention focused on the SDP in recent decades has centered around fiscal crises and state takeover of district management in the early 2000s, the SDP on the local level has continually developed its operational and disciplinary policies under a series of school superintendents, each with different goals and perspectives. This chapter begins with a brief history of the SDP, emphasizing the context in which policy decisions are made and remade in the district. The chapter then turns to discussion of the disciplinary policy of the SDP, focusing specifically on the establishment of the district's zero tolerance policy in the 2002-2003 school year, the dismantling of this policy ten years later, and the coinciding proliferation of school-based law enforcement in schools across the district. By leveraging two particular policy decisions regarding discipline policy and school police assignment made by the SDP in recent years, this study isolates and analyzes the effects on student behavior and

academics of both zero tolerance policies and policies which place police officers in schools.

Recent History of the SDP

Over the last 25 years, the School District of Philadelphia has continually evolved in its management structure, goals, and policy directives in an effort to overcome fiscal and academic crises. Perhaps the most significant shift in district operations occurred with state takeover of the district in the early 2000s; this decision continues to overshadow much of the rhetoric and debates that surround the school district today, as well as the various brands of school reform that have come and gone within the school district. The developments leading to the state takeover and the decisions made in its aftermath provide important context for understanding the policy decisions and constraints of the district as it operates today.

State takeover of the SDP occurred within an atmosphere of diminishing fiscal support for the district and an increasingly popular educational reform movement on the state and national levels. The state of Pennsylvania has traditionally funded a lower proportion of its schools than many other state governments, leaving districts to turn to local taxes to supplement state funds. In the early 1990s, the state legislature froze the funding formula such that the state would no longer automatically provide additional funds to school districts based on their enrollment or other needs. At the time, the state already funded its schools at a lower proportion of their total revenue than 70 percent of other states (author's calculations, based on Johnson, 1996). The SDP, which saw increasing enrollments throughout the 1990s but persistently low property values—and

thus low local property tax revenues—was particularly affected by this decision and recorded significant budgetary shortfalls throughout the decade (Travers, 2003).

At the same time, state and national policy had begun to center around accountability and school reform in the 1990s, promoting ideas such as the annual assessment of student and school performance and the elimination of bloated bureaucracies. Pennsylvania first administered its state assessment, the Pennsylvania System of School Assessment, in 1996. The results starkly highlighted the low performance of students enrolled in the SDP relative to other students across the state (Travers, 2003). The year-over-year findings leading into the early 2000s drove political will to reimagine the educational offerings of the district and coincided with support for eliminating bureaucratic excesses and addressing the SDP's fiscal problems. Thus in 2001, the state announced as a major reform effort its intention to eliminate the SDP's local school board and replace it with a committee of individuals appointed by the governor and the Philadelphia mayor, to be known as the School Reform Commission (SRC).

The SRC's policies in the years that have followed have proven dramatic and controversial. The SRC has encouraged the expansion of a parallel charter sector and placed management of a number of SDP schools into the hands of private educational management organizations. It has also voted to close a number of district schools and lay off teachers and staff—even canceling the local teachers' union contract—in light of declining enrollment and continued budget shortfalls. These and other decisions have proven unpopular with parts of the local community, who view the SRC as disconnected

from the community and indifferent to the value of local input and social ties (see, for example, Jones, 2016; Mezzacappa, 2015b). The controversies and vitriol surrounding the SRC and on-going reform efforts in Philadelphia place great pressure on the district superintendent—who operates as the CEO of the commission—to improve the state of the school system and remake the district according to his or her own vision (Nowak, 2015). This burden also contributes to quick turnover: the SDP has seen four superintendents since the establishment of the SRC, each with his or her own unique brand of initiatives. As a result, district policies have often differed from one superintendent to the next. It is within this context of constant evolution and revision of district policy that zero tolerance came to prominence in the SDP, and within this context that it fell out of favor.

Zero Tolerance in the SDP

The School District of Philadelphia first enacted its zero tolerance policy in its 2002-2003 student code of conduct under new superintendent Paul Vallas. Vallas came to Philadelphia after serving as superintendent of the Chicago Public Schools, where he established a legacy of strong accountability paired with investment in infrastructure and support for expanded school programming, including charter schools, magnet programs, after-school programs, and summer school (Russo, 2003). Vallas' brand of accountability, while most publicly focused on academics, expanded into student discipline as well: he emphasized strict reporting of serious offenses and removing offending students from the classroom through expulsions or placement in alternative programs (Johnson, 1999a). Upon his arrival in Philadelphia, the discipline policy of the

SDP soon reflected these priorities: Philadelphia's 2002-2003 zero tolerance policy, written into the student Code of Conduct published annually, emphasized student placement into alternative programs as the appropriate response to serious violations such as harassment or assault, and mandatory reporting of offenses to the district and, when necessary, the police department. The Code of Conduct summarizing the policy outlined two levels of violations. Level 1 incidents—including disruption of school, reckless endangerment, and possession of drugs or alcohol, among other infractions—obliged corrective actions ranging from meetings with teacher and administrators to suspension. Level II incidents—including repeated Level 1 violations, assault, and weapons possession, among other violations—obliged placement of the student in an alternative program or expulsion (School District of Philadelphia, 2002). Relevant sections of the 2002-2003 Code of Conduct are provided in Appendix A.

The policy was strictly implemented. Between the 2001-2002 and the 2008-2009 school years, enrollment in alternative disciplinary programs in the district grew by 436 percent; the student compositions of these schools tended to have higher proportions of black male students, socioeconomically disadvantaged students, and students scoring lower on state assessments, as well as lower proportions of students reaching graduation (Chiang & Gill, 2010). With the policy's implementation came both praise and criticism. Supporters commended the policy for eliminating ambiguities, removing disruptive and dangerous students from the classroom, and shining a light on incidents that previously may have gone unreported. Critics questioned the sweeping nature of the policy and whether it denied due rights to students most in need of academic and behavioral supports

(Snyder, 2002). These questions would not be answered during Paul Vallas' tenure; the superintendent left his position in 2007 to take charge of the Recovery School District in New Orleans. He was succeeded in Philadelphia by Arlene Ackerman, a seasoned administrator most recently leading the school districts of Washington, D.C. and San Francisco.

Ackerman's time as superintendent would primarily come to be defined by her Imagine 2014 plan, a wide-ranging five-year plan announced in 2009 to remake the district. The proposal touched on many areas of district policy, including reducing class sizes, expanding elective and summer school programs, and turning over the district's lowest-performing schools to new management (Graham, 2009; School District of Philadelphia, 2009; Socolar, 2009). Prior to the release of this plan, however, Ackerman revised disciplinary policies in the district according to her new vision to limit the use of alternative disciplinary schools, which had been a cornerstone of Vallas' discipline policy. Ackerman shifted the SDP's discipline policy away from transfers to alternative schools and toward expelling and suspending more students in response to serious transgressions. While expulsion had been a potential repercussion of Level II offenses under the previous Code of Conduct, no students were expelled between the 2004-2005 and 2007-2008 school years under Vallas' leadership; rather, students committing serious infractions were either transferred to alternative programs or were permitted to remain in their traditional schools (Graham, 2008). Ackerman's new 2008-2009 Code of Conduct instead outlined procedures for expulsion hearings to be held by district leadership and

raised the cap on the maximum number of suspension days for a single infraction from five to ten (School District of Philadelphia, 2008).

The new policies were intended to take a hard line in addressing persistent rates of violent student behavior in schools, and the subsequent school year saw increases in the district's rate of out-of-school suspensions and expulsions (Hardy, 2014). As described in 2011 by a collaboration of Philadelphia youth advocacy organizations,

“There may be no other large, urban school system that matches the District in its promotion of zero tolerance and in the heavy use of out-of-school suspensions, expulsions, disciplinary transfers to alternative schools, referrals to law enforcement, and school-based arrests.” (Youth United for Change et al., 2011)

Partially in response to this new burden of conducting large number of expulsion hearings, the SDP created a preliminary committee in 2011 to review cases prior to hearings and flag those which clearly do not merit expulsion (Snyder, 2011b). Still, rates of suspensions and expulsions remained high. While Ackerman had changed the operation of the district's disciplinary policy, she largely left the zero tolerance legacy of Vallas intact (Rouff, 2010).

Though more students were expelled and suspended in the immediate aftermath of the new discipline policy, reports of dangerous behaviors and misconduct in the school district continued. Some proved jarring and daunting. A seven-piece 2011 series on school crime published by the *Philadelphia Inquirer* reported that an average of 25 students, teachers, or other school staff per day were assaulted, robbed, or rendered victims of other violent crimes (Sullivan et al., 2011). Even then, evidence of underreporting of serious incidents persisted; teachers described an atmosphere in which they were encouraged not to report incidents in order to render statistics collected on their

schools more positive (Snyder, Sullivan, Graham, & Purcell, 2011). The SDP initiated a number of intervention programs throughout Ackerman's tenure to preempt misconduct and address underlying issues encouraging such behavior, but these were often sporadic, not fully implemented, or directed toward only a handful of the district's schools. For example, the district committed to instituting a positive behavioral support program in the 2009-2010 school year, but ultimately the program floundered due to administrative turnover, weak school-level implementation, and, ultimately, lack of funds (Public Citizens for Children and Youth, 2010). In another approach, Ackerman expanded in-school suspension programs in spring 2010, but funds for these programs were cut by nearly 40 percent in the 2011-2012 school year (Snyder, 2011a).

Not until new leadership entered the SDP would the district's discipline policy dramatically change. Ackerman left the SDP in August 2011, in light of mounting public pressure over the district's budgetary deficits, cheating allegations, and, in part, her handling of racial violence at a high school in South Philadelphia (Mezzacappa, 2011). She was replaced in June 2012 by William Hite, former schools chief of Prince George's County, MD. Just two months later, the SDP announced a large-scale revision of its zero tolerance policy as outlined in the Code of Conduct. The revision was based in part on a Blue Ribbon Commission on Safe Schools that Ackerman had convened in 2010 seeking district-wide solutions to student misconduct as well as specific solutions targeted at 46 particularly dangerous schools. The commission's report, published in January 2012, made a number of recommendations, including strengthening students' peer relationships, establishing positive support systems, and streamlining reporting practices

(Blue Ribbon Commission on Safe Schools, 2012). The 2012 revision of the district's Code of Conduct grew in part out of the commission's report, but extended even beyond its modest recommendations.

In August 2012, Hite announced the new Code of Conduct, giving administrators more authority to handle disciplinary cases and emphasizing that out-of-school suspensions, expulsions, and transfers be last resorts (School District of Philadelphia, 2012). The new policy did away with the Level I and Level II infractions outlined in prior Codes of Conduct. Instead, the policy outlined a number of disciplinary behaviors and five tiers of intervention permitted as administrative responses for each: (1) in-school intervention; (2) out-of-school suspension; (3) contract with intervention or lateral transfer; (4) disciplinary school assignment; and (5) disciplinary school assignment with expulsion referral. While the most severe consequences remained for dangerous behaviors such as weapons possession or assault, the most common behaviors—such as disruption and truancy—aligned only with the first tier of intervention, in-school supports, even when these behaviors were repeated. In-school interventions remained an option even for more serious behaviors such as fighting, harassment, or destruction of property. The revision effectively dismantled the district's zero tolerance legacy instituted under Vallas and established an early tone of misconduct prevention as opposed to punishment for Hite's tenure (Hardy, 2014; Mezzacappa, 2012; Pope, 2012).

The district's reversal on zero tolerance, while championed by some local policy organizations and based on district's own Blue Ribbon Commission on Safety, largely drew on descriptive evidence on high rates of suspensions and expulsions in the district,

anecdotal evidence on continued crime and misconduct in the district, and successes of alternative programs promoting positive behavioral supports. The decision was not based on any compelling research demonstrating that the zero tolerance discipline policy itself contributed to excessively high rates of punishments—a view which in itself assumes that many of the punishments recorded by the district were overreactions to student actions—or that crime and misconduct would have occurred at lower rates in its absence. Nor was it based on any such generalizable research from other locales, as such research is largely absent (see Chapter 2: Review of the Literature). This paper aims to address this void in the research literature on zero tolerance, leveraging the district’s stark reversal on the policy to examine student behavior and disciplinary responses in the absence of the zero tolerance policy.

School-based Law Enforcement

While the district’s move away from its zero tolerance policy was a sharp reversal, it came at the same time many other large districts were rewriting their discipline codes, including Chicago, New York City, Los Angeles, San Francisco, Oakland, Miami, and Denver (Anderson, 2015; Hood, 2012). Like these and other districts, however, Philadelphia has continued to maintain a legacy of widespread police presence in its schools. The SDP first established a district police department in 1993, under the supervision of the district and fully separate from the citywide Philadelphia Police Department (Vera Institute of Justice, 1999). School police in the SDP do not carry weapons and cannot arrest students, but may respond to incidents and contact city law enforcement, detaining students until city police arrive on the scene. These school

police officers are different from the School Resource Officers (SROs) that are common in many schools across the country. SROs first entered Pennsylvania schools in 1997, expanding to 26 districts by the 2003-2004 school year and 87 districts by the 2011-2012 school year (American Civil Liberties Union of Pennsylvania, 2015). As described in Chapter 2, SRO programs generally place a greater emphasis on community building, mentoring, education, and conflict resolution than on crime containment and consequences for misbehavior. The SDP's school police officers, however, unambiguously operate to reduce crime and misconduct. According to their current SDP job description, school police officers' responsibilities primarily include "[patrolling] School District facilities and grounds to prevent disruptive or illegal actions, access to restricted areas, theft or vandalism on an assigned or rotating shift" (School District of Philadelphia, 2016). The job description makes no mention of mentoring or educational engagement with students.

While the SDP shifted its disciplinary policy to focus on in-school interventions and deemphasize severe punishments, its reliance on school police officers has largely remained stable. The number of school police officers in the district's 200+ schools remained just above 400 from the mid-2000s through the 2013-2014 school year before dropping slightly to 386 in 2014-2015 (School District of Philadelphia Consolidated Budgets, 2011-2016). The lack of a more significant drop in that timeframe is particularly notable given that SDP has faced years of declining enrollment and repeated cuts in many areas of the district budget over the last decade, exemplified by the closing of 30 schools in the 2011-2012 and 2012-2013 school years and the conversion of many traditional

district schools to charter schools. Thus despite a recent disciplinary atmosphere promoting positive behavior supports and interventions, the actions of the SDP reveal that the district continues to view school police as an established and necessary element in its disciplinary approach.

The fiscal problems of the SDP continue to plague the district, however, and in recent years budget shortfalls have shaped the extent to which the district itself can maintain its outlined policies. In the 2013-2014 school year, the district had no choice but to eliminate some of its intended police officer positions. In all, 28 schools—23 elementary, 2 middle, and 3 high schools—were required to share a police officer with another school: each school would host the police officer for half of each week (McCorry, 2014). That the district did not simply eliminate police presence in any of its schools indicates its continued belief in the value of officers on school grounds. Still, by examining these changes in school assignment of police officers, this paper informs the research base on the effects of reducing the presence of school-based law enforcement outside of the context of zero tolerance.

• • •

As described in Chapter 2, discourse and research on zero tolerance discipline policies often link such policies to increased police presence on school campuses and ultimately to severe punishments for student misconduct and high rates of punishment overall. The descriptive evidence here indicates that while both areas of policy may have

grown out of the same theories of behavior and the conclusion that taking a hard line on crime and misbehavior can best protect students, they may not be as intimately intertwined as is often suggested. The next chapter describes how the SDP's elimination of its zero tolerance policy but maintenance of its school police force can be used to examine the general impact of both zero tolerance policies and police presence in schools, without conflating the two and while retaining the ability to make causal inferences.

CHAPTER 4: METHODOLOGY

Zero tolerance discipline policy and law enforcement in schools are two areas of education policy which are much discussed but little researched in a rigorous manner. This paper addresses this void in the literature, examining the interplay of student behaviors and administrative responses and the extent to which they are affected as district policies change. The chapter begins by defining the primary research questions addressed in this paper, their boundaries, and the assumptions underlying each. The data used in the paper are then described, including their origin and their composition. Particular attention is given to the reliability and validity of the school-level data drawn from publicly available sources. The chapter then describes the statistical methods used to examine the data and inform the research questions.

Primary Research Questions

The research questions that shape this dissertation are straightforward in subject matter but complex in the quasi-experimental methods and assumptions they invite. The two primary lines of inquiry examined in this paper are:

- (1) To what extent did the dismantling of the SDP's zero tolerance discipline policy affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?; and
- (2) To what extent did the reduction of police officer presence in particular schools in the SDP affect severe student misconduct and consequential

administrative reactions? To what extent did it affect schools' academic performance and truancy rates?

Severe student misconduct as defined in this paper represents those student actions which harm other individuals or place them in danger; those which entail possession of controlled substances or weapons; or those which could potentially be considered criminal acts. For example, acts of violence would all fall into this category, while classroom disruption would not. Schools are required to report all such on-campus incidents per year. School-level outcome variables representing severe student misconduct include the following variables per school per year: total number of incidents and total number of incidents involving law enforcement. For the purposes of the analyses in this paper, these variables are scaled to represent rates of occurrences per 1,000 students. Full definitions of these outcome variables as per the Pennsylvania Department of Education are provided in Appendix E.

Administrative reactions as defined in this paper are the consequences outlined by the district as potential responses to such behaviors. School-level outcome variables representing administrative responses include the following variables per school per year: number of out-of-school suspensions, disciplinary transfers, expulsions, and arrests. These variables are also scaled to represent rates of occurrences per 1,000 students for the analyses in this paper. Note that in-school consequences, such as in-school suspensions, are not included here, as these are exclusively used within the SDP to address minor student misbehaviors; further, these have not been shown to be associated with the negative effects on students that are associated with more severe administrative

consequences examined here. Outcome variables representing schools' academic performance include the proportion of students scoring proficient or above on mathematics and reading state standardized tests per school per year, reported as percentage points; proficiency rate is used in place of raw scale scores given data availability and the more concrete interpretation provided by proficiency rates. The outcome variable representing school truancy rates is the school-level proportion of students absent on a daily basis per school per year, reported as percentage points. Again, full definitions of these outcome variables are provided in Appendix E.

These four outcome categories—severe student misconduct, administrative reactions, student academic performance, and student truancy—are each examined as potential areas affected by disciplinary policies as identified by the literature. While student dropout is also a potential area affected by disciplinary policies, no reliable data are available across all schools in the dataset for the years examined in this paper, and this outcome is thus not included in the study. Further, this dissertation examines only aggregate effects of the treatment across all school types; potential moderating effects of the treatment—e.g., treatment effects that differ by school size, racial/ethnic composition, grade range, or measures not included in the dataset, such as principal or teacher turnover—are not included, for reasons described in the Analysis section of this chapter.

The research questions examine both severe student misconduct and administrative reactions given that each captures a different function of disciplinary policy. Discipline policies are intended to inform students of unacceptable and potentially criminal behaviors and their consequences in order to deter such behaviors; they are also

intended to provide clear guidance to administrators as to how to respond to misconduct. While related, these two behavioral aspects are distinct. For example, if student misconduct increases but administrative consequences remain stable, administrators may be under-responding to infractions; on the other hand, if student misconduct levels do not change but administrative consequences increase in severity, administrators may have increased the rate at which they address minor actions with severe punishments. Only by examining these two aspects in tandem can research begin to tease out the comprehensive effects of policy changes on each (particularly given that no data on minor student misbehaviors are available). For example, data indicating that out-of-school suspensions increased after a policy change fall short of revealing whether the increase is due to more student misconduct or expanded use of out-of-school suspensions on the part of administrators. By comparing the increase in out-of-school suspensions with data on incidents of student misconduct, however, the data can begin to shed light on whether the increase originated from student behavior or administrative response.

This paper focuses on severe student misconduct given that zero tolerance policies and the use of school-based law enforcement are predominantly concerned with protecting student safety, as opposed to reducing less threatening infractions such as profanity or dress code violations. Further, discourse surrounding zero tolerance policies often refers to the ‘criminalization’ of students, i.e., the purportedly routine introduction of students to the criminal justice system and the permanent labeling of these students as criminals and delinquents (see Chapter 2: Review of the Literature). Focusing on those serious acts of misconduct that could potentially be deemed criminal behavior—and the

extent to which administrators address those behaviors as criminal acts—is thus more germane to addressing critiques of zero tolerance than focusing on less serious acts of misconduct. In the SDP, these severe student misconduct behaviors are those that, both before and after the dismantling of the zero tolerance policy, have been aligned with consequences of out-of-school suspensions, transfers, and expulsions; these are the administrative consequences that are also examined in this paper.

Boundaries of the Research Questions

In order to address the research questions, this paper examines the ways in which student conduct and academic response change in the absence of each of the policy decisions, i.e., after the dismantling of the district's zero tolerance policy, and after the reduction of police officer presence in district schools. The paper cannot address questions of how the introduction of the zero tolerance discipline policy in the early 2000s affected student conduct and administrative response, nor can it address questions of how the introduction of school-based law enforcement first affected student conduct and administrative response across SDP schools. To do so would require analysis of data in the years preceding policy implementation, *or* the assumption that student and administrative behaviors in the absence of a policy inevitably return to the levels of those behaviors that existed prior to the enactment of the policy. The latter is a strong assumption that can only be proven or disproven with examination of data from all time periods before and after the policies. If reliable data existed to track behaviors in the SDP prior to the origin of both of these policies and beyond their removal, then a removed-treatment times-series design, or a time series design with multiple replications, could be

leveraged to examine the overall impact of the policies and the extent to which they establish new baselines levels of behavior in their aftermath (Shadish, Cook, & Campbell, 2002). The data available for this study, however, are not reliable in the years prior to the establishment of the SDP's zero tolerance policy and do not provide information on early assignment of school-based law enforcement; thus the design must focus solely on the decisions to relax the zero tolerance policy and eliminate specific school police positions. The reliability of the available data is examined in more depth in the next section.

A second element not addressed by this paper is the extent to which school policy affects student behavior off of school grounds. As previously described, overall juvenile arrest rates declined alongside measures of in-school misconduct throughout the 2000s (Federal Bureau of Investigation, 2009), suggesting a link between in-school and out-of-school youth behavior. Investigating this link could provide useful insight as to the supposed criminalization of students and whether norms established within schools affect students in their external, daily life. The data necessary to examine such a link, however, are not included in this study. Drawing such a link would require arrest data linked to individual students to be matched with individual student records, while this study instead uses school-level records to address its research questions. While arrest data by census block, if it could be obtained, might shed some light on patterns of crime in the neighborhood surrounding each school, such records would provide ambiguous conclusions at best, as they would not provide the opportunity to determine whether or not the local offenders attended the nearest school or any district school at all.

Particularly in a city such as Philadelphia, with a large charter school sector, many students attend schools beyond their nearest neighborhood schools or schools outside of the district. Still, such an investigation would represent an additional line of inquiry beyond the research questions posed in this study; much can still be learned in the absence of such an analysis.

A final element that this paper cannot address is the extent to which surveillance bias affects reporting of student misconduct as opposed to actual student misconduct. Surveillance bias is the notion that monitoring for a specific behavior or trait will inherently lead to higher reported rates of that behavior or trait (Haut & Pronovost, 2011). It may be that removal of a zero tolerance policy leads to fewer reported incidents of misconduct and fewer administrative consequences, but no change in actual misconduct, as reported and actual behaviors cannot be distinguished with the data available. Even if this were the case, however, this paper argues that the knowledge that reported rates decreased in the absence of zero tolerance is useful; the finding would still demonstrate a real effect of zero tolerance, one with potentially important implications for student educational careers and school safety. Still, this paper makes the assumption that administrative reporting of student misconduct and subsequent consequences is not affected by the two examined policy changes, as reporting of misconduct is emphasized by the district both before and after the policy changes. This does not mean that misconduct is assumed to be perfectly reported, only that the patterns of reporting do not change across the treatment conditions. As described in Chapter 3, some school personnel describe an atmosphere in the SDP in which school leaders discourage teachers and other

staff from reporting student misconduct, potentially because these school leaders do not want to contribute to a negative perception of the school. In this case, the data examined in this paper would not perfectly record infractions and consequences, and may instead serve as an underestimate of actual students and administrative behaviors. Still, if the incentives to underreport go unchanged as the district moves away from zero tolerance and as individual schools remove police officer positions, then the rate of underreporting would remain the same prior to and after the policy changes. In that case, the effects of the policies would still be captured, though the estimated effects might serve as conservative estimates of the true effects, proportional to the rates of underreporting. Because the policy changes examined here had no impact on school reporting requirements or district-mandated school-intervention costs, it is reasonable to expect that the policy changes did not affect the level of underreporting across the district.

To address the primary research questions, this paper examines the following decisions and points in time:

RQ1: The SDP replaced its zero tolerance discipline policy, written into its annual Code of Conduct, in August, 2012. The zero tolerance policy had been in place since the 2002-2003 school year with some minor adjustments, most notably a shift from the use of disciplinary transfers as a common consequence for severe student misconduct toward the use of expulsions in the 2008-2009 school year. Given that this shift likely affected the rates of student misconduct and administrative response, only

data from the 2009-2010 school year through the 2011-2012 school year are used to establish a baseline of behaviors prior to the district's shift away from zero tolerance. The analysis also considers two years of behavioral data after the policy decision: the 2012-2013 and 2013-2014 school years. The 2014-2015 school year is not included, given that the SDP enacted a new district-wide program to divert students from the criminal justice system in this year; this policy would contaminate estimates of the effect of the prior policy. All traditional SDP schools are included in the analysis, as are various groups of comparison schools across the state.

RQ2: While the SDP dismantled its zero tolerance policy, it largely retained its cadre of school-based police officers. Still, due to budget concerns, the SDP opted not to fill some school police officer vacancies in the 2013-2014 school year, requiring 28 schools to share police officers part-time. Analyses pertaining to Research Question (2) thus use two years of data, the 2012-2013 school year and the 2013-2014 school year. This allows for one baseline year of data and one year of data after the staffing change.

One potential limitation of the data examined in this paper is the relatively few post-treatment years available for study, given the historical interaction problems introduced by including the 2014-2015 school year, as well as the fact that publicly

available data for this year were not yet published at the time of this analysis. If the policies lead to lagged effects—i.e., effects that do not occur until two or more years post-policy—this study will not capture them. Still, given that school districts are reconsidering revamping their disciplinary policies in real time, and given the dearth of literature on the subject, this study can provide some immediate useful information on the near-term effects of dismantling zero tolerance policies. Future research that extends this analysis to additional post-treatment years would be valuable for capturing the full, long-term effect of such policy decisions.

The span of years informing each of the research questions overlap: post-treatment years of Research Question (1) serve as the analysis years for Research Question (2). This overlap is not problematic. Research Question (1) addresses a policy change in Philadelphia, as compared with the rest of Pennsylvania. Research Question (2) only examines and compares schools within Philadelphia, all of which were affected by the policy change in Research Question (1), thus eliminating the potential internal validity bias of differential policy exposure as a result of the policy change in Research Question (1).

Data

The data analyzed in this paper are collected from publicly available datasets published by the National Center for Education Statistics (NCES) and by the Pennsylvania Department of Education (PDE). Both sources provide yearly student data aggregated at the school level for all schools in Pennsylvania. The datasets can be combined using Pennsylvania school identifiers, which are standard four-digit IDs

included in both the state and federal data collections. As publicly available data, the datasets contain no personally identifiable material and can be accessed by visiting various webpages within the federal and state websites. While school-level records are limited in the depth of information that they can provide as compared to student-level records, they are valuable in identifying school-level trends and examining schools as a unit. School-level outcomes are the measures by which most student performance and behavior is publicly evaluated, and thus analyses using school-level records can be informative to policymaking at the district, state, and federal levels. While school-level records cannot be used to investigate which types of students are most affected by various policy changes, examination of school-level trends lays an important and illuminating foundation for corresponding analyses of student-level trends beyond the scope of this study.

The data published by NCES, housed in the Institute of Education Sciences operated by the U.S. Department of Education, are drawn from the Common Core of Data collected by the department. The collection includes aggregated student and fiscal data on all public schools, school districts, and state education departments across the country. Data are reported annually to the department by the state education departments. The data selected for use in this paper consist of school-level student demographic information: race/ethnicity, gender, receipt of free or reduced-price lunch (a proxy for socioeconomic disadvantage), and enrollment totals, for all students in schools nationwide. The data also include an indicator for whether a school is a charter school or an elementary, middle, or high school, as well as a variable indicating the number of

pupils per teacher. Schools are identified by their Pennsylvania school ID as well as their name, county, school district, and type, i.e., traditional school versus special education school, vocational school, or alternative school. For the purposes of descriptive analyses, this paper draws on NCES data from 2001-2002 through 2013-2014; for inferential analyses, this paper only draws on data from 2009-2010 through 2013-2014. NCES expanded its race/ethnicity categorizations in 2007, rendering race/ethnicity data prior to and after the change incomparable; thus descriptive analyses employing race/ethnicity variables only draw on NCES data after 2007-2008. All other NCES variables were collected in a consistently defined manner across all years of data.

The data drawn from PDE come from two sources: the state's Safe Schools Reports and the state's collection of student scores on annual assessments. The Office of Safe Schools within PDE releases annual Safe Schools Reports: these report aggregate school-level data on the number of incidents involving one or more victims and one or more offenders; these could include acts of violence on school property, as well as incidents involving weapon possession or possession, use, or sale of controlled substances. The categories include assaults, intimidation/harassment, sexual offenses, threats, robbery/theft/burglary, manslaughter, arson/vandalism, weapon possession, and possession of a controlled substance. These align with those behaviors defined by the SDP as warranting consequences of out-of-school suspension, transfer, and expulsion. The data also report the number of out-of-school suspensions, transfers to alternative programs, and expulsions issued annually per school, as well as average school truancy rates. For the purposes of descriptive analyses, this paper draws on Safe Schools Reports

from 2001-2002 through 2013-2014; for inferential analyses, this paper draws on reports only from 2009-2010 through 2013-2014. All variables used in this study have been consistently defined since 2001-2002, though some—out-of-school suspensions, expulsions, and truancy rate—have only been collected since 2007-2008, and are thus only included in descriptive analyses for those years for which data are available.

Additional data drawn from PDE include school-level aggregate student performance on state assessments. Pennsylvania has assessed its students using the Pennsylvania System of School Assessment (PSSA) since 1992. Since the 2005-2006 school year, the PSSAs have assessed mathematics and reading in grades 3 through 8 and grade 11, in addition to less frequent writing and science assessments. In 2011, the state created new end-of-course assessments for high school subjects of Algebra I, Biology, and Literature. Known as the Keystone Exams, these assessments replaced the PSSA for high school students. PDE publishes state-wide school-level data on the percentage of students scoring proficient on each assessment each year. This paper employs data on the percentages of students scoring proficient in mathematics/Algebra 1 and reading/Literature in each school across Pennsylvania from 2009-2010 through 2013-2014. One concern in examining rates of proficient students is that states change their proficiency cut scores, and even their assessment systems, over time. For example, Pennsylvania most recently revised the PSSAs to align to new state standards and raised their proficiency cut scores for the 2014-2015 school year (Mezzacappa, 2015a). Because such changes affect all schools across Pennsylvania, however, and thus all schools in the

study sample, these changes do not affect the analyses in this paper, which compare academic performance of groups of Pennsylvania schools.

This study aims to generalize the effects of disciplinary policy changes on student and administrative behaviors to standard public schools enrolling students in any grade 3–12. Thus while the NCES and PDE datasets contain records for all types of schools operated in Pennsylvania, this study limits the sample of schools in the data to traditional public schools, eliminating alternative, special education, career/technical, and early childhood education schools (i.e., any school enrolling no students beyond grade 2) from the dataset. The NCES school-type variable was used to remove these non-traditional schools.

Appendix E provides a full codebook defining the raw data variables included in the combined dataset used in this study.

Reliability and Validity

A key issue relevant to the analysis of administrative records is the quality of the data and the extent to which conclusions drawn from the data are credible. While the condition of administrative data records continues to improve as state and federal data systems are developed and standardized, administrative data quality remains a key concern of administrators and statisticians across the globe (Trepanier, 2014). Only by first addressing data quality can any claims be made regarding the utility of the research-based evidence drawn from a given data collection. Most literature on the subject, however, provides key aspects of data quality to examine, void of any standardized methods to guide such an examination; see, for example, publications from the U.S.

Census Bureau and the European Commission (Bergdahl et al., 2007; Iwig, Berning, Marck, & Prell, 2013). Still, these aspects of quality are central to evaluation of a dataset. In this vein, Herzog, Scheuren, & Winkler (2007) outline five key properties of administrative data quality: accuracy, completeness, relevance, timeliness, and comparability. Accuracy and completeness are indicators of data reliability, or the extent to which the data are free of error. Relevance, timeliness, and comparability are indicators of data validity, or the extent to which the data represent the underlying constructs they are intended to capture (Rossi, Lipsey, & Freeman, 2004).

Accuracy. Given the data examined in this paper, two aspects of accuracy are particularly relevant. First is the reliability of the PDE collection of student scores on state assessments. These data are unique in this study because they attempt to capture the latent construct of student knowledge, as opposed to all other data in this study, which aim to capture observable demographics and behaviors. In the context of testing and measurement, reliability refers to the extent to which assessment items, and thus students' assessment scores, are free of measurement error (Viswanathan, 2005). The PSSA and Keystone exams administered by PDE are thoroughly vetted and historically high in reliability, with internal consistencies of $\alpha > .90$ for mathematics and reading assessments of each grade (Pennsylvania Department of Education, 2015b). In other words, these data can reasonably be expected to be accurate.

The second key aspect of accuracy relevant to the data in this study is the extent to which observable data are reported and collected in a standardized and error-free manner across all units in the sample—here, schools in Pennsylvania. Error here could

arise from vague definitions of required variables; different interpretations by those inputting data; accidental mistakes in input; and/or purposeful distortion resulting from incentives to distort (Boruch, 2015). Both NCES and PDE make efforts to minimize threats of the first two issues; both organizations publish materials explicitly defining their variables and hold training sessions for those who input and compile data.

Accidental and purposeful errors are more difficult to anticipate and eradicate, though both are targeted through the standardization capabilities of the digital software used to collect the data.

The Pennsylvania Information Management System (PIMS) is the state's digital student record management and reporting system established in 2006 and continually developed through grants from the federal Institute for Education Sciences (National Center for Education Statistics, 2016). PIMS is used to assign unique records to individual students and to collect data from local departments of education, for statewide publication and for submission fulfilling federal reporting requirements. The system encompasses both the Safe Schools reports and the data reported by Pennsylvania to NCES. PIMS is aided by its built-in Data Quality Engine, which validates data within and across datasets. The Data Quality Engine only allows certain standard values or values within certain ranges to be entered into the PIMS reports, reducing errors and missing data and ultimately maximizing consistency. Other data are flagged for review, such as overlapping enrollment dates for the same student in two schools or a change in a student's free or reduced-price lunch status (Lind, 2015).

NCES's national collection of school-level data uses similar data quality checks to ensure standardization and accuracy. Since the 2009-2010 school year, NCES has enacted an editing strategy in which certain values are flagged, either because the categorical value of a school's operating status has changed or because a given value raises the variance of the item over the previous four years, a calculation described in more detail below (Glander, 2015). If a value is flagged, NCES returns to the state department of education and attempts to resolve or confirm the issue. NCES also publishes notes on non-resolved issues for each state annually; none of the issues related to Pennsylvania's submissions between the 2009-2010 and 2013-2014 school years affect the dataset analyzed in this study, as they primarily involve pre-Kindergarten enrollment, ungraded student enrollment, and staffing totals. Further, NCES suppresses those values for which states did not provide information that resolved the flagged issue. Since the 2011-2012 school year, NCES has published the number of suppressed data points by variable for each state. Only 33 values—spanning a dataset with over 300 variables and over 3000 schools per year—were suppressed in the Pennsylvania data between 2011-2012 and 2013-2014.

Given the programmed quality checks and streamlined reporting processes that shape the PDE and NCES data, it is reasonable to assume that the data *collection* processes provide maximum accuracy. The question remains as to whether the data are *reported* to the collection databases with fidelity, that is, the extent to which the reported data accurately represent the properties they are intended to measure. Demographic information is first reported by parents, and though surely mistakes happen, these data

can be reasonably assumed to manifest, at worst, minor random error. Academic and behavioral data, however, may be tied to policy consequences—such as mandated interventions tied to school statistics—that would incentivize school personnel to manipulate the data. Questions surrounding both have arisen in Pennsylvania. For example, the state published a 2009 report that found irregular statistical results and test score erasure patterns on recent state tests at over 200 schools across the state and over 50 schools in Philadelphia (Herold, 2011; Herold & Mezzacappa, 2012). Investigations into these potential cheating incidents continue to the present day, with some educators facing criminal charges (Dean, 2016). Another example is the potential underreporting of misconduct on school campuses, discussed briefly in Chapter 3. The *Philadelphia Inquirer* reported in 2011 on teachers describing an atmosphere in which they were encouraged not to report behavioral incidents in order to render statistics collected on their schools more positive (Snyder et al., 2011). Misreporting of crime and arrest is common even among police departments (Maltz, 1999). These examples of potential misrepresentation could threaten the credibility of any analyses employing the affected data, though their effect is less relevant if it remains constant before and after the policy change, as discussed previously.

Short of a full-scale retrospective audit, there is no surefire way to determine post hoc whether data have been purposely misreported, particularly when done in a systematic way over multiple years. Even in the case of the SDP's cheating scandal, any list of schools identified as having cheated is likely incomplete and thus removing these schools from the analysis would not provide reassurance that the data are any more

accurate. Some evidence of variation in data points across years, however, may highlight particular one-off incidents of misreporting. To examine the extent to which such incidents might have occurred in the current dataset, this paper follows the data editing approach of NCES by analyzing variation in one year of data versus variation across multiple years (Goldberg, Stillwell, & Little, 2013). NCES calculates the two following variables:

$$Y_1 = \text{mean}(|Y_i - Y_j|) \text{ for all years } i \text{ and } j \neq \text{to current year}$$

$$Y_2 = \text{mean}(|Y_{\text{current year}} - Y_j|) \text{ for all years } j \neq \text{to current year}$$

To compute variable Y_1 , variation over the nearest adjacent 4 years of a data point are calculated as the average difference between each data point. (Note that NCES uses the *prior* 4 years of data, not the nearest adjacent 4 years of data, because NCES edits each new year of data as it is collected; here, both prior and subsequent years are used for comparison.) To compute Y_2 , the variation between the data point and each of the four nearest adjacent years is calculated, and the average of those four values computed.

NCES flags data points as possible errors if both of the following criteria are met:

(1) Y_2 is greater than or equal to a minimum set value

(2) Y_2 is greater than Y_1 by a minimum set value

In each case, the minimum set value is determined separately for each variable by the Census Bureau. A final step examines the ratio of the data point to another value in the dataset, such as enrollment totals, and sets another value of variance acceptability. To be identified as a potential error, the data point must exceed the minimum set value of all three criteria. In this paper, these steps were carried out for select variables in the NCES

Table 1. *Raw Data Error Rates by Variable and Year*

Year	# Schools	NCES Demographics			Safe Schools Reports		
		Enrollment	Female	Free and Reduced- Price Lunch	Incidents	Arrests	Out-of- School Suspensions
09-10	2987	0.4%	0.3%	0.9%	1.6%	1.1%	0.9%
10-11	2964	0.1%	0.1%	0.7%	1.6%	0.9%	1.1%
11-12	2915	0.1%	0.0%	0.6%	1.4%	0.5%	0.9%
12-13	2828	0.0%	0.2%	0.5%	0.9%	0.9%	0.9%
13-14	2954	0.2%	0.1%	0.8%	1.0%	0.7%	0.8%

and PDE datasets. The minimum set value in each case was the 90th percentile of the range of values; only those data points exceeding the 90th percentile according to all three criteria were flagged. The rates of flagged errors are presented in Table 1. Note that data on school-level assessment results are not included here, as the assessment dataset contained no relevant variables that could be used to calculate a separate ratio for comparison, and as the nature of assessments raises separate reliability issues described previously.

Using this methodology, no tested variable's estimated error rate exceeds 2 percent of observations of that variable for any year between the 2009-2010 and 2013-2014 school years, with most falling below 1 percent. The NCES demographic variables see lower error rates than the Safe Schools Report data, as would be expected given that NCES uses this same method to identify and correct errors, albeit with different minimum cutoff values. Further, not all flagged values are necessarily errors, as in some cases school demographics or student behavior could potentially see actual large shifts from one year to the next. This is particularly true of student behavior, which could

explain why more data points from the Safe Schools Reports are flagged. For example, the most extreme case flagged among the variable representing number of female students per school is a Pittsburgh school which saw an increase in female students from 161 to 202 to 203 across three years, in the same time that enrollment increased from 372 to 399 to 408. While the increase in female in students is approximately one third greater than the total enrollment growth, this case does not appear clearly erroneous. On the other hand, the most extreme case flagged among the variable representing number of behavioral incidents per school saw 211 incidents in 2009-2010 but 3 to 5 incidents in every subsequent year in the data, while enrollment only varied by +/- 30 out of approximately 280 students in that range. This could be an example of an error, or of a school that saw a major disciplinary issue in a single year. While the former is more plausible, the latter cannot be ruled out with the data at hand. News articles suggest that the school did in fact see a large number of incidents in the 2009-2010 school year that prompted the school to take new actions to curb disciplinary problems, suggesting that the expectation of a decline in incidents is reasonable to a degree (Polke, 2014; Todd, 2010). Overall, given the results shown Table 1, this paper maintains that it is reasonable to assume that the data collected for analyses are accurate within a minimal range of error.

Completeness. Completeness of the dataset is an indicator of the extent to which certain data are missing. In this regard, the NCES data are nearly complete. For the 2009-2010 school year through the 2013-2014 school year, NCES data includes 15,321 school observations in Pennsylvania, or approximately 3,000 per year. Of these, enrollment is

listed as “not applicable” for 50 observations, and enrollment values are missing for an additional seven observations; these are exclusively career and technical centers not captured by limiting the NCES school “type” variable to traditional schools. Once removed, the remaining dataset of 15,264 school observations contains almost no missing data at all. Across all variables in the dataset, only two—the number of students on free or reduced-price lunch, and the pupil/teacher ratio—see any missing data, with neither missing more than 1 percent of all values: only 0.56 percent missing and 0.09 percent missing, respectively.

In terms of data reported, the Safe Schools Reports are also void of missing data. No variables in the 15,295 variable dataset between 2009-2010 and 2013-2014 are missing. Many values, however, are listed as zeros; in fact, the vast majority of schools reports zero values each year for variables such as the number of arrests or disciplinary transfers. While this appears reasonable, some of the zeros could in theory represent missing data if the editing strategy of the reports converts missing data to zeros. Reassuring is the fact that no enrollment totals are equal to zero, meaning that every school did in fact provide data of some sort.

Lastly, the school-level assessment proficiency rates published by the PDE are also nearly complete. Of the 14,848 school-level observations between 2009-2010 and 2013-2014, just 24 values are missing each for mathematics and reading proficiency, accounting for just 0.16 percent of the data. Combined, the NCES, Safe School Reports, and PDE assessment data are robust datasets with almost no missing data, and no

indication from any of the publishing sources that missing data have been masked or ignored.

Relevance and timeliness. The primary research questions of this paper address the extent to which district policy decisions affected severe student misconduct and consequential administrative reactions. The data used to answer these questions are both relevant and timely, then, as they cover multiple years before and after the policy decisions were put in place, and they include measures of some of the most common examples of severe student misconduct and consequential administrative reactions. Ultimately, the purpose of research is to inform the assignment of value to a given policy decision to inform future decisions; here, the policies might prove valuable if they were to positively impact a student's educational career and life more generally, and valueless if the converse were true. While these notions are not well defined or measured, research reviewed in Chapter 2 indicated the many ties between disruptive behaviors, administrative actions, and negative life outcomes. Thus the variables examined here lay the groundwork for extension of this type of research into other student outcomes, at a time when many districts are considering similar policy decisions.

Comparability. When combining multiple datasets, comparability refers to the extent to which the datasets overlap and confirm each other, suggesting that they represent the same underlying constructs intended to be examined. In this case, each of three main datasets in this paper contain a column of four-digit school IDs assigned by the state of Pennsylvania. These are used to match the records from each dataset. In theory, each dataset should contain the same schools, with the exception that the state

assessment data only includes schools which enroll students in grade 3 through grade 8 or high school, whereas the NCES and Safe Schools Data publish data on all schools in the state.

For the purposes of this comparability analysis, data were merged from the 2009-2010 school year through the 2013-2014 school year, the years used in the inferential analyses used in this paper. First, the NCES and Safe Schools Data were matched. Of the 15,264 observations in the NCES data from the 2009-2010 through 2013-2014 school years, only 185 did not find a match; and of the 15,295 observations in the Safe Schools Data, only 216 did not find a match. These represent 1.4 percent and 1.2 percent of their original datasets, respectively. Upon close inspection, the vast majority of these unmatched cases are career technical centers or vocational/alternative programs. Overall, the combined dataset resulted in 15,079 matches. Given that the majority of the observations without matches fall outside the purview of this paper, these non-matched observations were dropped.

Second, the school-level proficiency results were matched with the merged NCES/Safe Schools Report data. Here, 520 of the 15,079 observations from the prior merged data did not find a match, while 289 of the 14,848 school-level proficiency observations did not find a match. These represent 3.4 percent and 1.9 percent of their datasets, respectively. Upon close inspection, the vast majority of unmatched observations from the NCES/Safe Schools Report data are early childhood education programs that do not enroll students as advanced as grade 3; while the vast majority of unmatched observations from the school proficiency dataset are vocational/alternative

programs or learning centers. Again, given that these observations fall outside the purview of this dissertation, the unmatched observations were dropped. The final remaining dataset contains 14,559 observations from the 2009-2010 through 2013-2014 school years.

The high rate of comparability of school IDs across each of these datasets reinforces the consistency and the legitimacy of each of the sources. Overall, while users of administrative data often face problems of accuracy, completeness, relevance, timeliness, and comparability, the data used in this paper surpass minimum expectations of quality in each category and can be reasonably presumed to be useful in addressing the research questions outlined in this paper.

Methods

The policies examined in this paper were not conceived prior to their implementation as experimental processes to be evaluated. That is, the population of affected schools and students, the timing of the policy intervals, and the related measurements collected were not initially manipulated to produce an experimental setting from which unbiased estimates of policy effects could be drawn with maximum statistical efficiency. Research methods in this context, then, refer to the post hoc combination of three aspects of the quasi-experimental research process: (1) the identification of a research design that best represents the policy changes and the data available, (2) the application of sampling methods to the data to support and enhance the design; and (3) the analysis of data using appropriate statistical methods given the assumptions and limitations of the design. The ultimate aim of these methods is to produce an artificial

setting closely approaching an experimental setting in which treatment is randomized to individuals or groups and perhaps even treatment intervals are randomized. The methods applied in this paper borrow from both the statistical and econometric literature and are layered upon each other here to provide evidence of potential effects from multiple viewpoints. For the purposes of the discussion below, “treatment” schools refer to schools affected by the SDP policy decisions examined in the research questions; “control” schools refer to schools not affected by the SDP policy decisions, sampled in various ways from the population of all traditional public schools in the state of Pennsylvania.

Design

Research design here is explored in the tradition of Campbell and Stanley (1966), who refer to the aspect of the research process in which measurements are manipulated and their relationships to other measurements explored, given a fixed setting in which the researcher has no control over the data available. Thus the research designs used in this paper are quasi-experimental in that they are designs applied to a natural social setting after policy enactment in order to approach the statistical properties of a fully experimental design. The aim of selecting an appropriate research design is to determine which threats to validity of causal estimates are eliminated by the design and which remain. Ideally, all research would imitate designs in which groups are randomized to approach equivalence in baseline characteristics; data are collected over an extended period of time both before and after the intervention of interest; and the intervention is randomly timed across multiple groups in addition to being staggered, reversed, and/or

repeated. A research design utilizing all of these elements is most likely to eliminate validity threats arising from history, maturation, attrition, regression, and selection, and their potential differential effects across control and treatment groups.¹ In the context of this dissertation, an ideal design for Research Question (1) would have been one in which Pennsylvania districts with zero tolerance policies were randomized into two groups, one maintaining the policies and one removing the policies; for Research Question (2), an ideal design would have been one in which Philadelphia schools with police officers were randomly sorted into two groups, one maintaining the current level of police officer presence and one limiting the current level of police officer presence. In both cases, an ideal design would also include data on multiple years of stable policies prior to and following the policy changes. In this study, however, while multiple years of data are available in some cases, the policy changes were not randomly assigned to districts or schools.

Given the panel data available for all schools across Pennsylvania, the policy changes addressed in Research Questions (1) and (2) can, however, be formulated as non-equivalent control group designs. In both cases, a treatment group experiencing the policy change can be paired with a non-equivalent control group that did not experience the policy change. The groups are non-equivalent because they could differ on any number of baseline characteristics, given that they were not randomly sampled from the same population of schools. Designs using non-equivalent groups are generally perceived as

¹ (Campbell & Stanley, 1966) also identify a number of other threats to validity—including testing, instrumentation, and their interactive effects—but these are less relevant to the setting and data explored in this paper and thus are not the focus of the above discussion on research design.

more robust than those using a one-group design—i.e., a design without a control group—as the inclusion of the comparison group eliminates bias resulting from the main effects of history, maturation, attrition, regression, and selection. In a one-group design, any one of these elements could affect the group's outcomes and account for any change or lack of change in the outcome variable. On the other hand, the main effect of any of these threats to bias would affect both groups in the same way in a non-equivalent control group design, and thus the difference in their change in outcome remains an unbiased estimate of the intervention effect.

Non-equivalent control group designs, however, must make the necessary assumption that the *interaction* effects of these threats are null. In other words, the assumption is made that the two groups do not mature at different rates, or that the mortality rate in one group is different than in the other. To the extent that the analysis does not account for or ignores these differences, these omitted variables can bias causal estimates.

Some weaknesses inherent to this large assumption of non-equivalent control group designs can be alleviated through the addition of design elements (i.e., multiple years of data, multiple groups) which render the design more robust; through sampling methods which identify control groups that approach equivalence with the treatment group; and through analysis methods which layer different approaches atop one another to form a multifaceted picture of causal effects. This paper takes this multi-pronged approach, using multiple data points, multiple comparison groups, and multiple analyses. Interpreted in tandem, these layers can help build a body of evidence to support (or

Table 2. *Design Elements, Sampling, and Analytic Methods*

Research Question 1			
Additional Design Elements	Analytic Methods	Sampling	
		Purposive Sampling Control Group	Propensity Score Matched Control Group
None: Nonequivalent Control Pretest/ Posttest Design	DD	Phil. Charter Schools	Phil. Charter Schools
		Pittsburgh Schools	Pittsburgh Schools
		PA Schools	PA Schools
Multiple Years of Data: Nonequivalent Control Group Time Series	GDD	Phil. Charter Schools	Phil. Charter Schools
		Pittsburgh Schools	Pittsburgh Schools
		PA Schools	PA Schools
	CITS	Phil. Charter Schools	Phil. Charter Schools
		Pittsburgh Schools	Pittsburgh Schools
		PA Schools	PA Schools
Research Question 2			
Additional Design Elements	Analytic Methods	Sampling	
		Purposive Sampling Control Group	Propensity Score Matched Control Group
None: Nonequivalent Control Pretest/ Posttest Design	DD	Other SDP Schools	Other SDP Schools
	FRD	Other SDP Schools	-

Note: DD = difference-in-differences; GDD = generalized difference-in-differences; CITS = comparative interrupted time series; FRD = fuzzy regression discontinuity; SDP = School District of Philadelphia.

dispel) the conclusions drawn from a non-equivalent control design. Specifically, design elements of multiple years and groups are added to Research Question (1), while sampling and the use of multiple analyses are applied to both research questions. The design elements added to Research Question (1) are described below, followed by the sampling and analyses applied to both research questions. The full analytic approach for both research questions is provided in Table 2.

This paper begins by framing both research questions in the most basic nonequivalent control group design: a simple nonequivalent pretest/posttest comparison group setting in which one data point is drawn immediately prior to the intervention and one is drawn immediately after. Differential rates of change between pretest and posttest for the two groups are used to calculate treatment effects.

Design elements of multiple years of data and multiple groups are then added to the Research Question (1) analysis. As previously discussed, the basic nonequivalent control group design is threatened by the nonequivalence of the two groups: selection bias could interact with threats such as history, maturation, attrition, and regression, leading to biased causal estimates (Shadish et al., 2002). To a limited extent, selection bias can be explored and controlled for using pretest measurements. As such, this paper extends the pretest/posttest design of Research Question (1) to a time series setting in which multiple data points are drawn both before and after the intervention. Interrupted time series provide more information on baseline values and posttreatment values, permitting estimation of changes in mean values over time as well as changes in trends (i.e., rates of change) over time. Further, in a non-equivalent comparison group time series, more information is available to explore the selection-interaction threats of history, maturation, attrition, and regression, and—potentially—to rule out each as a viable threat.

Further, for Research Question (1), multiple group comparisons are included, namely, charter schools in Philadelphia, schools operated by the Pittsburgh Public Schools, and all other schools in Pennsylvania. These groups are purposively chosen as

control groups based on their characteristics: charter schools in Philadelphia are likely the most similar group of schools to SDP schools, as they operate in the same neighborhoods and serve students who would otherwise attend the SDP; Pittsburgh's school district is the nearest district in size to the SDP, and also operates in a low-income urban environment; and all other schools in Pennsylvania provide a robust comparison group given the size and diversity of the schools and districts in Pennsylvania.

As explored in the next section, purposive sampling is limited in the extent to which it can approach equivalence across treatment and control groups. The use of multiple control groups, however, can help eliminate the threats to validity inherent in a single comparison and potentially build a case for the generalizability of treatment effects. Note that, with the use of other Pennsylvania schools as control subjects, one concern is the extent to which other districts in Pennsylvania may have experienced similar policy changes in the same time frame as the SDP, which would result in treatment receipt in both the treatment and control groups. Evidence suggests, however, districts across Pennsylvania did not as a whole dismantle their zero tolerance policies at the same moment as the SDP. Pittsburgh Public Schools, for example, is known for slightly tweaking its Code of Conduct on an annual basis; as recently as June 2015, the district published a new Code of Conduct adding the requirement of a hearing for any student recommended for a long-term suspension, just one element in its gradual move away from prior punitive disciplinary policies (Chute, 2015). While a full review of the disciplinary policies across Pennsylvania's 500 districts—and even more charter school organizations—is beyond the scope of this dissertation, it is clear that these districts vary

in their use of punitive responses to behavior. Despite the SDP's standing as the largest urban district in Pennsylvania, 20 districts reported higher rates of out-of-school suspensions per 1,000 students than the SDP in 2011-2012 school year, while the SDP ranked 30th in its rate of arrest per 1,000 students (American Civil Liberties Union of Pennsylvania, 2015). While it is likely that these and many other districts did have zero tolerance policies, it is unlikely that all simultaneously removed these policies at the outset of the 2012-2013 school year. Still, if a handful of districts were to do so at the same time, the changes would only serve to underestimate any treatment effects of the SDP policy change in comparison.

Note that, to address Research Question (2), non-affected SDP schools serve as the sole control group in order to maintain comparability with analyses using fuzzy regression discontinuity methods, discussed later in this chapter. Further, no additional baseline or posttreatment years of analysis are added, also maintaining comparability with the fuzzy regression discontinuity analyses.

Sample

The predominant weakness of a non-equivalent control group design based on purposive sampling is by definition the non-equivalence of groups at baseline. Given this lack of established comparability, estimates of treatment effects are likely biased, as analyses applied to nonequivalent control group designs make the assumption that the counterfactual baseline and posttreatment values of the outcome for treated subjects are those for the control subjects. If the groups are not statistically equivalent at baseline, this assumption is false. Pretest values can give some indication of differences across the

treatment and control groups, and covariate measures can help control for some of these differences. Still, all relevant differential aspects would need to be included in the model to eliminate bias completely, and these aspects are nearly always unknown and/or unmeasured.

This dissertation pairs clearly defined treatment groups with nonequivalent control groups sampled purposively by convenience. For Research Question (1), these include all other charter schools in Philadelphia, all public schools in Pittsburgh Public Schools, and all other public schools across Pennsylvania. For Research Question (2), the control group includes all other SDP schools. Examining multiple contrasts of the treatment group with these control group samples can give some broad evidence of potential treatment effects, but each comparison is still subject to potential bias.

In addition to analyses using this purposive convenience sampling, this paper conducts parallel analyses using propensity score matching (PSM) methods. PSM is a method for amassing a control group sample that can better approximate group equivalence and reduce bias. PSM methods calculate a conditional probability of assignment to treatment for all potential control group members, based on observed characteristics of the treatment group and observed characteristics of each potential control group member (Rosenbaum & Rubin, 1983). Treatment group members are then matched to control group members of similar propensity score. This matching approach can eliminate bias when all characteristics related to both treatment assignment and outcome are included in the calculation of the conditional probability; or, in other words, when the treatment assignment and the outcome are conditionally independent given the

observed characteristics. The richer the set of covariates used to calculate propensity scores, and the more likely those covariates are to have influenced the treatment selection process, the closer an analysis can come to meeting this assumption of strong ignorability; on the other hand, covariates of convenience, if not related to the selection process, can retain or even increase bias (Shadish, 2012; Shadish, Clark, & Steiner, 2008). The covariates used to calculate the propensity scores in this paper include observations of school-level enrollment, level (primary school, middle school, high school, or other), race/ethnicity, proportion of students receiving free or reduced-price lunch, proportion of students who are female, and pupil-teacher ratio. While these measures are collected in administrative records and not for the purposes of this study, they are likely to be related to the “selection” process of SDP policy decisions, in that schools selected for the treatment group are those operated by the SDP; these schools differ from other schools across the state—and the SDP is often defined—by their demographics and resources, both of which are captured by these covariates.

Many approaches to calculating propensity scores have been proposed. This paper relies on those that have shown significant bias reduction as compared with unadjusted ordinary least squares approaches. Propensity scores here are calculated using logistic regression and transformed onto the logit scale to maintain their linearity (Rubin, 2001). Control groups are formed by matching each treatment school to a control school using nearest neighbor matching based on the Mahalanobis distance between scores within a caliper of ± 0.2 standard deviations of the propensity score logits; though many alternatives have been suggested, studies have not demonstrably found other methods

with clear advantages over this method proposed by Rosenbaum and Rubin (1985). Allowing one-to-many matching with replacement within a caliper range increases the size of the control group sample over that of simple one-to-one matching. Generally, larger analysis samples produce less biased results: Shadish (2012) finds that a PSM analysis with sample size of 200 reduces bias 83-85 percent of the time, while a sample size of 500 reduces bias 97-99 percent of the time. Intact group matching and multi-level matching are also common procedures in PSM analyses (see, for example, Aiken, West, Schwalm, Carroll, & Hsiung, 1998; Diaz & Handa, 2006; or Michalopoulos, Bloom, & Hill, 2004), but given that schools in this study are nested within districts and only one district (SDP) received treatment, intact group matching is not a viable option.

PSM can face hurdles when the distribution of treatment group propensity scores and that of the matched control group do not substantially overlap. Thus the viability of the propensity scores used here is evaluated according to three criteria aimed at maximizing balance, motivated by Rubin (2001):

- (1) The standardized difference in the treatment and control groups' mean propensity score should be near zero;
- (2) The ratio of the groups' propensity score variable should be near one; and
- (3) The ratio of the variance of each matched group's covariate should be between 0.80 and 1.25.

While PSM methods do not move the research design out of the nonequivalent control group framework entirely, they can reduce the bias due to nonequivalence when applied thoughtfully. Paired with analyses employing purposive convenience sampling,

the analyses using propensity score matching in this paper are considered more reliable and less biased, with the purposive sampling analyses included as a form of sensitivity check on the magnitude and direction of the propensity score matched analyses.

Analyses

The statistical analysis methods in this paper range from the simple to the complex. They build on the framework of the nonequivalent control group design, beginning first in the setting of a nonequivalent control pretest/posttest design and then extending to a nonequivalent control group time series design. All else equal, simpler methods are preferred; yet if those methods do not accurately capture relationships in the data, more complex methods are warranted. In this paper, then, both simple and complex methods are conducted, and their interpretation based on which methods capture statistically significant and important features of the treatment function. As all the empirical methods described below employ serially correlated panel data—correlated measures of the same school across time, and correlated school-level measures within the same districts—this paper calculates bootstrap standard errors to inform all inferential tests of significance drawn from the analyses (Bertrand, Duflo, & Mullainathan, 2004).

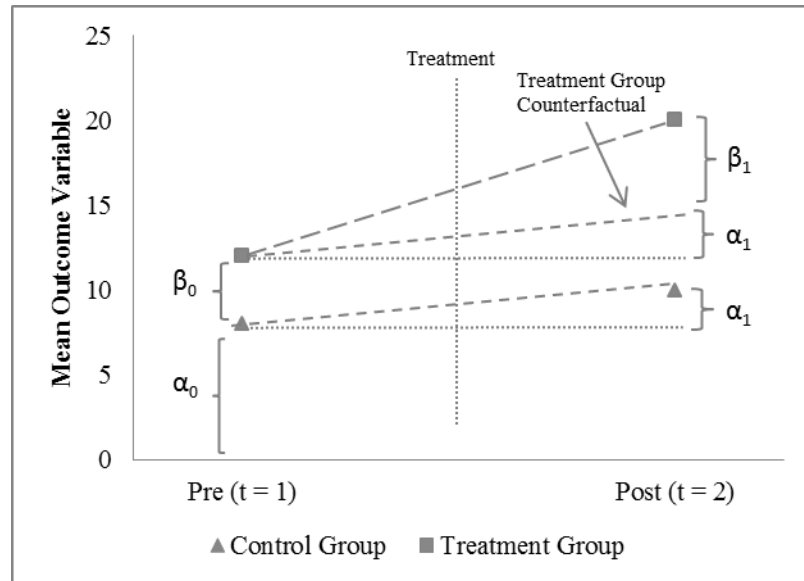
Difference-in-differences (DD). Difference-in-differences (DD) analysis compares change over time of some measure of a treatment group relative to control group. The design requires some setting comprised of at least two time periods of data—before and after an intervention—and at least two groups, one experiencing the intervention and a control group not experiencing the intervention. The estimand of the basic DD analysis is the difference between the two groups' change over time from their

baseline means to their follow-up measures (Angrist & Pischke, 2009). Basic DD analysis thus aligns with a nonequivalent control pretest/posttest design, as shown in Table 2. This paper considers the differences in the change in school-level outcomes between treatment schools and control schools before and after the SDP policy changes described in the research questions. Basic DD is commonly represented by equation E1:

$$E1. Y_{st} = \alpha_0 + \beta_0 Treat_Group_s + \alpha_1 Post_t + \beta_1 (Treat_Group_s \times Post_t) + \varepsilon_{st}$$

In the context of this paper, s denotes schools from $s = 1 \dots S$, and t denotes time, where $t = 0$ represents time prior to the treatment and $t = 1$ represents time after the treatment. Y_{st} is the outcome for school s in time period t . $Treat_Group_s$ is an indicator for whether school s was in the treatment group, set to 0 if no and 1 if yes. $Post_t$ is an indicator equal to 1 if the observation is from the post-treatment time point ($t = 1$) and 0 otherwise. The intercept α_0 represents the estimate of the baseline mean outcome for the control group, while $\alpha_0 + \alpha_1$ represents the mean outcome for that group after the policy change; $\alpha_0 + \beta_0$ represents the estimate of the baseline mean outcome for the treatment group, while $\alpha_0 + \beta_0 + \alpha_1 + \beta_1$ represents the mean outcome for that group after the policy change. The difference of the differences of these quantities, which can be expressed as $[(\alpha_0 + \beta_0 + \alpha_1 + \beta_1) - (\alpha_0 + \beta_0)] - [(\alpha_0 + \alpha_1) - \alpha_0] = \beta_1$, is the estimate of the average effect of the policy change on the outcome.

Equation E1 is shown graphically for a hypothetical dataset in Figure 1. The four data points represent the mean outcome variable for each group at each time point. The counterfactual trajectory of the treatment group is plotted based on the actual trajectory of



Notes. Figure presents hypothetical difference-in-differences (DD) analysis.

Figure 1. Graphical Portrayal of DD Analysis

the control group. The causal estimate, β_1 , can be seen as the change in the pre/post mean treatment values over and above the change in the pre/post control mean values.

The specific DD model employed in this analysis is represented by equation E2:

$$E2. Y_{st} = \alpha_0 + \beta_0 Treat_Group_s + \alpha_1 Post_t + \beta_1 (Treat_Group_s \times Post_t) \\ + x_{st} + v_s + \varepsilon_{st}$$

In this particular case, school-level covariates (x_{st}) are included to account for time-varying compositional changes, and dummy variables (v_s) are included to account for time-invariant fixed effects. These additional measures help account for error and reduce bias resulting from the omission of variables related to the given outcome. The school-level covariates include racial/ethnic composition, grade span, school size, proportion female, proportion of students

receiving free or reduced-price lunch, and pupil-teacher ratio. Relevant time-invariant characteristics represented by the dummy variables could include, for example, school location, static funding level, or on-going leadership. With the inclusion of the fixed effects, the coefficients on the covariates must be interpreted as the relationship of the outcome with schools *changing* across the covariate measures, not simply schools static in those measures, as the static characteristics are conjointly captured by the fixed effects dummies. In the same vein, the model intercept and treatment coefficient are no longer interpretable as the baseline means of the treatment and control groups, because these variables are now also confounded by the fixed effects dummies. As the covariates are not interpretable in the traditional sense, testing for moderating effects by interacting the covariates with the treatment does not lead to estimates of how treatment might be moderated by static school characteristics. Thus moderating effects are not examined in this study, nor are coefficients on covariates interpreted in discussion of the results (though all model results across all analyses are reported in Appendix A). Note that, with the inclusion of the time-varying covariates and the time-invariant fixed effects, the estimate of the effect of the policy change, β_1 , is unaffected but potentially less biased.

Different families of linear regressions, namely regressions based in the normal (Gaussian), Poisson, and negative binomial distributions, are applied to the data within the framework of the DD analysis in order to determine which type of linear regression best fits the data. These steps are described in detail in Chapters 6 and 7, prior to reporting and interpretation of model results.

Generalized difference-in-differences (GDD). As applied to any nonequivalent pretest/posttest design, DD analysis is limited in the information it can provide. DD analysis only examines the mean difference in the outcome variable across two points in time. It relies on a single baseline measure to represent baseline differences and cannot examine any delayed effects of the intervention. To expand on the capabilities of basic DD analysis, this dissertation shifts from a nonequivalent pretest/posttest framework to a nonequivalent control group time series framework, which utilizes multiple years of data to examine treatment effects. Within this expanded framework, this dissertation applies a generalized difference-in-differences (GDD) analysis to the data. GDD analysis can incorporate multiple time points and multiple groups into the DD framework. The multiple pre-/post-policy years of data available relevant to the policy change examined in Research Question (1) invite this more complex iteration of DD analysis. (See Table 2 for the full strategy of analyses applied to each research question.) GDD analysis is commonly represented by equation E3:

$$E3. Y_{sgt} = \gamma_g + \lambda_t + \delta D_{gt} + \varepsilon_{ist}$$

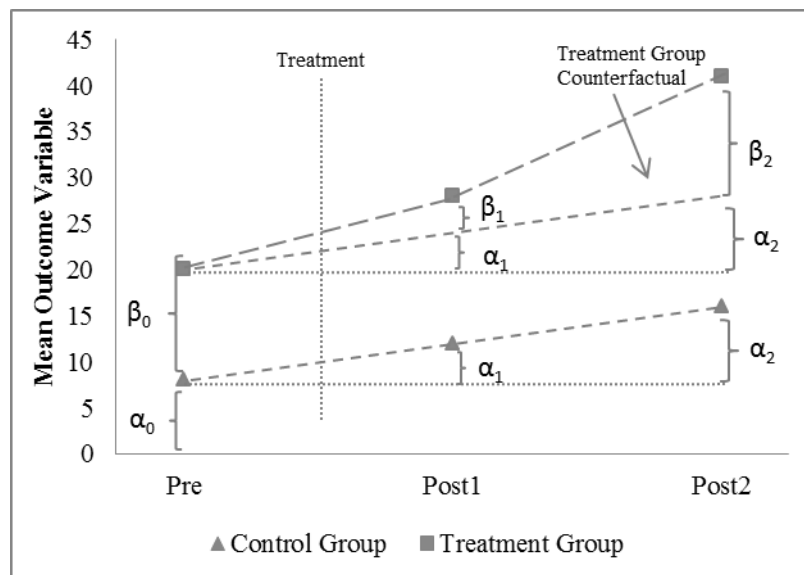
In this case, γ_g and λ_t are dummy variables representing group g (any number of treatment or control groups) and time t , respectively. D_{gt} is a dummy representing treatment schools in the time periods after treatment. The estimate of the effect of the policy on the outcome is δ . The model no longer provides the simple intuition of a difference between rates of change across two groups, but the underlying interpretation is the same as for equation E1. When limited to two time periods, E1 and E3 are computationally equivalent.

This paper applies GDD methods to include three school years of pretest data (2009-2010 through 2011-2012) and two school years of posttest data (2012-2013 and 2013-2014) to address Research Question (1). In generic form, this expanded model can be represented by equation E4:

$$\begin{aligned} \text{E4. } Y_{st} = & \alpha_0 + \beta_0 \text{Treat_Group}_s + \alpha_1 \text{Post1}_t + \beta_1 \text{Treat_Group}_s \times \text{Post1}_t \\ & + \alpha_2 \text{Post2}_t + \beta_2 \text{Treat_Group}_s \times \text{Post2}_t + \varepsilon_{st} \end{aligned}$$

Here, s again denotes schools and t denotes time. Y_{st} is the outcome for school s in time period t . Treat_Group_s is an indicator for whether school s was in the treatment group, set to 0 if no and 1 if yes. Post1_t is a dummy variable indicating the first year after the policy change, set to 1 if the school year is 2012-2013 and 0 otherwise. Post2_t is a dummy variable indicating the second year after the policy change, set to 1 if the school year is 2013-2014 and 0 otherwise.

Like the basic DD approach, this GDD model estimates the baseline mean for the control schools (α_0), now aggregated over three pretreatment years, and for the treatment schools ($\alpha_0 + \beta_0$). It adds the ability to model the changes from these baseline values one and two years after the policy change: α_1 is the change in baseline mean for the control group one year after the policy change; α_2 is the change in baseline mean for the control group two years after the policy change; $\alpha_1 + \beta_1$ is the change in baseline mean for the treatment group one year after the policy change; and $\alpha_2 + \beta_2$ is the change in baseline mean for the treatment group two years after the policy change. The estimated treatment



Notes. Figure presents hypothetical generalized difference-in-differences (GDD) analysis.

Figure 2. Graphical Portrayal of GDD Analysis

effect one year after the policy change is thus β_1 , while the estimated treatment effect two years after the policy change is β_2 .

The treatment effect across two posttreatment years are shown graphically for a hypothetical dataset in Figure 2. The six data points represent the mean outcome variable for each group at each time point. As in Figure 1, the counterfactual trajectory of the treatment group is plotted based on the actual trajectory of the control group. The causal treatment estimates for each of the two years post-treatment, β_1 and β_2 , can be seen as the change in the pre/post mean treatment values over and above the change in the pre/post control mean values at each time point. Were these same data modeled using basic DD analysis, results would fail to indicate a growing treatment effect over time, at least two years out from the treatment.

As with the DD analysis, the specific GDD model applied in this study includes time-varying school covariates and school-level fixed effects, resulting in equation E5:

$$E5. Y_{st} = \alpha_0 + \beta_0 Treat_Group_s + \alpha_1 Post1_t + \beta_1 Treat_Group_s \times Post1_t \\ + \alpha_2 Post2_t + \beta_2 Treat_Group_s \times Post2_t + x_{st} + v_s + \varepsilon_{st}$$

The fixed effect dummy variables (v_s) confound the estimate of the intercept and the coefficient on the treatment group variable, again rendering these estimates non-interpretable; the time-varying covariates (x_{st}) capture variance due to changing school compositions over time, not static school characteristics, and thus are not interpretable in the traditional sense. While serial correlation of multiple yearly measures within schools could theoretically drive the use of a random effects mixed model here, a fixed effects model is used given that the schools are not randomly distributed, nor are the schools in this analysis a random sample from a larger population.

GDD analysis has been used in many well-respected econometric papers. Autor (2003) examines whether policies limiting employers' right to fire employees at will leads to their hiring temporary workers at greater rates, and finds that employers begin hiring temporary workers at higher rates in the year before the enactment of such limiting policies in anticipation. Gruber and Kleiner (2012) analyze data on nurses' strikes across New York from 1984 to 2004 and find that strikes increase mortality and readmission rates of patients admitted during the strikes. Wing and Marier (2014) examine policies limiting certain dental procedures to performance by a dentist versus a dental hygienist and find that such policies, across multiple states and multiple intervals, increase the pricing of such procedures. GDD analyses can aid in identifying and controlling for

Ashenfelter's dip—a scenario in which the outcome variable declines (or increases) just prior to policy implementation in the group affected by the policy, first noted in Ashenfelter (1978)—as lags and leads to the intervention can be included. Ashenfelter's dip is just one example of a regression-selection interaction threat that can be alleviated through GDD. These can also serve as sensitivity checks in comparison to the estimated intervention effect, as the treatment timing can be artificially adjusted. These papers and methods demonstrate the flexibility of GDD approaches in their application to multiple time points, instances, and groups.

The GDD analyses in this paper, with their estimation of baseline measures across three pretreatment years and treatment effects across two posttreatment years, are more complex than the DD analyses. Were the treatment effect to prove stable in both posttreatment years, the GDD model would not be necessary, and the DD model would be preferred as it represents the simpler model of the two. The magnitude and statistical significance of the treatment effect estimate in the second year posttreatment suggests whether the GDD analysis is warranted. Where the GDD models return the same treatment effect for both posttreatment years, the DD estimates are preferred, and these drive interpretation of the treatment effects; where they do not return a stable treatment effect for both posttreatment years, the GDD treatment effects are interpreted.

GDD analyses are not without their limitations. Like DD analyses, GDD analyses must assume that threats to validity, such as maturation, do not interact with group selection. For example, one might expect the Philadelphia charter sector to experience differential fluctuations in its incident rate across the years in this analysis, as the sector

expanded rapidly between the 2009-2010 and the 2013-2014 school years, growing from 20 percent to 30 percent of Philadelphia's public school market share while traditional schools shrank from 80 percent to 70 percent of the market. Multiple years of data can be included in the GDD model to provide more reliable estimates of stable baseline and posttreatment means, but they do not provide information on such differential baseline or posttreatment trends. Potential differential rates of maturation are a source of bias that cannot be eliminated with GDD analyses (Somers, Zhu, Jacob, & Bloom, 2013). To gain more insight on changes in group trends over time, this paper turns to comparative interrupted time series analysis.

Comparative interrupted time series (CITS). Comparative interrupted time series (CITS) are an elaboration on GDD models, though they emerge from the tradition of basic interrupted time series. DD analyses focus on mean differences estimated across group membership, post-treatment time indicators, and their interaction; GDD analyses expand this approach to account for multiple groups and time points. On the other hand, interrupted time series analyses focus on mean and trend differences of one group across time estimated using variables of time index, post-treatment time indicators, and their interaction. The basic interrupted time series model is given in equation E6:

$$E6. Y_{st} = \beta_0 + \beta_1 Year_t + \beta_2 Post_t + \beta_3 (Year_t \times Post_t) + \varepsilon_{st}$$

Equation E6 is nearly identical to equation E1, except that the variable $Year_t$ has replaced the variable $Treat_Group_s$. The basic interrupted time series design includes only one uniform sample of treated observations, rendering $Treat_Group_s$ obsolete, given that it indicates treatment status across groups. In its place, $Year_t$ represents an

index of all times t in the dataset. β_0 and β_1 serve as estimates of the intercept and slope of the baseline trend, respectively, while β_2 and β_3 serve as estimates of the deviations from the baseline slope and trend in the posttreatment years, respectively.

The addition of group membership as a main and interaction effect shifts an interrupted time series into the CITS framework. The basic CITS model in the context of this study is given in equation E7:

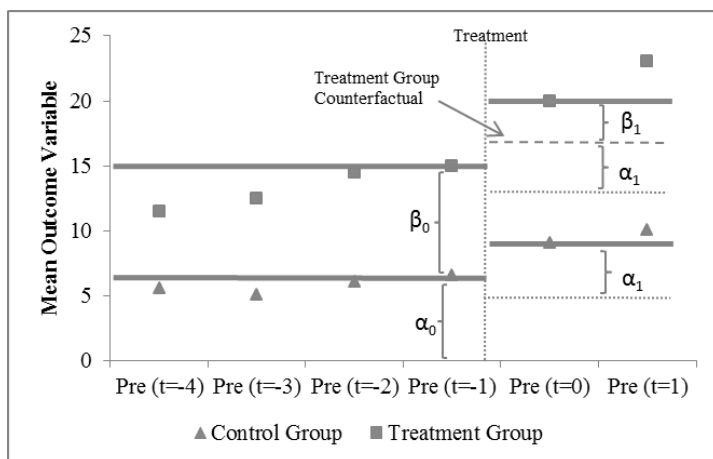
$$\begin{aligned} \text{E7: } Y_{st} = & \alpha_0 + \phi_0 \text{Year}_t + \beta_0 \text{Treat_Group}_s + \phi_1 \text{Year}_t \times \text{Treat_Group}_s \\ & + \alpha_1 \text{Post1}_t + \beta_1 \text{Treat_Group}_s \times \text{Post1}_t + \alpha_2 \text{Post2}_t \\ & + \beta_2 \text{Treat_Group}_s \times \text{Post2}_t + \varepsilon_{st} \end{aligned}$$

Equation E7 is identical to equation E4, with the addition of Year_t and its interaction with Treat_Group_s . Year_t is an index of observation year, centered at the last pretreatment year. Given this centered variable, α_0 now represents the baseline mean in the year prior to treatment for the control group, and $\alpha_0 + \beta_0$ represents the baseline mean in that year for the treatment group. Unlike equation E4, however, Equation E7 introduces the estimation of trend. Here, ϕ_0 and $\phi_0 + \phi_1$ represent the baseline slope for control and treatment schools, respectively. The estimates α_1 and α_2 now represent the deviation from baseline trend for control schools one and two years after the policy change, respectively; $\alpha_1 + \beta_1$ and $\alpha_2 + \beta_2$ represent the deviation from baseline trend for treatment schools one and two years after the policy change, respectively.

Because the method accounts for baseline trend differences, the CITS model helps alleviate concerns relative to differential maturation and regression—such as potential differential baseline trends of the Philadelphia traditional and charter schools, for

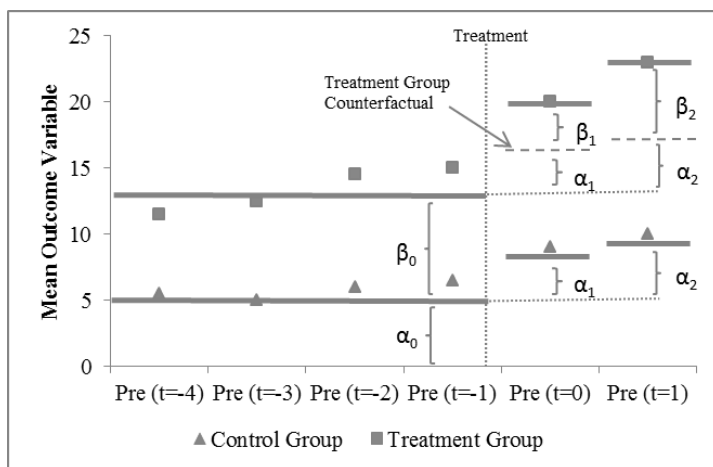
example—whereas a GDD model could ignore these important elements of a comparison. Accounting for baseline trend differences is especially important when extending impact analyses beyond the first year posttreatment, as the additional data points must be compared relative to counterfactual potential posttreatment trends. Given the additional information provided by CITS analysis, the approach may appear superior to that of a DD or GDD analysis; still, the additional complexity of the CITS analysis may not prove necessary, if treatment and control groups do not exhibit differential baseline trajectories. Where the CITS model produces statistically significant estimates of differential baseline trends across treatment and control groups—and thus the additional complexity of the CITS model is justified—the CITS model results will be interpreted in place of the GDD or DD results; where these effects are not statistically significant, the GDD or DD results will drive interpretation, as the more complex CITS model is not warranted.

Figure 3 presents a graphical representation of a CITS approach versus a DD or GDD approach as applied to the same hypothetical dataset. In the first panel representing DD analysis, differential treatment effects are estimated only taking into account one baseline year of data and one posttreatment year of data, leaving β_1 as the overall treatment effect. In the second panel representing GDD analysis, two posttreatment years are compared with baseline means for each group aggregated over four prior years of data. β_1 and β_2 represent the treatment effects one and two years posttreatment. These estimates provide more information of the treatment effect over time than the results of the first panel, but still fail to take into account differential baseline trends pre-treatment. In the third panel representing CITS analysis, four prior years of data are used to estimate



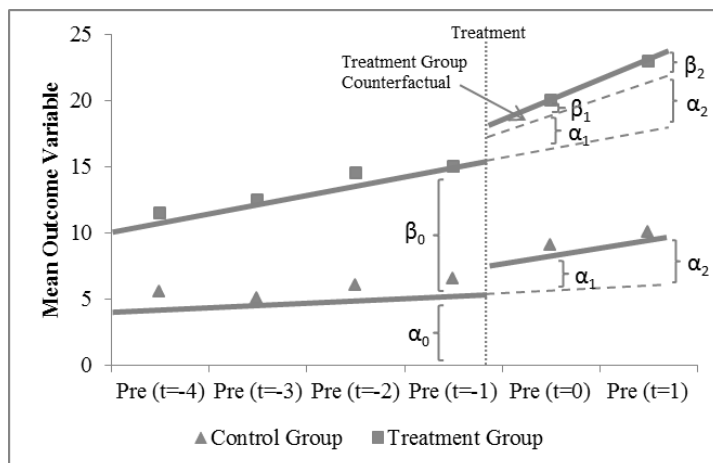
Difference-in-Differences (DD).

Only one baseline year of data and one posttreatment year of data inform estimation of the treatment effect, β_1 . Other years of data are not included in the analysis.



Generalized difference-in-differences (GDD).

Yearly dummies are added to estimate treatment effects by year. Here, two posttreatment years are compared to aggregate pretreatment years.



Comparative interrupted time series (CITS).

Baseline slopes are modeled and baseline trajectories inform estimation of deviation from baseline trends for treatment and control groups.

Notes. Figure presents difference-in-differences (DD), generalized difference-in-differences (GDD), and comparative interrupted time series (CITS) analysis as applied to an identical hypothetical dataset.

Figure 3. Graphical Display of DD, GDD, and CITS Analyses with Identical Dataset

baseline trend and two posttreatment years of data are used to estimate posttreatment trend. β_1 and β_2 again represent the treatment effects one and two years posttreatment, but here they are smaller than those estimated in the first and second panels because the CITS model accounts for differences in baseline trend across groups. Ignoring these baseline trends would have resulted in inflated estimates of the treatment effect.

Following in the vein of the DD and GDD analyses, the specific CITS model applied in this study includes time-varying school covariates and school-level fixed effects, resulting in equation E8:

$$\begin{aligned} \text{E8: } Y_{st} = & \alpha_0 + \phi_0 \text{Year}_t + \beta_0 \text{Treat_Group}_s + \phi_1 \text{Year}_t \times \text{Treat_Group}_s \\ & + \alpha_1 \text{Post1}_t + \beta_1 \text{Treat_Group}_s \times \text{Post1}_t + \alpha_2 \text{Post2}_t \\ & + \beta_2 \text{Treat_Group}_s \times \text{Post2}_t + x_{st} + v_s + \varepsilon_{st} \end{aligned}$$

Once more, these fixed effect dummy variables (v_s) confound the estimate of the intercept and the coefficient on the treatment group variable, rendering these estimates non-interpretable in this particular case; the time-varying covariates (x_{st}) capture variance due to changing school compositions over time, not static school characteristics, and thus are not interpretable in the traditional sense. All other variables are still interpretable as described in the more general case of equation E7.

The CITS approach has seen less widespread usage, at least in theoretical framing by researchers, than the GDD design. Some examples of CITS research include Wong, Cook, and Steiner (2015), who apply CITS to the impact of No Child Left Behind on public versus private schools, and Dee and Jacob (2011), who apply CITS to the impact of No Child Left Behind on states with and without prior school accountability policies.

Beyond the educational setting, Bloom and Riccio (2005) use CITS to examine the effects of an employment program aimed at public housing residents. Though its use has been more limited, CITS has been shown to produce consistent and unbiased results in line with those of regression discontinuity analyses, which more closely imitate experimental designs (Somers et al., 2013). Perhaps CITS has not seen widespread usage because it requires sufficient number of baseline data points to be able to establish a reliable baseline trend; some literature suggests that four data points are necessary at minimum (Somers et al., 2013). The three data points used here fall just below this threshold. Still, the application of CITS analyses here, while limited to three baseline years of data, is intended to provide a close approximation of an experimental analysis.

Fuzzy regression discontinuity. Regression discontinuity is a research design and analysis approach typically applied in situations in which a predetermined cutoff value of some continuous variable determines subjects' assignment to treatment or control. Under certain assumptions of appropriate functional form, regression discontinuity analyses can produce statistically unbiased estimates of a treatment effect—the only quasi-experimental method proven to do so (Shadish et al., 2002). In a typical regression discontinuity analysis, regression lines are fit separately for the treatment and control groups near the cutoff value; the effect is measured as the difference between the two lines. The intuition supporting the analysis rests on the notion of measurement error and the imprecision of the cutoff value: that subjects near the cutoff are randomly distributed above and below the cutoff given the random error inherent in measurement;

thus comparing the two groups is akin to comparing equivalent groups formed through random assignment.

In the context of the SDP policy decision examined in Research Question (2), no explicit cutoff was established to determine whether a school would retain its full-time school police officer or whether its position would be turned into a part-time position. It is a reasonable premise, however, that the schools which were selected to face part-time officer vacancies were those with fewer disciplinary issues and safer school environments. Examining the extent to which this is true for baseline variables such as prior school-level academic performance and/or incident rate is the first step in applying regression discontinuity methods to the current dataset. As described in Chapter 7, this paper examines the school-level rate of incidents involving law enforcement as a potential variable from which to generate an artificial treatment assignment cutoff. The cutoff is determined where a given value maximizes the margin between the proportion of treatment and control schools on each side; this variable is also compared to others to determine whether any other variable could produce a larger margin.

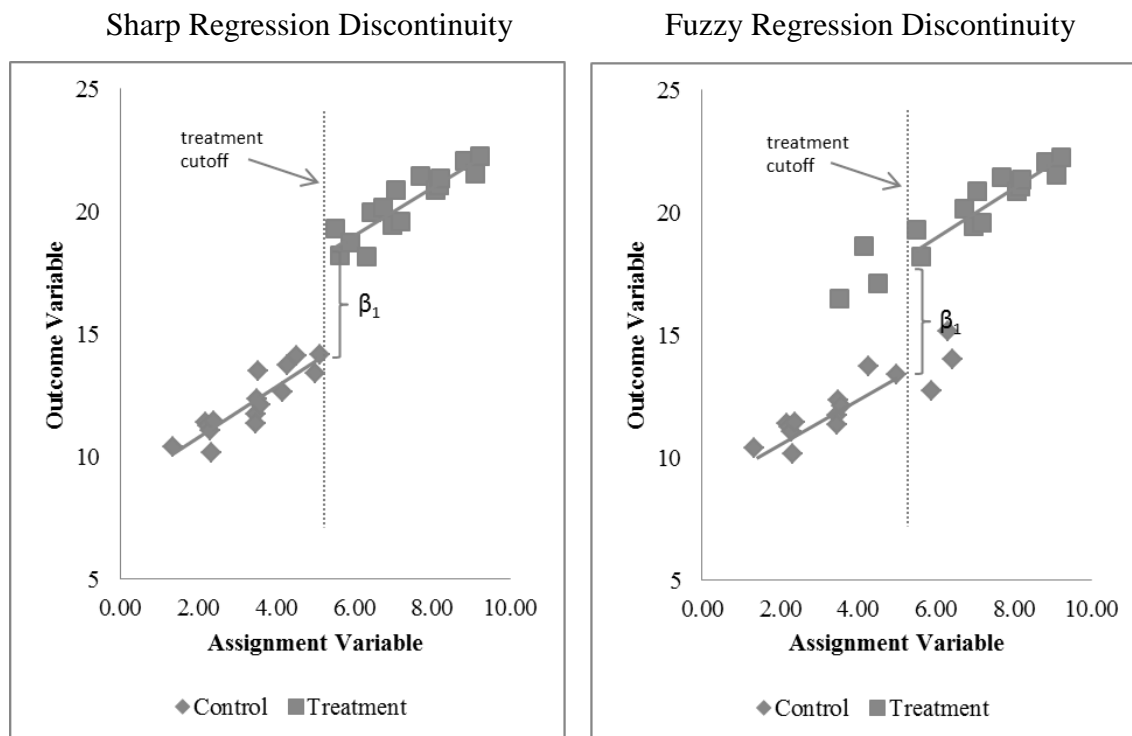
While it seems reasonable to expect that treated schools generally displayed a higher or lower value on one of these variables, it is not reasonable to expect that only treated schools will fall above/below the given value, and all control schools will fall on the opposite end of the spectrum. Thus, a sharp regression discontinuity design cannot be applied to the data. Instead, a fuzzy regression discontinuity design (FRD) is appropriate. FRD methods are used when a sharp cutoff is intended, but some crossover or misassignment from treatment to control groups and vice versa occurs. In other words,

the probability of treatment given a subject's measurement on the cutoff variable is not strictly 0 or 1. The fewer the proportion of participants who are misassigned, the less misassignment threatens to bias causal estimates. Yet just as with a randomized study, certain methods can be applied to correct for misassignment and estimate the effect of the treatment on those subjects that received the treatment; most commonly, this involves the use of instrumental variables.

After determining a viable cutoff value, linear FRD strategy begins with equation E9, an expression of basic regression discontinuity (Jacob, Zhu, & Bloom, 2012):

$$E9. \quad Y_s = \alpha + \beta_1 D_s + \beta_2 Z_s + \beta_3 D_s Z_s + \varepsilon_i$$

In this equation, Y_s represents the outcome for school s ; D_s represents treatment status for school s ; and Z_s represents the assignment variable value for school s . The local average treatment effect is represented by β_1 ; β_2 represents the relationship between the assignment variable and the outcome in the control group, and β_3 represents the potential change in the slope of that relationship due to treatment. Like basic regression discontinuity, FRD rests on the assumption that the functional form has been correctly specified, i.e., the relationship between the assignment variable and the outcome is correctly modeled as linear, quadratic, cubic, etc. The analysis thus compares linear models with quadratic models for each outcome variable, calculating F statistics on the addition of the quadratic terms and favoring the linear model where this term is not significant. As a visual check on this relationship, pretest values of each outcome variable are simultaneously plotted against the assignment variable to visually inspect the credibility of this relationship between the assignment variable and the response function.



Notes. In the hypothetical examples above, the cutoff indicates which units should receive treatment, i.e., those above the cutoff should receive the treatment, and those below should not. In the right panel, all units complied with this assignment. In the left panel, some units below the cutoff received treatment, and some above the cutoff did not, creating the setting for a fuzzy regression discontinuity (FRD) analysis.

Figure 4. Graphical Representation of Sharp Regression Discontinuity and FRD

Similarly, nonparametric local linear regression is applied to posttreatment values of the outcome data, to visually explore the potential for treatment effects. Local linear regression can be informative in a regression discontinuity framework, as the method estimates a regression using only data points near the cutoff, but is best used as a visual complement to parametric estimation, as it is not possible to determine from an extant dataset which introduces less bias (Lee & Lemieux, 2010).

Figure 4 graphically represents a hypothetical basic regression discontinuity analysis. In the left-most panel, subjects have been assigned to treatment based on their

scores along a continuum, with “5” as the hypothetical cutoff; all those with scores greater than 5 have received treatment, and β_1 represents the treatment effect at the cutoff. In the right-most panel, the sharp assignment of subjects to treatment above scores of 5 no longer holds; in this FRD case, some subjects with scores lower than 5 received treatment, while others with scores greater than 5 did not receive treatment.

Given the fuzzy nature of misassignment in a FRD, equation E9 is not sufficient. In this basic form, the average treatment effect represents a measure of the effect for all subjects intended to receive the treatment—e.g., all those above or below the cutoff—irrespective of whether or not they received the treatment. To account for misassignment, a local average treatment effect must instead be calculated, to represent the average effect for those subjects that actually received the treatment (Imbens & Angrist, 1994). This treatment effect is calculated as the difference in mean outcomes for the treatment and comparison groups (β_1), divided by the difference in treatment receipt rates for both groups near the cutoff. This calculation is commonly expressed by equation E10 (Hahn, Todd, & Van der Klaauw, 2001):

$$\text{E10. } \frac{\lim_{z \uparrow z_c} E[Y_S | z_S = z_c] - \lim_{z \downarrow z_c} E[Y_S | z_S = z_c]}{\lim_{z \uparrow z_c} E[D_S | z_S = z_c] - \lim_{z \downarrow z_c} E[D_S | z_S = z_c]}$$

The numerator of equation E10 represents the intent-to-treat effect in the regression discontinuity framework; the denominator represents the proportion of intended treatment schools that actually received the treatment minus the proportion of intended control schools that actually received the treatment.

This analysis runs parallel to analyses of instrumental variables in experimental settings with noncompliance, and requires similar assumptions. First, that the treatment assignment is correlated with the likelihood of receiving the treatment; and second, that the treatment assignment is only related to the outcome through receipt of the treatment (Lee & Lemieux, 2010). The analysis also requires the assumption that no external effect potentially confounds the treatment effect at the cutoff, i.e., any estimated effect is the result of the treatment and no external effect. To provide some evidence on this possibility, compositional differences of the two groups at the cutoff is examined.

One concern with the application of FRD to Research Question (2) is the statistical power of the analysis. Given the small treatment sample size of 28 schools, the minimum detectable effect size (MDES)—with an 80 percent chance (power) of producing a statistically significant treatment effect estimate at the 0.05 level, assuming a balanced normal density distribution across the values of the cut variable and a constant rate of crossovers of 5 percent—is 0.55 (Jacob et al., 2012). Modifying the assumptions to an overall crossover rate of 5 percent with a 20 percent crossover rate clustered at the cutoff increases the MDES to 0.77. These MDES would represent relatively large effects for the average educational study, and thus some concern remains that the FRD analysis is underpowered. Still, the calculations are based on hypothetical assumptions surrounding a largely unstudied construct—police in schools; perhaps the effect of police in schools is as large as or larger than these minimum effect sizes, or perhaps the actual analysis will be able to identify statistically significant effect sizes of smaller degrees

given its actual parameters. At minimum, the FRD analysis could prove valuable in an exploratory framework that could inform future similar but larger-scale analyses.

FRD analyses have not seen extensive use in education literature, though they are common in other fields. One exception is research by Shobo, Wong, and Bell (2014), who examine a merit-based program for college tuition aid and find significant positive effects on various students' outcomes; because the aid is based on multiple criteria, a FRD design could be applied to the setting. Other education research has used basic regression discontinuity approaches, such as Wong, Cook, Barnett, and Jung (2008), who examine state pre-school programs using date of birth as an enrollment cutoff and find mixed results by state and student outcome. While the use of these methods is limited in education thus far, their use is likely to grow, as education provides a setting in which treatment cutoffs based on age and/or performance are common.

CHAPTER 5: PENNSYLVANIA SCHOOL DEMOGRAPHICS AND BEHAVIOR

To understand the effects of disciplinary policy changes on student and administrative behaviors, this paper contrasts behavior in the School District of Philadelphia (SDP) with that in other groups of schools across Pennsylvania. These groups are outlined in Table 2 of Chapter 4; some are purposively selected existing groups, and others are uniquely sampled from these groups using propensity score matching. This chapter examines demographics and school-level outcomes of each of the purposive groups as well as the state as a whole. The descriptive analyses presented in this chapter are not intended to serve as evidence of actual relationships or differences among groups or their outcomes; instead, they are intended to provide some understanding of the background and operation of the Pennsylvania schools and the school groups included in the analysis.

Demographics included in the descriptive analysis in this chapter include total school enrollment, school-level proportions of race/ethnicity, gender, and students receiving free or reduced-price lunch; and pupil-teacher ratios. The chapter traces aggregate trends of each group across the four categories of school-level outcomes examined in the research questions: (1) severe student misconduct, represented by number of incidents and number of incidents involving law enforcement per 1,000 students; (2) administrative reactions, represented by out-of-school suspensions, disciplinary transfers, expulsions, and arrests per 1,000 students; (3) student academic performance, represented by percentages of students achieving or surpassing proficiency

in reading and mathematics; and (4) student truancy, defined as the average number of students in 100 absent on a daily basis. As these overview analyses concern the student population of Pennsylvania and descriptive comparisons of groups within the state, no inferential tests of statistical differences across groups are included in this chapter.

The chapter begins with an overview of current student demographics in Pennsylvania and statewide trends in student and administrative behavior since 2001-2002. Next, following the bounds of the research questions, the chapter disaggregates these data to examine school groups and timeframe relevant to Research Question (1). This includes contrasting behavior in the SDP with extant groups of Philadelphia charter schools, Pittsburgh schools, and all other Pennsylvanian schools from 2009-2010 through 2013-2014. The chapter then addresses schools and timeframe relevant to Research Question (2). This includes behavior in SDP schools affected by police officer vacancies versus all other SDP schools from 2012-2013 through 2013-2014. Finally, the chapter describes the strengths and limitations of the propensity score matched samples drawn for each research question, and whether or not these represent viable improvements over the purposive samples. Evaluation of the matches is based on reduction of matched group mean differences and variance ratios, as per the discussion of propensity score matching presented in Chapter 4.

The Pennsylvania Context: Trends and Demographics

The state of Pennsylvania enrolled 1.7 million students in 2,923 schools in the 2013-2014 school year (the most recent year of data available). As shown in Table 3, primary grades comprised the largest proportion of these schools, enrolling 46 percent of

Table 3. *Pennsylvania School Demographics, 2013-2014 School Year*

	All Schools	Primary Schools	Middle Schools	High Schools	Other Schools
Number of Schools	2,923	1,695	550	591	87
Total Enrollment	1,710,240	783,374	333,759	503,103	90,004
Mean School Size	585	462	607	854	1035
Race/Ethnicity %					
% Asian	0.03	0.03	0.03	0.02	0.02
% Black	0.14	0.15	0.11	0.13	0.33
% Hispanic	0.08	0.09	0.07	0.05	0.16
% White	0.72	0.69	0.77	0.77	0.47
% Other	0.03	0.04	0.02	0.02	0.03
Female %	0.49	0.49	0.49	0.49	0.50
Free or Reduced-Price Lunch %	0.45	0.47	0.41	0.40	0.58
Mean Pupil/Teacher Ratio	15.14	15.39	14.30	15.24	15.02

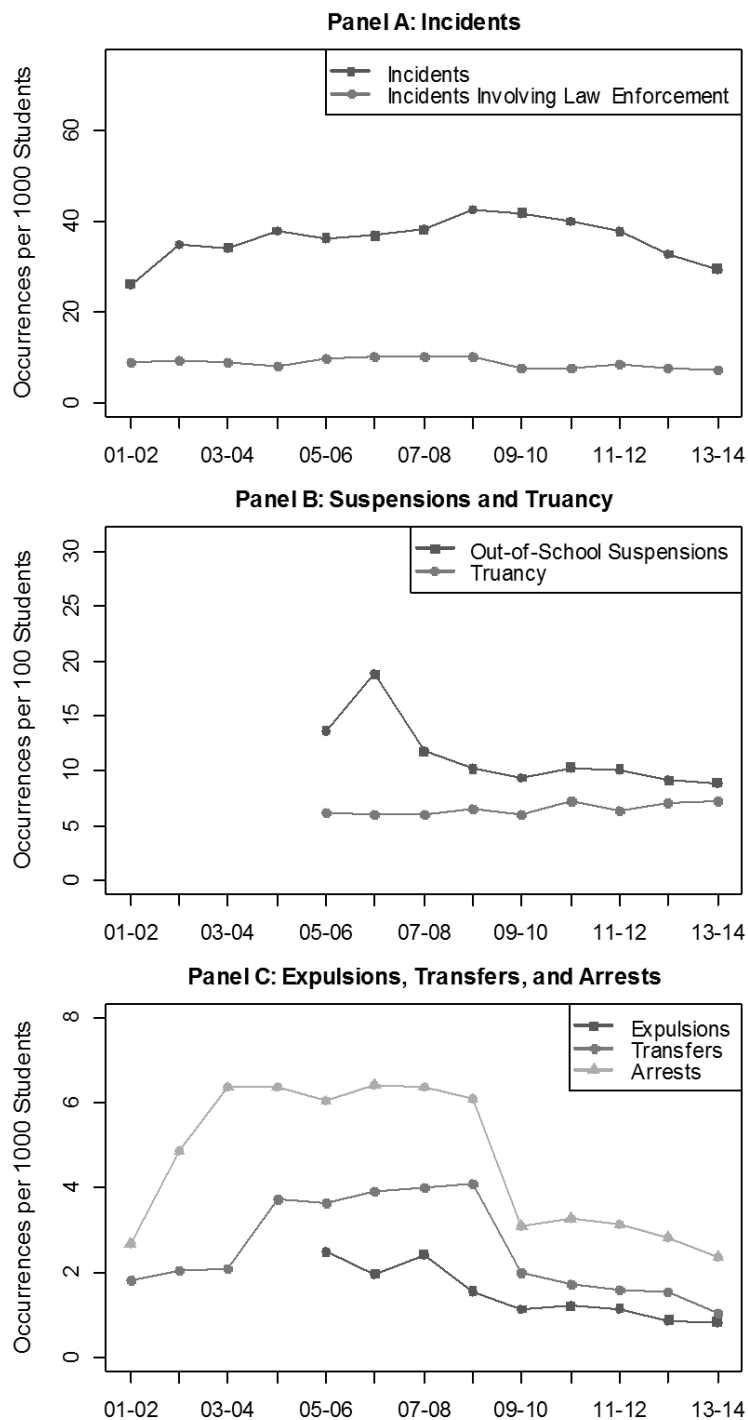
Notes. Figures represent school-level averages for all schools across Pennsylvania.

all students. Across the entire state in the 2013-2014 school year, 72 percent of the state's students were white, while just 14 percent were black and only 8 percent were Hispanic. Less than half of the student body—45 percent—received free or reduced-price lunch. Younger students were more likely to be minority and receive free or reduced-price lunch. For example, 69 percent of primary school students were white, compared with 77 percent of high school students, while 9 percent of primary school students were Hispanic, compared with just 5 percent of high school students; 47 percent of primary school students received free or reduced-price lunch, while just 40 percent of high school students did so. These data suggest that more diverse, lower income cohorts of Pennsylvanian students are replacing less diverse, higher income cohorts with each passing year. A quick check of Pennsylvania demographics in previous years confirms

this trend: five years prior to the 2013-2014 school year, 74 percent of students were white and just 41 percent received free or reduced-price lunch; ten years prior, 77 percent of students were white and just 30 percent of students received free or reduced-price lunch.

Just as demographics shifted across the state since the early 2000s, so too did student outcomes. These trends are depicted in Figure 5. (Note that, for the purpose of visually plotting the data, the variable representing out-of-school suspensions is rescaled to represent number of suspensions per 100 students.) Overall, the data show a heightened plateau of student misconduct and severe administrative responses in the late 2000s, with such occurrences decreasing since the early 2010s. Specifically, the number of student incidents, out-of-school suspensions, expulsions, transfers, and arrests all peaked between the 2006-2007 and 2008-2009 school years. On the other hand, the number of incidents involving law enforcement remained largely stable across the time period. As shown in Figure 6, this suggests a higher rate of arrests per incident involving law enforcement in the late 2000s; in other words, arrests were made in as many as 80 percent of incidents involving law enforcement in the late 2000s, compared with just 40 percent of those incidents in the early 2010s, while the overall rate of arrests per all incidents remained relatively stable, below 20 percent across all years of data.

A complex picture of student outcomes can be drawn from these data: More student incidents were recorded in the late 2000s than prior to or after these years, but not the types of incidents that would entail the involvement of law enforcement. Still, for those incidents which did involve law enforcement, students were arrested at higher rates



Note. No data were available on suspensions, expulsions, and truancy until the 2005-2006 school year.

Figure 5. Pennsylvania Trends in Student Outcomes, 2001-2002 to 2013-2014

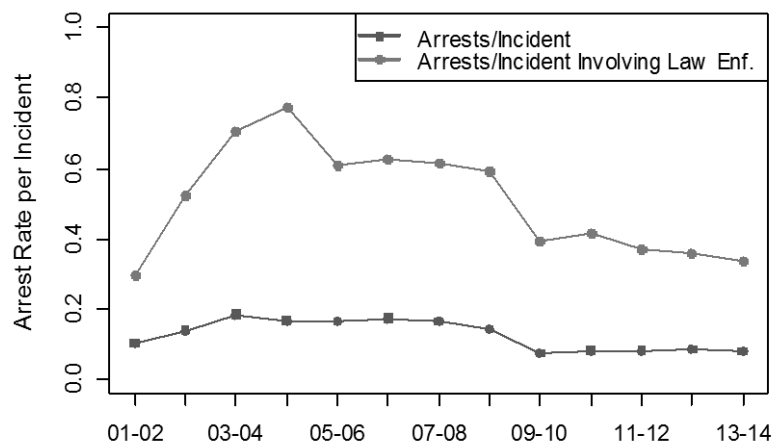


Figure 6. Arrests per Incident Type

in these same years. Concurrently, severe administrative responses increased, including out-of-school suspensions, expulsions, and transfers. Thus less severe incidents and severe administrative responses peaked at the same time.

The increased activity seen in the late 2000s coincided with the nationwide height of zero tolerance policies and strict disciplinary codes, as described in Chapter 2. More recent years have seen a national emphasis on less severe punishments and increased behavioral supports. Though the data thus suggest a correlational relationship between severe disciplinary codes, student misconduct, and severe administrative responses to behavior, no causal relationship can be drawn from these basic descriptive trends. To build evidence for or against such a causal relationship, this paper contrasts the SDP's policy changes with other groups of schools within Pennsylvania. As a first step in this

contrast, the next section examines Philadelphia alongside these other comparison groups.

Research Question (1) School Groups

To investigate the potential causal relationship between zero tolerance disciplinary policies and student and administrative behavior, this paper contrasts SDP schools with Philadelphia charter schools, Pittsburgh schools, and other schools across Pennsylvania. Baseline demographics and trends in student outcomes for these groups are described below in order to provide contextual information on each group.

Demographic Baseline

Table 4 tabulates baseline demographics of each of these purposive groups. The largest district in the state, the SDP comprised 8.4 percent of all Pennsylvania students in the 2011-2012 school year, the baseline year for analyses pertaining to Research Question (1). The Philadelphia charter sector and Pittsburgh Public Schools were smaller, enrolling just 2.6 and 1.4 percent of Pennsylvania schools, respectively. Most SDP students were enrolled in primary schools (67 percent). On average, the district enrolled larger primary schools and smaller middle and high schools than the rest of the state.

The SDP enrolled larger proportions of low income and minority students than the larger state, with black students comprising 62 percent and Hispanic students comprising 16 percent of the student body. While the Philadelphia charter sector saw similar proportions of these students (71 percent and 13 percent, respectively), Pittsburgh's student body was 57 percent black and just 2 percent Hispanic, and the rest of Pennsylvania was just 9 percent black and 7 percent Hispanic. A full 82 percent of

Table 4. *RQ1 Purposive Sample Baseline Demographics*

	Philadelphia District	Philadelphia Charters	Pittsburgh Schools	PA Schools
Number of Schools	237	73	52	2,517
Total Enrollment				
Primary School	92,694	18,795	13,700	624,288
Middle School	11,653	2,291	2,559	306,021
High School	34,113	9,941	4,072	458,700
Other	275	12,481	2,664	49,927
Mean School Size				
Primary School	555	553	381	436
Middle School	507	382	366	610
High School	742	585	814	875
Other	275	780	666	832
Race/Ethnicity %				
% Asian	0.06	0.03	0.03	0.03
% Black	0.62	0.71	0.57	0.09
% Hispanic	0.16	0.13	0.02	0.07
% White	0.12	0.12	0.32	0.80
% Other	0.04	0.01	0.07	0.02
Female %	0.49	0.51	0.49	0.49
Free or Reduced-Price Lunch %	0.82	0.73	0.73	0.37
Mean Pupil/Teacher Ratio	15.89	16.07	15.03	14.83

Notes. Baseline demographics refers to demographics of each school prior to the policy change examined in Research Question 1 (RQ1), i.e., in the 2011-2012 school year.

Philadelphia students received free or reduced-price lunch, compared with 73 percent of students in both Philadelphia charters and in Pittsburgh, and with 37 percent of students in remaining schools across Pennsylvania.

These data confirm that Philadelphia charters are demographically similar to Philadelphia traditional schools and are thus a reasonable intact comparison group—

though this comparison is by no means perfect, as described in Chapter 4. Pittsburgh schools are less similar, enrolling almost no Hispanic students, though they do comprise a similarly large black population of students and only a slightly smaller proportion of students on free and reduced-price lunch. As a whole, other schools across Pennsylvania are quite distinct from Philadelphia schools. Propensity score matching is employed later in this chapter to render these groups more similar to the SDP schools; here, the chapter continues with a comparison of behavioral trends in each of these purposively selected groups.

Behavioral Trends

School-level mean trends in student outcomes for each of these groups are pictured in Figure 7. The panels show that both student misbehavior and administrative responses decreased in the SDP in the 2012-2013 and 2013-2014 school years, i.e., after the change in disciplinary policy. The SDP saw each measured student outcome decrease between 2011-2012 and 2012-2013, and again decrease further between 2012-2013 and 2013-2014. This was not necessarily the case for the other comparison groups. Pittsburgh in particular saw considerable volatility in its number of incidents, suspensions, expulsions, and transfers, suggesting no particular directional pattern. Both Philadelphia charters and the larger body of Pennsylvanian schools did not experience much volatility in their trends, maintaining relatively stable trajectories of student outcomes from the 2009-2010 school year through the 2013-2014 school year.

The data suggest that the SDP indeed experienced distinct shifts in student outcomes between the 2011-2012 and the 2012-2013 school years and after. The data

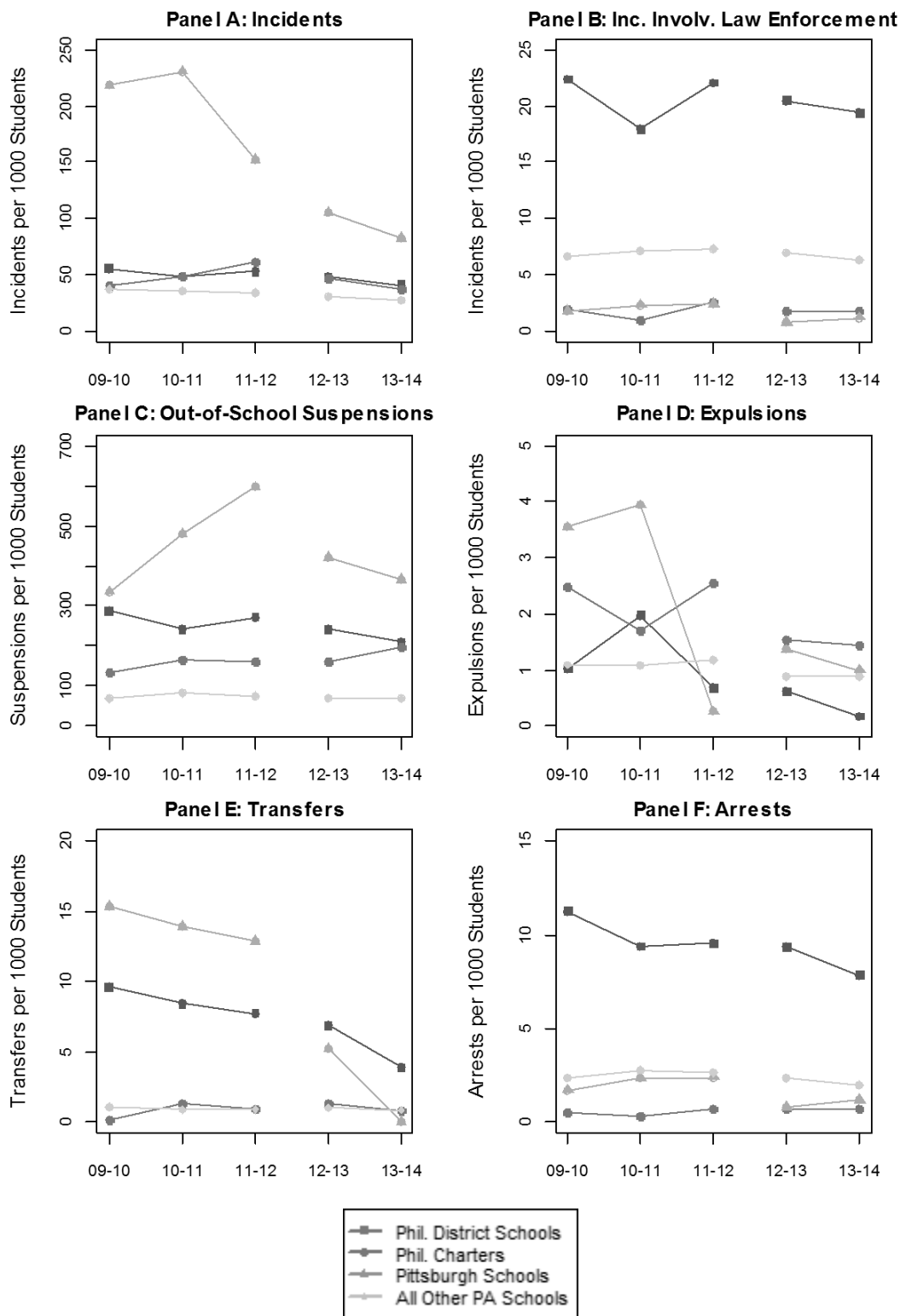


Figure 7. School-level Mean Student Outcomes by RQ1 Group and Year

cannot be taken at face value to suggest that these shifts resulted from the SDP's 2012 disciplinary policy change; to explore this possibility, the study turns to the inferential analyses of the next chapter. Still, Figure 7 raises the possibility of these relationships, at minimum highlighting the unique trajectory of student outcomes in the SDP across the timeframe, as other groups of schools across Pennsylvania did not follow the same outcome trajectories across the same years.

Research Question (2) School Groups

To investigate the potential causal relationship between a reduction in policy officer presence in schools and student and administrative behavior, this paper contrasts SDP schools which saw a reduction in police officer staffing with other non-affected SDP schools. Baseline demographics and trends in student outcomes of each of these groups are described below.

Demographic Baseline

Table 5 tabulates baseline demographics of these two groups of SDP schools, i.e., those "treatment" schools affected by the staffing change, and those "control" schools not affected by the staffing change. Affected schools enrolled 12 percent of the SDP student body in 2012-2013, the baseline year of data for this analysis. Primary schools were more likely to see their police officer staff reduced, with primary school students making up 88 percent of the affected schools compared to just 65 percent of the non-affected schools. This suggests that the SDP perceived law enforcement as less critical in primary grades than in the middle and high school years.

Table 5. RQ2 Purposive Sample Baseline Demographics

	Affected by Staffing Change	Not Affected by Staffing Change
Number of Schools	27	201
Total Enrollment		
Primary School	13,512	76,074
Middle School	502	9,794
High School	1,431	30,606
Mean School Size		
Primary School	587	551
Middle School	251	466
High School	716	729
Race/Ethnicity %		
% Asian	0.10	0.06
% Black	0.48	0.62
% Hispanic	0.17	0.16
% White	0.19	0.12
% Other	0.06	0.05
Female %	0.49	0.48
Free or Reduced-Price Lunch %	0.83	0.85
Mean Pupil/Teacher Ratio	16.37	16.51

Notes. Baseline demographics refers to demographics of each school prior to the policy change examined in Research Question 2 (RQ2), i.e., in the 2012-2013 school year.

Affected schools were more likely to enroll white students and less likely to enroll black students: white students comprised 19 percent of the student body of affected schools versus just 12 percent in non-affected schools; black students comprised 48 percent of affected schools versus 62 percent of non-affected schools. That the affected schools enroll proportionally fewer black students than non-affected schools is of particular note, given the narrative of the ‘school to prison’ pipeline described in Chapter

2 and its perceived consequences for black—particularly black male—students. The data suggest that schools with greater white student populations and smaller black student populations are those which the SDP perceived as safer environments in which to reduce police officer presence.

Behavioral Trends

School-level mean trends in student outcomes for these two SDP groups are portrayed in Figure 8. Affected schools consistently saw less student misbehavior and fewer administrative responses to misbehavior than non-affected schools, both prior to and after the staffing changes. Thus affected schools were in fact those schools with fewer behavioral issues at baseline, indicating that the SDP did take into these and other student outcomes when selecting schools in which to reduce police presence. This is key to the fuzzy regression discontinuity design used to analyze these data in Chapter 7. Further, in every instance, the gap in outcomes between the two groups decreased from 2012-2013 to 2013-2014. In the case of student incidents, suspensions, expulsions, and transfers, the number of occurrences per 1,000 students decreased for both groups in 2013-2014 but decreased at a smaller rate of change for the affected schools. In the case of student incidents involving law enforcement and arrests, occurrences decreased for the non-affected schools but increased for the affected schools. Thus those schools with less police officer presence saw more incidents involving law enforcement and more arrests, while schools with unchanged police officer staffing saw less.

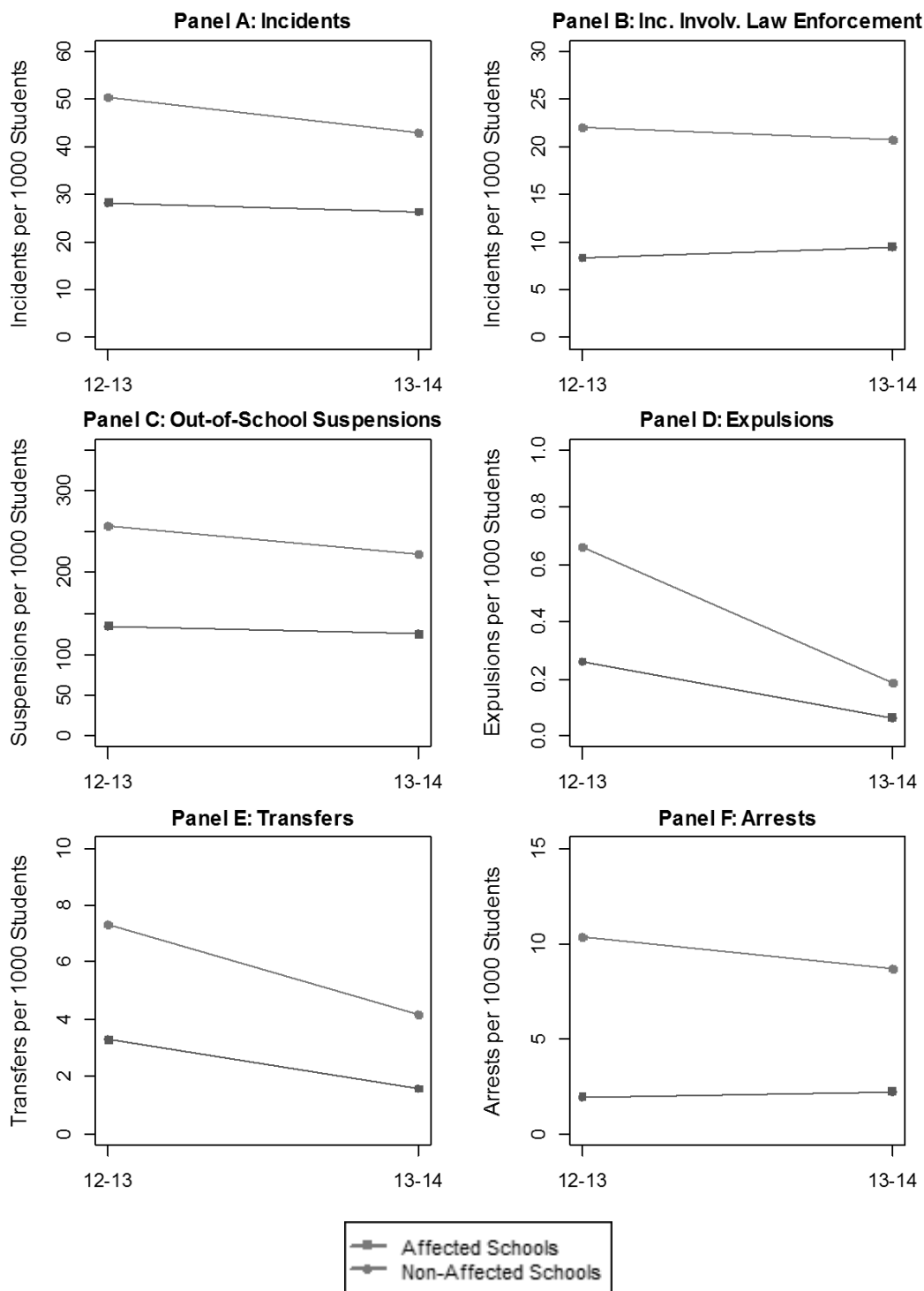


Figure 8. School-level Mean Student Outcomes by RQ2 Group and Year

Propensity Score Matched Samples

As described in Chapter 4, baseline differences among treatment and comparison groups call into question the validity of inferences drawn from their comparison. Non-equivalent group comparisons can lead to biased results, potentially suggesting incorrect causal relationships—hence why the descriptive analyses in this chapter are not taken as evidence of such relationships. Propensity score matching can be used to create alike samples to strengthen the validity of such statistical conclusions. Using the methods outlined in Chapter 4, matched samples were created by drawing schools from each of the purposive groups of schools. The matching process is evaluated based on differential group mean and variance. Viable group matches are those with a minimal standardized mean difference in propensity scores, a variance ratio of propensity scores near one, and variance ratios of all group demographic variables near one.

Evaluation of Matched RQ1 Groups

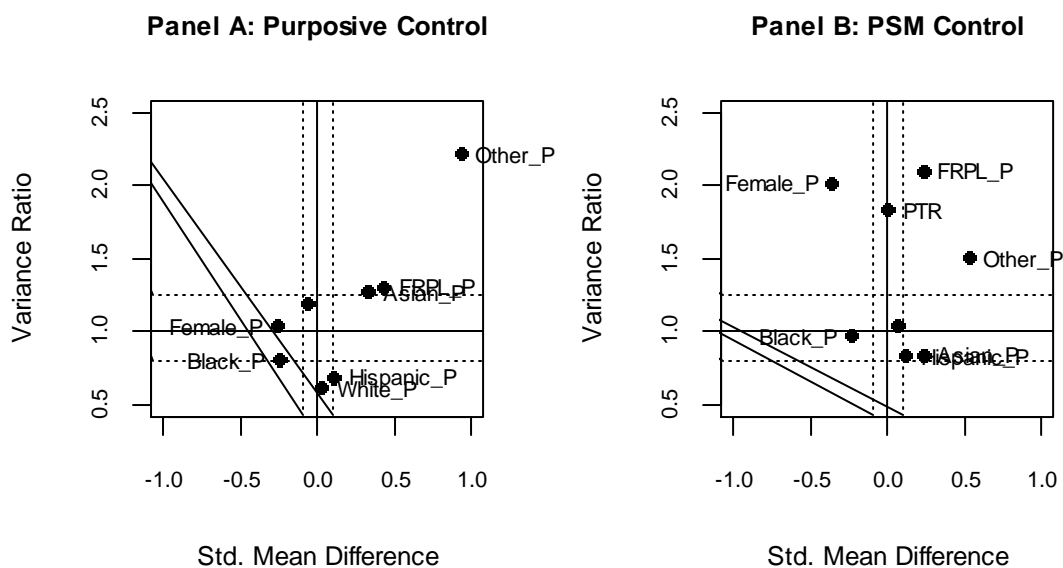
Table 6 provides these balance statistics for each matched comparison relevant to Research Question (1) before and after matching. In no case did the groups' propensity scores return recommended magnitudes of mean differences and variance ratios prior to matching. Still, the matching process proved satisfactory for two of the three pairs.

While Philadelphia charter schools were already relatively similar to the SDP schools across school-level characteristics prior to matching, matching represented an improvement. Prior to matching, the standardized mean difference of the groups' propensity scores was 1.48; after matching, it was 0.88, closer to zero. At the same time, the variance ratio of the groups' propensity scores prior to matching was 0.56; after

Table 6. Balance of RQ1 Propensity Score Matches

	Purposive Sample		Matched Sample	
	Standardized Mean Difference	Variance Ratio	Standardized Mean Difference	Variance Ratio
<i>Panel A: SDP versus Philadelphia Charter Schools</i>				
Calculated Propensity Scores	1.48	0.56	0.88	0.74
% Asian	0.33	1.27	0.23	0.84
% Black	0.25	0.81	0.23	0.98
% Hispanic	0.11	0.69	0.12	0.84
% White	0.02	0.62	0.06	1.05
% Other	0.93	2.22	0.53	1.51
% Free/Reduced-Price Lunch	0.43	1.31	0.23	2.11
% Female	0.26	1.04	0.36	2.02
Pupil/Teacher Ratio	0.07	1.20	0.00	1.84
<i>Panel B: SDP versus Pittsburgh Schools</i>				
Calculated Propensity Scores	2.96	0.16	1.56	0.10
% Asian	0.46	8.32	0.47	7.45
% Black	0.19	1.64	0.36	1.51
% Hispanic	0.95	94.05	0.95	56.65
% White	0.98	0.78	0.31	0.89
% Other	1.07	1.04	0.42	1.14
% Free/Reduced-Price Lunch	0.47	2.62	0.21	2.73
% Female	0.01	2.82	0.05	3.99
Pupil/Teacher Ratio	0.30	1.08	0.03	1.19
<i>Panel C: SDP versus Other PA Schools</i>				
Calculated Propensity Scores	2.59	5.07	0.70	1.01
% Asian	0.45	6.10	0.11	1.15
% Black	2.04	3.44	0.24	0.82
% Hispanic	0.53	2.58	0.11	0.98
% White	3.06	0.61	0.53	0.52
% Other	0.56	1.18	0.24	1.12
% Free/Reduced-Price Lunch	1.98	1.04	0.48	1.09
% Female	0.06	6.51	0.01	1.95
Pupil/Teacher Ratio	0.27	0.37	0.28	0.72

Notes. RQ1 = research question (1). Ideal parameters of matched sample include standardized mean differences near zero and variance ratios between .80 and 1.25.

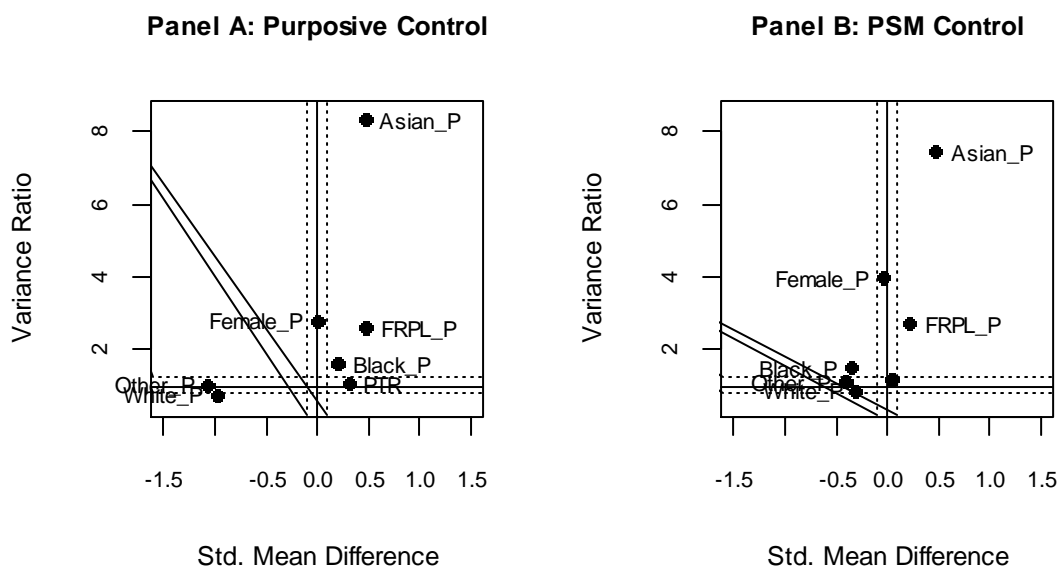


Notes. Standardized mean differences of baseline characteristics across treatment and control groups are plotted against the variance ratios of the same characteristics. Dotted lines suggest boundaries of ideal values for PSM samples: between $-.10$ and $.10$ standardized mean differences, and between $.80$ and 1.25 variance ratios. RQ1 = research question (1); PSM = propensity score matched.

Figure 9. Balance Statistics of SDP and Philadelphia Charters, RQ1

matching, it was 0.74, closer to one. The demographic mean differences and variance ratios of this matched comparison are also plotted in Figure 9. As the figure shows, many of the demographic contrasts approached ideal values even prior to matching; after, their divergence from a mean difference of zero and a variance ratio of one was similar. Given the similarity at baseline and the improvement in propensity score balance with matching, the matched sample is considered a viable equivalent group.

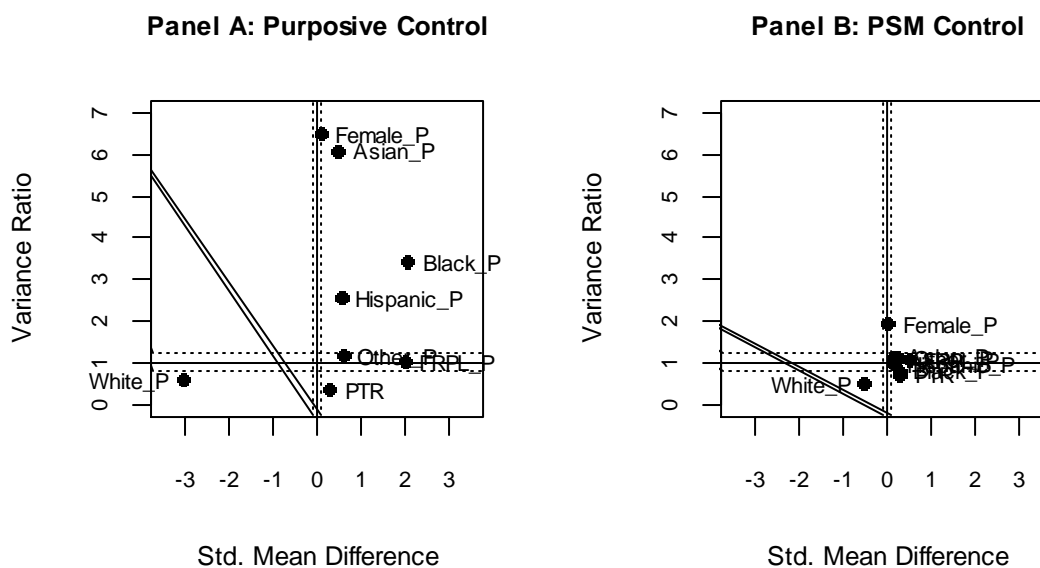
Matching was not successful when using Pittsburgh schools as the comparison group. This is likely because Pittsburgh, a smaller district than the SDP, enrolls almost no



Notes. Standardized mean differences of baseline characteristics across treatment and control groups are plotted against the variance ratios of the same characteristics. Dotted lines suggest boundaries of ideal values for PSM samples: between -0.10 and 0.10 standardized mean differences, and between 0.80 and 1.25 variance ratios. One variable is omitted from the graph due to plot margins: the variance ratio of the proportion of Hispanic students in each group, which equaled 94.05 prior to matching and 56.65 after matching. RQ1 = research question (1); PSM = propensity score matched.

Figure 10. Balance Statistics of SDP and Pittsburgh Schools, RQ1

Hispanic students and almost three times as many white students as the SDP (see Table 4). Pittsburgh simply does not provide enough schools demographically similar to Philadelphia to form equivalent matches. As a result, the matched group of schools did not result in propensity scores or demographic variables with standardized mean differences near zero nor variance ratios near one, as shown in Table 6. These demographic values are plotted in Figure 10, with the exception of the variable representing percent Hispanic, given its extremely large value. As the figure shows, almost no variables fell within ideal value ranges either before or after matching. Given

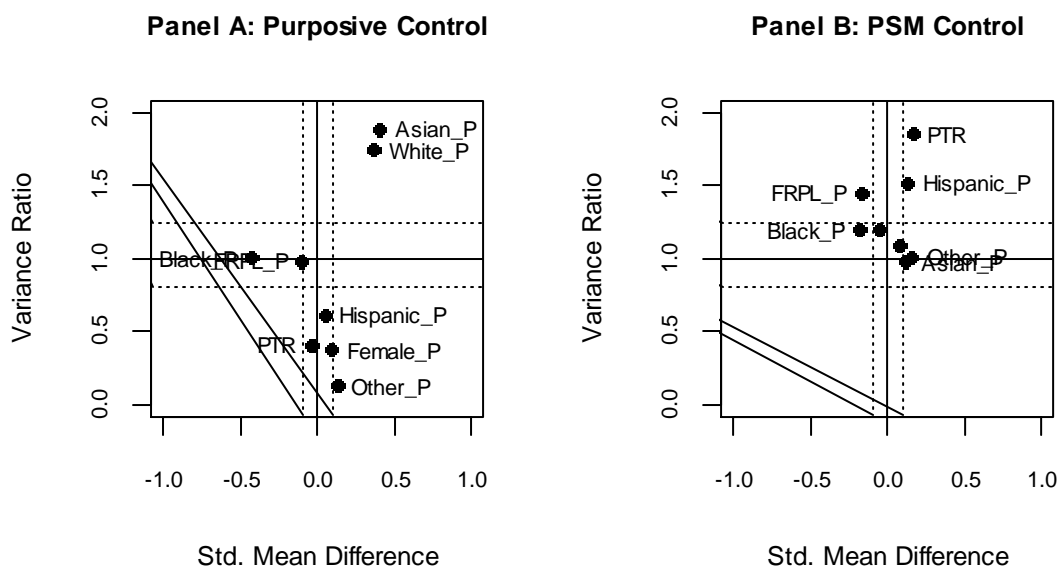


Notes. Standardized mean differences of baseline characteristics across treatment and control groups are plotted against the variance ratios of the same characteristics. Dotted lines suggest boundaries of ideal values for PSM samples: between $-.10$ and $.10$ standardized mean differences, and between $.80$ and 1.25 variance ratios. RQ1 = research question (1); PSM = propensity score matched.

Figure 11. Balance Statistics of SDP and Pennsylvania Schools, RQ1

these results, Pittsburgh is removed from the analysis as a feasible comparison group, both as a purposive sample and as a propensity score matched sample: as a comparable group could not be drawn from among all Pittsburgh schools, using the group as a whole would serve no purpose. Though Pittsburgh—the second largest district in Pennsylvania and a district in an urban setting—was included as a school group potentially similar to Philadelphia schools, demographically it is in fact quite different from the SDP.

On the other hand, while other Pennsylvania schools as a whole are demographically quite different from SDP schools, the large sample of schools that



Notes. Standardized mean differences of baseline characteristics across treatment and control groups are plotted against the variance ratios of the same characteristics. Dotted lines suggest boundaries of ideal values for PSM samples: between $-.10$ and $.10$ standardized mean differences, and between $.80$ and 1.25 variance ratios. RQ2 = research question (2); PSM = propensity score matched.

Figure 12. Balance Statistics of Affected and Non-Affected Schools, RQ2

comprise Pennsylvania—and their demographic variance—render these schools a useful group from which to draw matches. As shown in Table 6, the standardized mean difference of their matched propensity scores was $.70$ and their variance ratio 1.01 , even closer to zero and one, respectively, than the matched pairs drawn from the Philadelphia charter sector. The demographic contrasts are plotted in Figure 11. The reduction in both standardized mean differences and variance ratios is clear, with a number of values falling within ideal value ranges after matching. Given the marked improvements

Table 7. *Balance of RQ2 Propensity Score Matches*

	Purposive Sample		Matched Sample	
	Standardized Mean Difference	Variance Ratio	Standardized Mean Difference	Variance Ratio
<i>Panel A: SDP versus Philadelphia Charter Schools</i>				
Calculated Propensity Scores	0.64	1.53	0.06	0.93
School-level Characteristics				
% Asian	0.41	1.89	0.11	0.98
% Black	0.43	1.01	0.18	1.21
% Hispanic	0.04	0.62	0.13	1.52
% White	0.37	1.75	0.08	1.10
% Other	0.12	0.13	0.16	1.02
% Free/Reduced-Price Lunch	0.10	0.99	0.17	1.46
% Female	0.09	0.38	0.05	1.20
Pupil/Teacher Ratio	0.05	0.41	0.16	1.86

Notes. RQ2 = research question (2). Ideal parameters of matched sample include standardized mean differences near zero and variance ratios between .80 and 1.25.

resulting from the propensity score matches, the matched sample of Pennsylvania schools is considered a viable equivalent group.

While neither the propensity score matches drawn from Philadelphia charter schools nor from Pennsylvania as a whole perfectly met the ideal conditions outlined in Chapter 4, both matched groups represented an improvement over their unmatched groups, and thus both matched groups are retained for descriptive analysis in this chapter and inferential analyses in Chapter 6. The matched group drawn from all Pennsylvania schools is, by the metrics in Table 6, the most evenly matched comparison group. Still, the contextual advantage of the Philadelphia charter schools cannot be overlooked; that these schools operate in the same urban district and thus same contextual setting provides this sample some face validity of equivalence lacking in the matched Pennsylvania group.

Evaluation of Matched RQ2 Group

To address Research Question (2), propensity score matching was used to draw an equivalent sample of non-affected “control” schools from within the SDP. The matching process is again evaluated based on the standards outlined in Chapter 4 and described briefly above. Table 7 provides balance statistics of standardized mean differences and variance ratios of affected and non-affected schools before and after matching. Matching resulted in considerable improvement: the standardized mean difference in the groups’ propensity scores was reduced from 0.64 to 0.06, while the variance ratio approached one, shifting from 1.53 to 0.93. As shown in Figure 12, nearly all of the demographic contrasts fell within or very near ideal bounds. Given these values, the matched sample of “control” schools is considered a viable equivalent group.

• • •

The context provided in this chapter reveals that schools of the SDP—and those selected within it for staffing changes—were demographically distinct groups with unique trajectories of student and administrative behaviors. The chapter provides some preliminary indication that changes in disciplinary policies could have affected these behaviors, but more importantly, the chapter objectively demonstrates that student outcomes in Philadelphia followed different patterns than in other locales over the same timeframe, introducing the potential for causal effects. Further, the chapter finds matched groups of schools that approach equivalence with the “treatment” groups on baseline

measures, a key step in reducing potential bias of inferential conclusions. The next two chapters analyze these same data with the aim of producing such unbiased conclusions on the effects of disciplinary changes on student and administrative behaviors.

CHAPTER 6: REMOVAL OF ZERO TOLERANCE AND BEHAVIOR

The School District of Philadelphia (SDP) dismantled its disciplinary zero tolerance policy entering into the 2012-2013 school year. The policy—which had once outlined out-of-school suspensions, transfers, and expulsions as perfunctory responses to cascading misconduct—now permitted these reprimands only as last resorts used to address the most severe offenses. The SDP’s new policy came at a time when many school districts across the country were also reshaping their disciplinary policies. The sharp removal of zero tolerance in Philadelphia schools, however, creates an opportunity to contrast student and administrative behavioral trends before and after the policy revision with that of behaviors in other districts across the state of Pennsylvania. This chapter explores this contrast using multiple methodological approaches, addressing the first research question (RQ1) outlined in the larger study:

RQ1: To what extent did the dismantling of the SDP’s zero tolerance discipline policy affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools’ academic performance and truancy rates?

The chapter includes a discussion of preliminary analyses gauging model fit for the difference-in-differences (DD), generalized difference-in-differences (GDD), and comparative interrupted time series (CITS) analyses applied to the research question. Results of the three approaches are then examined and contrasted. A discussion of the results and their interpretation follows.

Preliminary Model Selection

Prior to estimating treatment effects, multiple forms of linear regressions were explored as potential models for the basic difference-in-differences analyses applied to RQ1, and model type selected based on model fit. Different types of linear regressions were considered given the distributions of the behavioral outcome variables, pictured in Figure 13. The plots show the distributions of the variables in the year prior to the policy change across all of Pennsylvania. Behavioral outcomes such as school-level rates of student incidents, suspensions, transfers, expulsions, and arrests all exhibit distributions that peak at or near zero and skew outward to the right. The shapes of these distributions result from the prevalence of school-level misbehavior and punishments: many schools see few incidents of misconduct and thus non-normal distributions of these outcomes. Only the variables representing mathematics and reading proficiency rates exhibit approximately normal, though skewed, distributions; note that ceiling and floor effects for these variables are largely absent, as few schools approach zero or 100 percent proficiency.

While ordinary least squares regression does not assume that the outcome variable is normally distributed, it does require that the residuals of the given model are normally distributed. An outcome variable with a highly skewed distribution holds the potential to produce residuals with a similarly skewed distribution given inadequate estimation of the outcome variable. For this reason, the fit of the DD models estimated using ordinary least squares regression—premised on the normal, or Gaussian, distribution of random variables—is compared with linear regressions based in both Poisson and negative

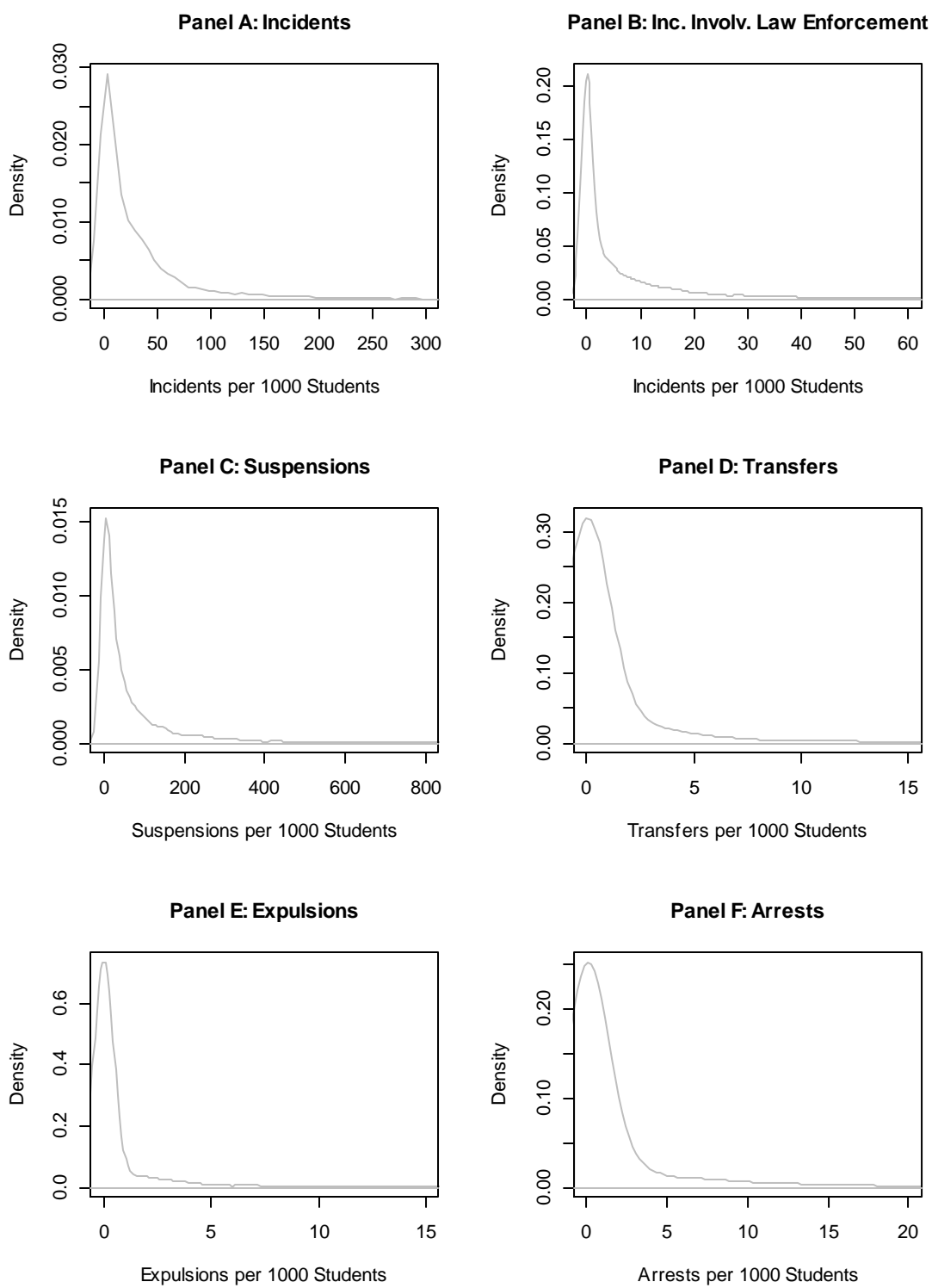


Figure 13. Outcome Variable Distributions, All PA Schools in 2012-2013

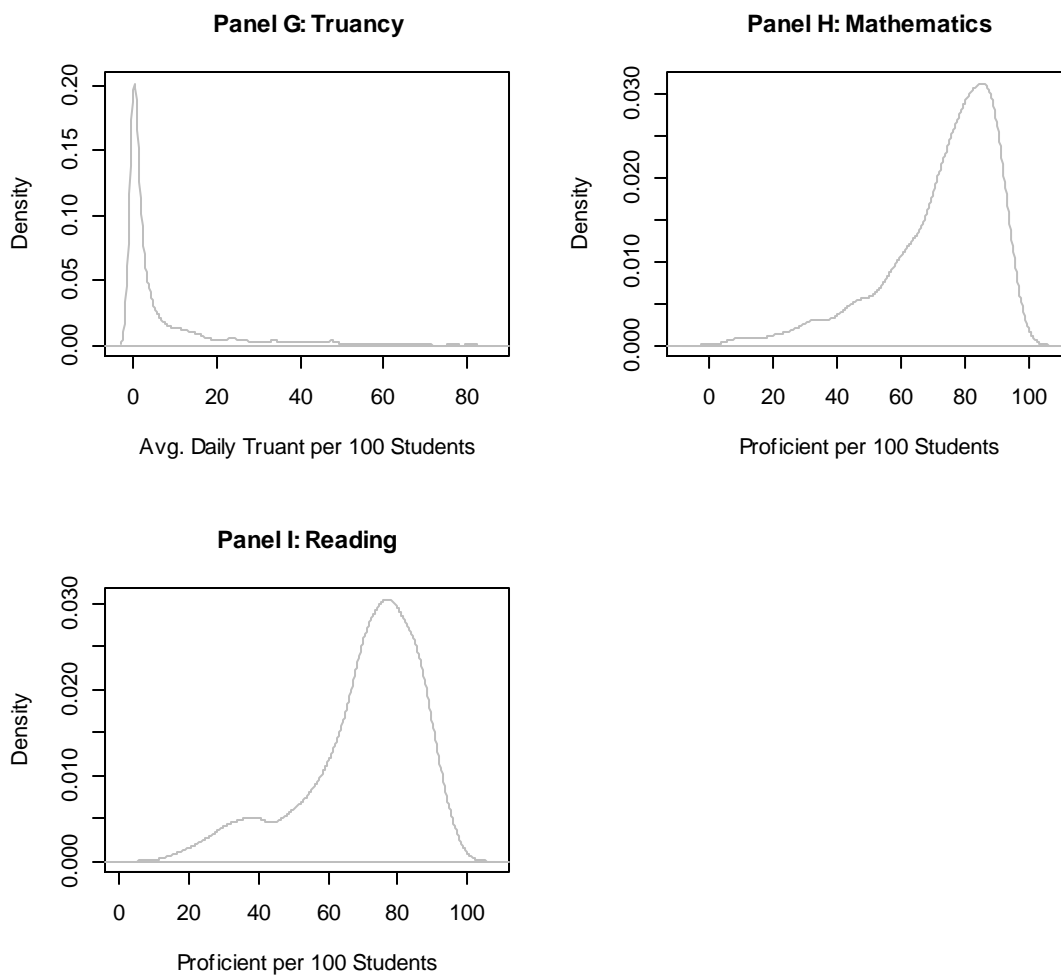


Figure 13, continued

binomial distributions, which follow more closely that of skewed data and which can be more appropriate for modeling count data. All models are estimated as generalized linear models using maximum likelihood estimation in order to produce comparable results. Note that the behavioral outcome variables are not strictly count data, as prior to modeling they are transformed to represent the number of occurrences per 1,000

Table 8. *AIC and BIC Values by Sample and Model Type*

Sample Comparison	AIC			BIC		
	Gaussian	Poisson	Negative Binomial	Gaussian	Poisson	Negative Binomial
Philadelphia Charters	5797	31357	7835	7161	32717	9199
PSM Philadelphia Charters	5692	33233	7280	6956	34492	8544
Pennsylvania Schools	58424	267880	60109	76229	285679	77914
PSM Pennsylvania Schools	8393	51770	9497	10759	54129	11863

Notes. PSM = propensity score matched.

students—values which need not be integers. For the purposes of model comparisons, these values are scaled by a factor of ten and rounded to the nearest integer. Further, note that while Poisson and negative binomial regressions both rely on a logarithmic link function between the predictors and the outcome, this difference in link function is irrelevant to model fit comparisons.

Model fit is examined using the example outcome variable representing annual rate of incidents per school. Fit comparisons are informed by methods outlined by Beaujean and Morgan (2016). To compare model fit, AIC and BIC values are considered, as well as predicted versus residual distributions and influential data points. Model fit is examined for all four control-group contrasts: purposive groups of Philadelphia charter schools and other Pennsylvania schools, as well as propensity score matched groups of Philadelphia charter schools and other Pennsylvania schools. AIC and BIC values of the different model types across these control-group comparisons are provided in Table 8. In

each case, Gaussian regression provides the best fit, returning the lowest AIC and BIC values.

Figures B1 through B8, provided in Appendix B, present graphs of model residuals plotted against predicted values, as well as plots of Cook's D estimates of each observation. Residual versus predicted plots suggest no clear superior model, as values are clustered in each graph and LOWESS lines suggest generally even spread across positive and negative residuals. The Cook's D plots show each observation by a measure of that observation's influence. Thick dashed lines represent $4/n$, a common criterion for identifying observations with high influence. The Gaussian regressions return the fewest observations above this line, suggesting the fewest number of observations with unusually high influence which could bias results.

As the Gaussian models returns equivalent or better estimates of fit than the Poisson or negative binomial models, and as the Gaussian model provides the identity link function and thus most straightforward interpretation, the Gaussian model is considered to fit the data reasonably well and ordinary least squares regression is selected to form the basis for the DD, GDD, and CITS analyses pertinent to RQ1.

Results

The treatment effects estimated by the DD, GDD, and CITS analyses provide complex but informative results. Treatment effects of the DD analyses contrasting treatment schools with the four different control groups—purposive Philadelphia charter schools, propensity score matched Philadelphia charter schools, purposive Pennsylvania schools, and propensity score matched Pennsylvania schools—are provided in Table 9.

Table 9. *RQ1 DD Treatment Estimates, by Sample*

School-Level Outcome Variables	Control Group			
	Purposive Philadelphia Charters	PSM Philadelphia Charters	Purposive Pennsylvania Schools	PSM Pennsylvania Schools
Student Misbehavior				
Incidents	17.55**	67.55*	2.04	9.65
Incidents Involving Law Enforcement	1.37	0.56	-0.26	0.65
Administrative Responses				
Out-of-School Suspensions	22.07	71.39*	-9.67	69.76
Transfers	-0.23	0.07	-0.81	0.32
Expulsions	1.92**	1.10	0.12	0.07
Arrests	1.13*	1.32*	0.58*	0.53
Other Student Outcomes				
Truancy	9.21***	10.86***	8.13***	9.31***
Mathematics Proficiency Rate	-1.58	0.21	-3.06***	-0.82
Reading Proficiency Rate	0.34	2.05	0.08	0.02

Notes. RQ1 = research question (1); DD = difference in differences. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

(Full DD model results including all coefficient estimates are provided in Tables A1, A2, A3, and A4 in Appendix A.) As discussed in Chapter 5, the propensity score matched samples more closely reflect the average baseline characteristics of the treated schools than do the purposive samples; thus the results of the matched samples are considered to reflect latent treatment effects more reliably, and treatment estimates of the purposive samples demonstrate the sensitivity of the effects in comparison.

The results provide some preliminary evidence of treatment effects. The strongest evidence suggests a treatment effect on truancy rate, with treatment increasing the average truancy rate by 9 to 11 percentage points ($p < .01$) as estimated by the PSM models, and by 8 to 9 percentage points ($p < .01$) as estimated by the purposive sampling models. Other evidence is less robust. The table shows a positive effect on incidents, out-of-school suspensions, and arrests across both PSM models, but these do not reach statistical significance against the matched Pennsylvania schools and are only marginally significant ($p < .10$) against the matched Philadelphia charter schools. The latter models suggest treatment caused an additional 67.6 incidents per 1,000 students, an additional 71.4 out-of-school suspensions per 1,000 students, and an additional 1.3 arrests per 1,000 students. These magnitudes vary, however, across the other control group samples. No conclusive evidence is found for treatment effects on transfers, expulsions, or mathematics and reading proficiency rates

Taken at face value, these results suggest that treatment led to increased misbehavior and truancy, but did not affect student performance; administrators

responded with additional out-of-school suspensions, but not transfers or expulsions, though they did resort to arrest slightly more often.

These results only account for school-level outcomes in the years prior to and immediately following the policy change; they do not account for random variance in the baseline outcome measures, nor do they consider longer-term effects of the policy change. The GDD analysis includes two additional pre-treatment years of data to establish aggregate baseline values, as well as a second year of outcomes posttreatment. Results of the GDD analyses are provided in Table 10. (Full GDD model results including all coefficient estimates are provided in Tables A5, A6, A7, and A8 in Appendix A.)

The estimation of treatment effects in the second year posttreatment suggests time-variant treatment effects for many of the outcome variables, and thus an improvement in these models over the DD estimates. Like the DD estimates, the GDD models find a positive, though small, effect on arrests in the first year posttreatment—between 1.0 and 1.3 arrests per 1,000 students ($p < .10$)—but now show that this effect disappears by the second year posttreatment. On the other hand, the evidence here for treatment effects on the rate of incidents and suspensions is now statistically weak, though these estimates remain consistently positive.

The GDD analyses provide some entirely new insights as well. With the inclusion of additional baseline years of data, a positive treatment effect on the rate of incidents involving law enforcement can be seen in the first posttreatment year: across all control samples, the treatment is estimated to increase incidents involving law enforcement by

Table 10. *RQ1 GDD Treatment Estimates, by Sample*

School-Level Outcome Variables	Control Group							
	Philadelphia Charters		PSM Philadelphia Charters		All Pennsylvania Schools		PSM Pennsylvania Schools	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Student Misbehavior								
Incidents	10.84	5.59	43.18	72.52**	14.41***	11.98***	21.67	34.95
Incidents Involving Law Enforcement	2.89***	0.88	3.01***	0.83	3.00***	1.97***	2.81**	-0.34
Administrative Responses								
Out-of-School Suspensions	32.72	31.92	66.93**	38.04	24.47*	10.45	9.52	7.24
Transfers	0.31	-2.21***	1.05	-2.00**	0.39	-2.62***	0.56	-2.64***
Expulsions	0.99*	-0.71	0.77	-0.60	-0.04	-0.49***	0.11	-0.69**
Arrests	1.03*	-0.38	1.30**	-0.11	1.01***	-0.01	0.42	-1.41*
Other Student Outcomes								
Truancy	7.25***	13.60***	6.72**	7.74	4.67***	10.25***	5.23***	11.38***
Mathematics Proficiency Rate	-7.02***	-6.46***	-5.86***	-6.43***	-8.60***	-9.78***	-3.30***	-4.70***
Reading Proficiency Rate	-3.92***	-3.60**	-3.19*	-3.47	-4.77***	-5.72***	-1.72*	-2.22*

Notes. RQ1 = research question (1); GDD = generalized difference in differences. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

2.8 and 3.0 incidents per 1,000 students ($p < .01$). This treatment effect largely disappears in the second posttreatment year. At the same time, the rate of transfers is unaffected in the first year posttreatment, but decreases in the second year by 2 to 2.6 transfers per 1,000 students ($p < .05$). Least ambiguous are the estimated treatment effects on truancy and proficiency rates. Though magnitudes vary somewhat, truancy is estimated to increase year over year posttreatment; mathematics and reading proficiency are estimated to decrease year over year.

The DD and GDD analyses, as their names suggest, model the outcome as the difference in the treatment and control groups' changes in outcome pre- and posttreatment. Both analyses measure the pretreatment outcome as a stationary value, the former as the outcome value in the year immediately prior to treatment, and the latter as the aggregate mean of the outcome in the three years prior to treatment. For a stable pretreatment outcome variable, this approach is sufficient; in the case of an outcome variable that is increasing or decreasing over time for the treatment or control group, however, this approach fails to account for the pretreatment trends in calculating treatment and control groups' changes in outcome pre- and posttreatment. The CITS approach estimates this additional, potentially important, parameter.

Results of the CITS analyses are provided in Table 11. (Full CITS model results including all coefficient estimates are provided in Tables A9, A10, A11, and A12 in Appendix A.) The results conflict somewhat with those of the DD and GDD analyses. The full model results provided in the appendix provide some insight as to why the estimated treatment effects differ from the previous analyses. Across all four control

Table 11. *RQ1 CITS Treatment Estimates, by Sample*

School-Level Outcome Variables	Control Group							
	Philadelphia Charters		PSM Philadelphia Charters		All Pennsylvania Schools		PSM Pennsylvania Schools	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Student Misbehavior								
Incidents	41.73**	52.25	118.25**	186.19**	0.21	-9.31**	-22.07	-31.63
Incidents Involving Law Enforcement	0.66	-2.37	0.54	-2.66	0.24	-2.19	0.85	-3.10
Administrative Responses								
Out-of-School Suspensions	55.74*	67.56	140.27***	150.76**	9.65	-11.95	52.87*	75.12
Transfers	0.42	-2.03	0.61	-2.63**	-0.01	-3.22***	-0.55	-4.30***
Expulsions	1.28	-0.27	0.15	-1.54	-0.08	-0.55	-0.56	-1.72***
Arrests	0.44	-1.24	0.75	-0.89	0.64	-0.57	-0.07	-2.11
Other Student Outcomes								
Truancy	5.84*	11.53***	9.97***	12.75**	3.82***	8.97***	5.05***	11.21***
Mathematics Proficiency Rate	4.59***	10.81***	4.66*	9.00**	-0.80	1.94	1.12	1.52
Reading Proficiency Rate	4.96***	9.62***	5.78***	9.76***	2.12***	4.65***	1.89*	2.90*

Notes. RQ1 = research question (1); CITS = comparative interrupted time series. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

samples, the treatment group is estimated to have exhibited a differential baseline outcome trend than the control group, irrespective of the treatment itself, for four of the outcome variables: the number of incidents per 1,000 students, the number of incidents involving law enforcement per 1,000 students, mathematics proficiency rate, and reading proficiency rate. Accounting for these differential trends results in wholly different treatment estimates for these four outcomes than are estimated by the GDD models. Treatment estimates on these four outcomes are interpreted here; for all other models, the GDD model is preferred as it represents the simpler design:

- (1) The estimated treatment effect on school-level incident rate is now inconclusive: compared with matched Philadelphia charter schools, the treatment effect is large in both posttreatment years, with treatment leading to an additional 118.3 and 186.2 incidents per 1,000 students one and two years posttreatment, respectively ($p < .05$); compared with matched Pennsylvania schools, however, this effect is null, and in the negative direction. The full model results for these analyses show that the control groups' baseline trajectories were different relative to the treatment group: the matched Philadelphia charter schools are estimated to have seen a positively sloped rate of baseline incidents relative to the treatment group, while the matched Pennsylvania schools are estimated to have seen a negatively sloped rate of baseline incidents relative to the treatment group. These different baseline outcome trajectories lead to diverging treatment estimates. This was not the

case for any other outcome variable; it suggests that the PSM control groups themselves are not equivalent across all influences related to school-level incident rates.

- (2) The estimated treatment effect on school-level incidents involving law enforcement is null, though it appears consistently positive in the first posttreatment year and consistently negative in the second posttreatment year across the samples.
- (3) The estimated treatment effect on mathematics proficiency is positive in both years posttreatment, though the estimates are only statistically significant ($p < .05$) when treatment schools are compared with matched Philadelphia charter schools.
- (4) The estimated treatment effect on reading proficiency is positive and statistically significant post years posttreatment across all samples, with estimated effects ranging from an additional 1.9 percentage points proficient ($p < .10$) in the first year posttreatment to an additional 9.8 percentage points proficient ($p < .01$) in the second year posttreatment. The difference in direction of these estimates as compared with those of the GDD models is likely due to the statistically significant differential reading proficiency trends for the treatment group estimated in the CITS model. The treatment group is estimated to have exhibited a preexisting negative baseline trend in proficiency relative to the control group; once accounted for, the treatment is no longer estimated to have negatively affected proficiency; rather, it is

estimated to positively affect proficiency. This outcome is one example where the CITS approach proved essential over the GDD and DD analyses.

Combining the CITS and GDD analyses, the results provide some evidence that treatment led to higher rates of student incidents and out-of-school suspensions. Stronger evidence suggests that treatment led to an increase in truancy; an increase in the rate of arrests in the first posttreatment year, balanced against decreases in the rate of transfers and expulsions in the second posttreatment year; and an increase in the reading proficiency rates both years posttreatment.

Discussion

In the years following the elimination of the zero tolerance policy, the SDP counted fewer incidents of misconduct and fewer student punishments—both school-based and criminal—than it had in the years prior to the policy change, as described in Chapter 5. When compared against multiple control groups, however, the effect of dismantling the discipline policy is less clear cut. The analyses suggest that students skipped school more often, and may have misbehaved at higher rates. Students were initially arrested at higher rates as well. Two years after the policy change, however, this increase in the arrest rate dissipated, and rates of transfers and expulsions declined. Reading proficiency improved. Thus the removal of the policy ultimately did not affect the student behavior and outcomes that are most often tied to the ‘criminalization’ of students in schools—those incidents involving law enforcement and arrests. On the contrary, the elimination of the zero tolerance policy led to an increase in school-level

reading proficiency rates and a decrease in the administrative responses to behavior that reduce students' opportunity to learn—i.e., transfers and expulsions.

The lack of treatment effect on the rate of incidents involving law enforcement and the rate of arrest indicates that dismantling zero tolerance policies did not change the extent to which students encounter law enforcement or the criminal justice system. The criminalization of students is one of the most common arguments against zero tolerance systems; yet this finding suggests that aiming to reduce student exposure to the judicial system by dismantling zero tolerance policies may be a misguided endeavor. In fact, removing the zero tolerance policy appears to increase the rate at which students misbehave and skip school.

Removing the zero tolerance policy did achieve desired effects in a few areas, however, with the reduction in transfers and expulsions, and the increase in reading proficiency. As transfers and expulsions are generally viewed as more damaging punishments long-term than suspensions, this is a positive effect and one in line with arguments in favor of eliminating zero tolerance policies.

Key to these findings was the estimation of the comparative interrupted time series (CITS) model, with its additional parameter estimation of the difference in group trend prior to treatment. This parameter proved statistically significant in the estimation of the treatment effect on the number of incidents per 1,000 students, the number of incidents involving law enforcement per 1,000 students, and mathematics and reading proficiency rates. In each case, the CITS model estimated statistically significant differential group trends in these outcome variables pretreatment, resulting in estimates

that diverged from the previous DD and GDD analyses. The DD and GDD analyses are included in this study as, all else equal, simpler models are preferable over more complex models, and in many cases prove sufficient—such as in the estimation of the effect on transfers and expulsions. Were groups truly equivalent in baseline characteristics, no trend differences would need to be modeled. The significance of the additional trend parameter suggests here that, while propensity score matching approached equivalence of groups in their static baseline characteristics pretreatment, it did not fully capture equivalence of groups in terms of their baseline outcome trends.

CHAPTER 7: SCHOOL POLICE OFFICER STAFFING AND BEHAVIOR

In the 2013-2014 school year, the School District of Philadelphia (SDP) altered police officer staffing positions in 28 of its schools, eliminating a full-time position in each school and replacing it with a half-time position in which a single officer split his or her time across two schools. This selective policy decision provides an opportunity to investigate the effects of reduction in school police officer presence on school-level student and administrative behaviors. This chapter explores the effects of this policy decision, addressing the second research question (RQ2) outlined in the larger study:

RQ2. To what extent did the reduction of police officer presence in particular schools in the SDP affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?

The methods applied in this chapter are outlined in detail in Chapter 4. The chapter includes a discussion of difference-in-difference (DD) analyses comparing the affected schools to other schools in the school district, employing both purposive sampling and the propensity score matched sample described in Chapter 5. The chapter also includes an examination of the strengths and weaknesses of a fuzzy regression discontinuity (FRD) analysis, including intent-to-treat (ITT) and treatment-on-the-treated (TOT) estimates, and the conclusions that can be drawn from such an approach. Combined findings of all analyses are then discussed.

Preliminary Model Selection

Prior to estimating treatment effects, multiple types of linear regressions were explored as potential models for the difference-in-differences analyses applied to RQ2, and model type selected based on model fit. Further, cutoff values were explored for the fuzzy regression discontinuity analysis, as well as linear versus quadratic relationships between predictors and outcome variables, and visual estimation of discontinuity in the outcomes. These preliminary analyses are described here.

DD Model Type Selection

As in Chapter 6, the behavioral outcome variables considered here are not normally distributed; rather, their distributions represent rates of count variables generally clustered at lower values and skewed to the right. Ordinary least squares regression does not require the assumption that the outcome variable is normally distributed, only that the residuals of the given model are normally distributed. Still, an outcome variable with a highly skewed distribution holds the potential to produce residuals with a similarly skewed distribution. For this reason, the fit of the DD models estimated using ordinary least squares regression—based on the normal, or Gaussian, distribution—is compared with linear regressions based in both Poisson and negative binomial distributions, the distributions of which follow more closely that of skewed data and which can be more appropriate for modeling count data. All models are estimated as generalized linear models using maximum likelihood estimation in order to produce comparable results. Note that the behavioral outcome variables are not strictly count data, as prior to modeling they are transformed to represent the number of occurrences per

1,000 students—values which need not be integers. For the purposes of model comparisons, these values are scaled by a factor of ten and rounded to the nearest integer. Note that while Poisson and negative binomial regressions both rely on a logarithmic link function between the predictors and the outcome, this difference in link function is irrelevant to model fit comparisons.

Model fit is examined using the example outcome variable representing annual rate of incidents per school. Fit comparisons follow the comparison methods outlined by Beaujean and Morgan (2016). To compare model fit, AIC and BIC values are considered, as well as residual distributions and influential data points. The models returned the following AIC values: Gaussian regression, 3566.58; Poisson regression, 14712.11; negative binomial regression, 5179.55. Of the three, Gaussian regression returns the lowest AIC value, suggesting the best fit. Further, the models returned the following BIC values: Gaussian regression, 4393.66; Poisson regression, 15535.22; negative binomial regression, 6006.63. Again, Gaussian regression provides the best fit.

Model residuals are plotted against predicted values in Figure B9, provided in Appendix B. Residuals of the all three models are similarly clustered and distributed across positive and negative values, as indicated by the dashed LOWESS lines. Cook's D estimates of the observations are plotted in Figure B10. Thick dashed lines represent $4/n$, a common criterion for identifying observations with high influence. The Gaussian regression returns the fewest observations above this line, while the Poisson returns the most.

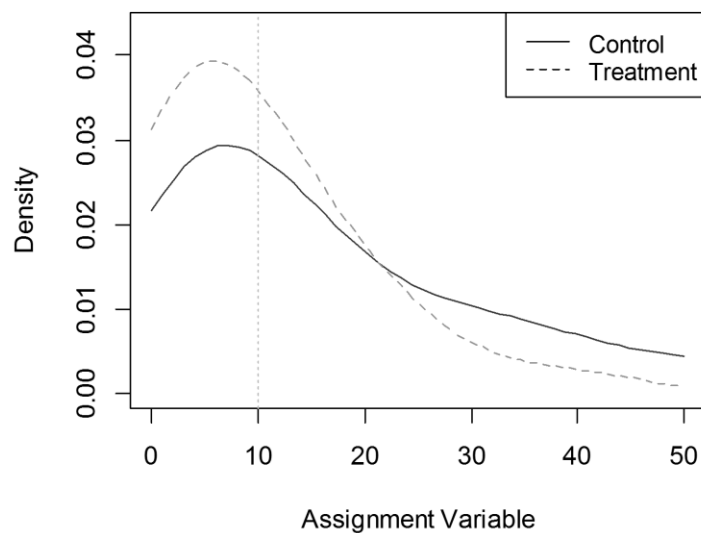
As the Gaussian model returns equivalent or better estimates of fit than the Poisson or negative binomial models, and as the Gaussian model provides the identity link function and thus most straightforward interpretation, the Gaussian model is considered to fit the data reasonably well and ordinary least squares regression is selected to form the basis for the DD analyses pertinent to Research Question (2).

FRD Assignment Variable and Cutoff Value

The decision to reduce police officer presence in 28 SDP schools was likely not based on evaluation of a single variable. Still, this study theorized that schools affected by the policy likely saw fewer behavioral incidents and less perceived need for police presence than non-affected schools. In this case, it might be possible to rank schools according to a given variable and determine a cutoff value which maximizes the proportion of affected—i.e., treated—schools on one side of the cutoff while simultaneously maximizing the proportion of non-affected—i.e., control—schools on the other. Such a cutoff would set the stage for a FRD analysis. Among the school-level measures in this study, the variable representing the school-level number of incidents involving law enforcement in the year prior to the staffing change (the 2012-2013 school year) was determined to provide the largest margin between treatment and control schools, with treatment schools clustering at or below a rate of 10 incidents involving law enforcement per 1,000 students, and control schools clustering above a rate of 10 incidents involving law enforcement per 1,000 students represent. The face validity of this variable holds as the school-level rate of incidents involving law enforcement likely influences the perceived need for police presence in a school.

In a sharp regression discontinuity design, all treatment schools would fall on one side of the cutoff and nearly all control schools on the other. In practice, no treatment assignment cutoff resulted in such a clean divide, as many treatment and control schools share similar values. While a cutoff value of 10 incidents involving law enforcement per 1,000 provides the largest margin, both treatment and control schools span the cutoff: 67 percent of treatment schools fall below this value, but 33 percent fall above it; on the other hand, 55 percent of control schools fall above the value, but 45 percent fall below it. Further, while the majorities of the treatment and control schools fall on opposite sides of the cutoff, the larger sample of control schools overwhelm the treatment schools on both sides of the cutoff: 91 percent of schools above the cutoff are control schools, while still 81 percent of schools below the cutoff—suggesting treatment—are control schools. The density of treatment and control schools are plotted in Figure 14. As the figure shows, the mode of each group falls below the cutoff of 10 incidents involving law enforcement per 1,000 students. The control group, however, is less dense than the treatment group below the cutoff and shows a thicker tail of density above the cutoff, resulting in the majority of control schools falling above the cutoff. Ideally, the mode of the control schools would also fall above the cutoff. Still, no other variable or cutoff value resulted in a wider margin between treatment and control schools.

The inability to separate more cleanly the treatment and control schools, combined with the small sample of treatment schools, likely increases the variance in the outcome below the cutoff and renders the ability to estimate treatment effects imprecise. More formally, the FRD analyses presented in this paper combine traditional regression



Notes. Assignment variable represents the school-level number of incidents involving law enforcement per 1,000 students. A theorized treatment cutoff value of 10 is shown. RQ2 = research question (2).

Figure 14. Density of Treatment and Control Groups by Assignment Variable, RQ2

discontinuity approach with an instrumental variables approach. To apply an instrumental variables approach to an analysis, the independent variable and its instrument must be closely related. In this case, treatment assignment—based on the identified assignment variable—and treatment take-up—based on the schools which actually experienced reductions in police staffing—represent the instrument and the independent variable, respectively. Their relationship, measured as the F statistic associated with regressing one on the other, is significant ($p < .05$) at 4.563 on 1 and 195 degrees of freedom, with a p -value of 0.03. Though statistically significant, the relationship is weak, sharing a correlation r equal to just .15. These initial measures suggest that the assumptions of the

FRD analyses are tenuously, if at all, met. Thus the FRD analyses that follow in this study are *not* intended to be taken as valid estimates of the treatment effect. Rather, they are included as examples of the instability of treatment estimates under tenuous conditions. They should be viewed as exploratory examples of how researchers might examine treatment effects in this context and how sample size and extensive overlap of treatment and control units across the assignment variable affects FRD analyses. They are included in this study, and not discounted all together, as the parallel estimation of DD treatment effects provides a unique opportunity to examine the FRD estimates in comparison with valid treatment estimate counterparts.

FRD Framework, Group Differences, and Models

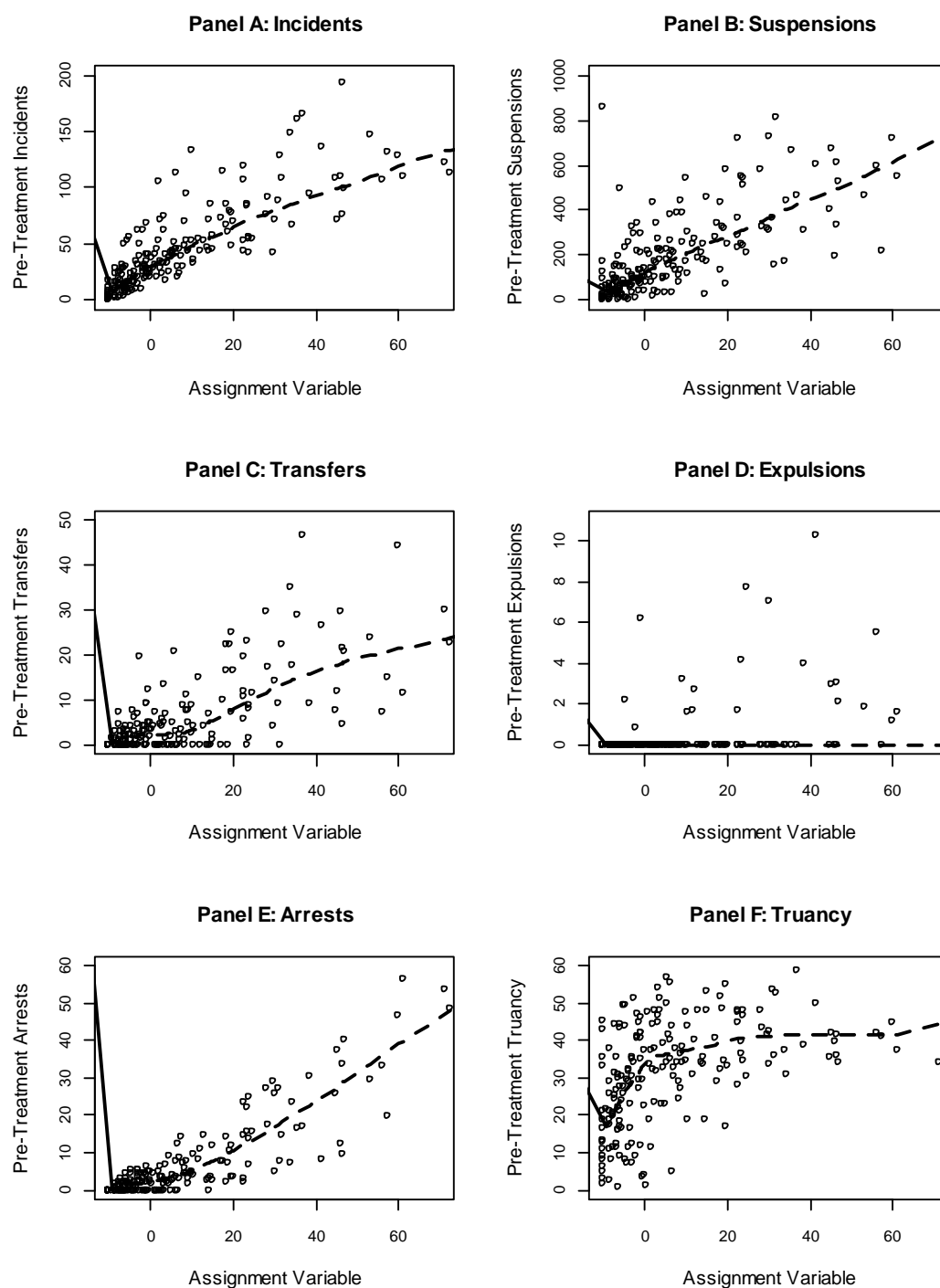
The FRD analyses conducted as part of this study consider a scenario in which a theoretical treatment assignment variable was not strictly applied—that is, some schools assigned to treatment did not receive it, while some not assigned to treatment did. This sorting sets the stage for paired intent-to-treat (ITT) and treatment-on-the-treated (TOT) analyses. The ITT analysis accounts for treatment effects across all units theoretically assigned to treatment—in this case, all schools below the cutoff, whether or not those schools received the treatment—and is measured by a basic regression discontinuity analysis. The TOT analysis represents the treatment effect for those units that did receive treatment, and is a scaled measure of the ITT sample estimate. The TOT estimate is modeled in the FRD framework. Both ITT (regression discontinuity) and TOT (FRD) analyses are conducted in this study.

Table 12. *Mean Differences in RQ2 ITT Samples at Cutoff*

School-level Characteristics	Mean Difference	SE	p-value
% Asian	0.014	(0.025)	0.572
% Black	-0.138	(0.097)	0.161
% Hispanic	0.021	(0.060)	0.729
% White	0.092	(0.058)	0.123
% Other	0.010	(0.010)	0.337
% Free/Reduced-Price Lunch	-0.040	(0.057)	0.490
% Female	0.043	(0.026)	0.097
Pupil/Teacher Ratio	1.575	(1.017)	0.128

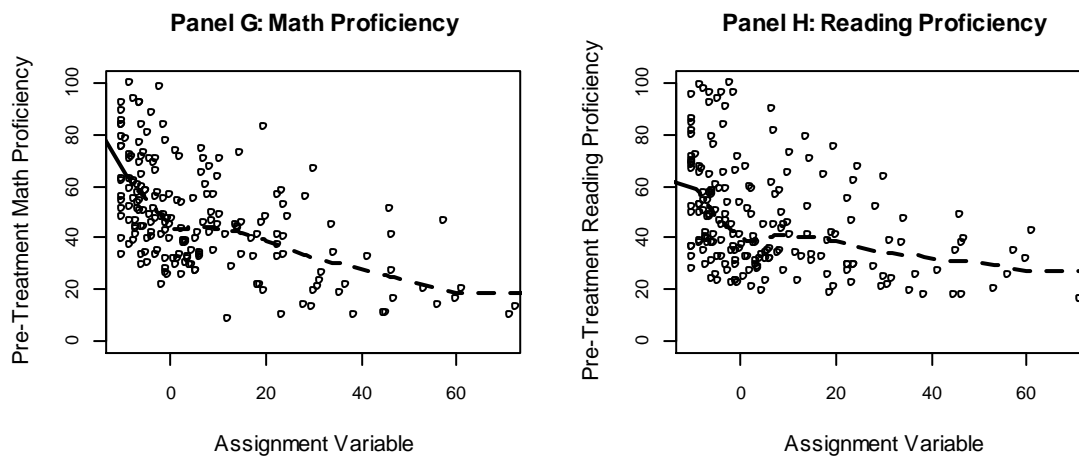
Notes. RQ2 = research question (2). Intent-to-treat (ITT) sample represents treatment and control group assignment based on assignment variable cutoff, not observed receipt of treatment.

Basic regression discontinuity analyses leverage the random sorting of units across a cutoff value used to assign a given treatment, resulting in treatment and control groups which are equivalent in characteristics near the cutoff. This feature of regression discontinuity must be valid for the design to hold. Group equivalence here is examined for the ITT sample, that is, treatment and control groups based on treatment assignment, not treatment take-up. Table 12 presents mean differences in school-level characteristics for these groups within a bandwidth of +/- 2 around the cutoff. The two groups do not differ significantly ($p < .05$) in any characteristic. This group equivalence also suggests that no external effect potentially confounds the treatment effect at the cutoff, as the cutoff is artificially determined and the schools on both sides are statistically equivalent across key characteristics. Thus this assumption of basic regression discontinuity is maintained.



Notes. LOWESS lines plotted to suggest relationship between assignment variable and pre-treatment outcome variables. Outcomes are measured as rate per 1,000 students. RQ2 = research question (2).

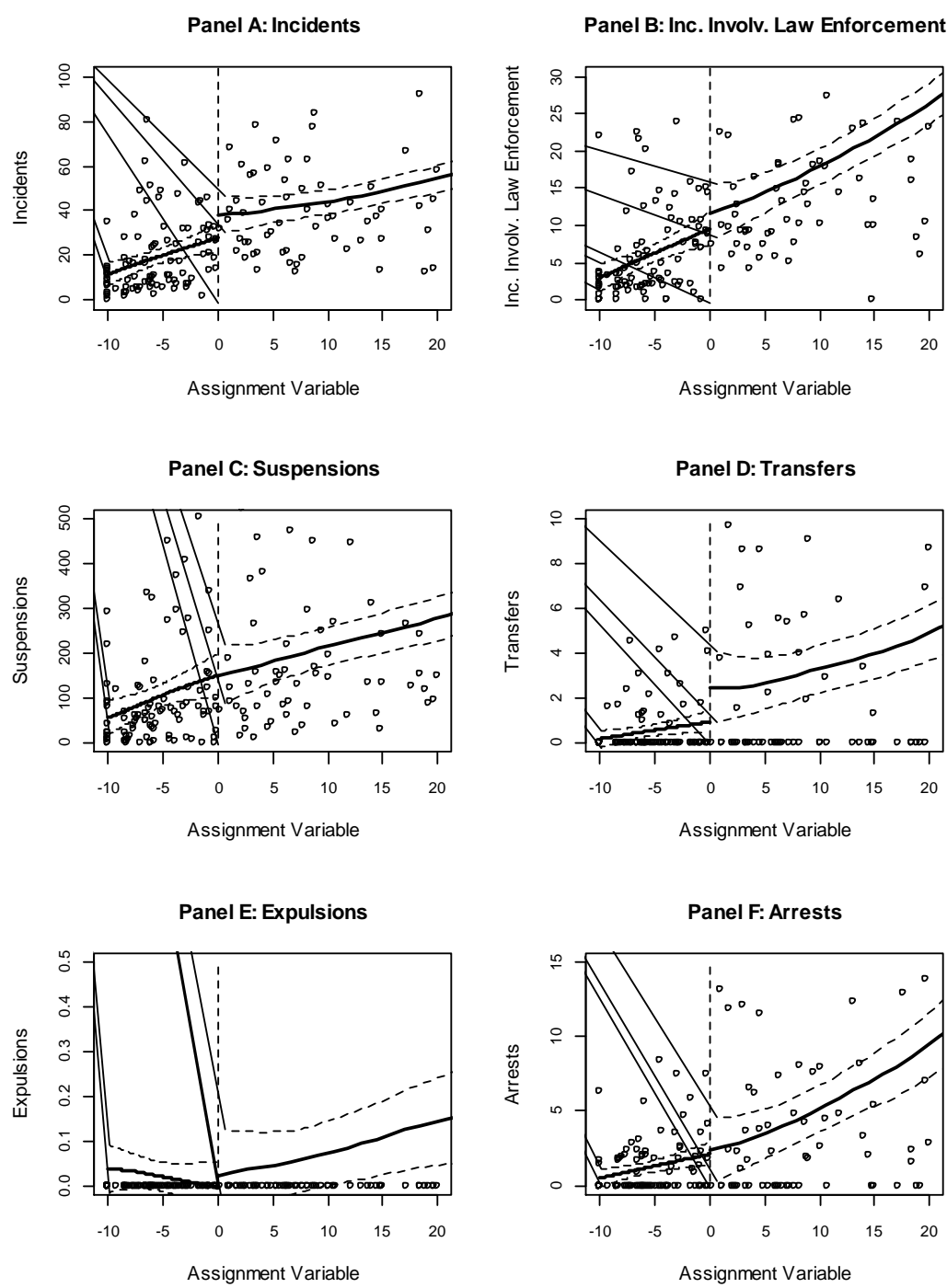
Figure 15. FRD Assignment Variable versus Baseline Outcome Measures, RQ2



Notes. LOWESS lines plotted to suggest relationship between assignment variable and pre-treatment outcome variables. Outcomes are measured as rate per 1,000 students. RQ2 = research question (2).

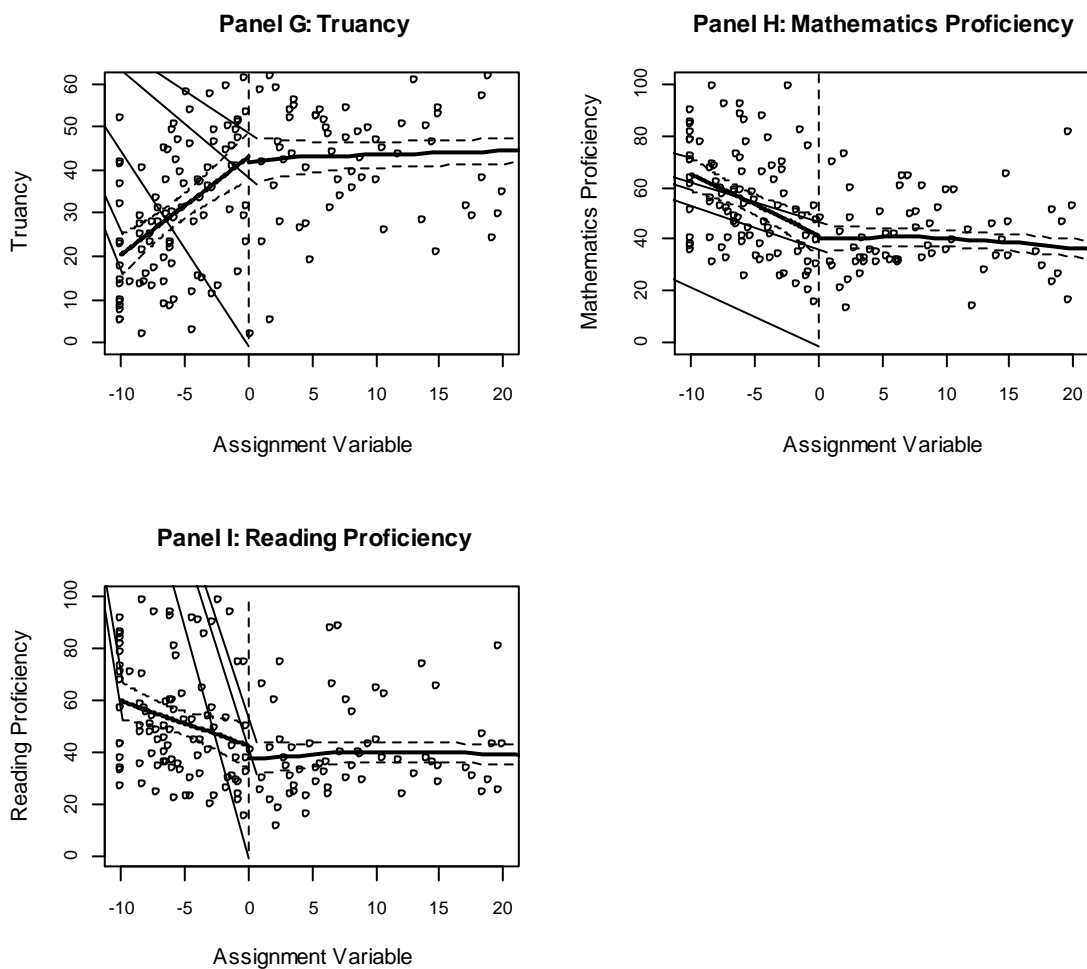
Figure 15, continued

As with the DD analyses, the relationship between the assignment variable predictor and the outcome must be correctly modeled. In Figure 15, baseline outcome measures—those captured in 2012-2013 school year, prior to the staffing change—are plotted against the assignment variable as a proxy for the relationship between the assignment variable and the post-treatment outcomes. The plots suggest that the assignment variable shares linear relationships with outcomes representing incidents, suspensions, transfers, and arrests, and potentially quadratic relationships with outcomes of representing truancy, mathematics proficiency, and reading proficiency. Note that almost no relationship exists between the assignment variable and expulsions, as schools saw almost no expulsions; in addition, the relationship between the assignment variable and the outcome representing number of incidents involving law enforcement is not



Notes. Nonparametric regression lines plotted to demonstrate discontinuity in outcome variables at centered assignment variable cutoff. Outcomes measured as rate per 1,000 students. RQ2 = research question (2).

Figure 16. Treatment Discontinuity at the Assignment Cutoff, RQ2



Notes. Nonparametric regression lines plotted to demonstrate discontinuity in outcome variables at centered assignment variable cutoff. Outcomes measured as rate per 1,000 students. RQ2 = research question (2).

Figure 16, continued

plotted but assumed to be linear, as the assignment variable is this same measure pre-treatment.

Building on these descriptive patterns, linear and quadratic regression fits were compared for each outcome by calculating F statistics on the significance of the inclusion

of the quadratic term. In no case was the quadratic model a significantly better fit than the linear model ($p < .05$), and thus all outcomes are modeled as sharing a linear relationship with the assignment variable.

Finally, prior to conducting the ITT and TOT analyses, the discontinuity in the outcome is plotted for each posttreatment outcome using nonparametric regression. These plots are shown in Figure 16. Though formal inferential tests follow, the plots suggest minimal treatment effects, as few plots indicate a discontinuity in the outcome at the cutoff value. These plots graphically display the limitations of the FRD analysis in which the true treatment schools comprise a small proportion of the assigned treatment group. Because the treatment effect is measured across all schools below the cutoff, the ITT effect is predominantly measured on non-treatment schools, and is inherently small; the TOT effect then scales the ITT effect based on treatment probability below the cutoff, which is also small. The result is a ballooning but unstable effect, as will be seen in the final treatment effect estimates.

Results

Treatment estimates of the DD analyses and FRD analyses are provided in Table 13. Full results of these models are provided in Tables A13, A14, A15, and A16 in Appendix A. Overall, the estimates provide weak evidence of treatment effects across the school-level outcome variables. In the left-most column, estimates are drawn from DD models using all non-affected SDP schools as the purposive control comparison group. These models provide the greatest number of significant treatment effects across the outcome variables: treatment is estimated to have reduced the number of incidents

Table 13. *RQ2 Treatment Estimates, by Analysis Type and Sample*

School-Level Outcome Variables	DD Analysis		FRD Analysis	
	Purposive Sample	Matched Sample	ITT	TOT
Student Misbehavior				
Incidents	-0.33	11.13	-3.49	-37.88
Incidents Involving Law Enforcement	-4.10**	-3.25	1.44	15.62
Administrative Responses				
Out-of-School Suspensions	4.53	8.10	18.49	200.75
Transfers	0.53	1.34	-0.56	-6.07
Expulsions	-0.26	0.38*	0.04	0.47
Arrests	-2.05**	-1.06	1.40	15.18
Other Student Outcomes				
Truancy	2.38*	2.54	2.48	26.97
Mathematics Proficiency Rate	-2.60**	-0.57	-0.32	-3.52
Reading Proficiency Rate	-2.25**	-1.33	0.56	6.07

Notes. RQ2 = research question (2); DD = difference in differences; FRD = fuzzy regression discontinuity; ITT = intent to treat; TOT = treatment on the treated. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

involving law enforcement by 4.1 per 1,000 students ($p < .05$) relative to the control group, as well as the number of arrests by 2.1 per 1,000 students ($p < .05$). Note that these estimates counter the raw group trends depicted in Figure 8, which do not account for school-level fixed effects or time varying covariate measures. At the same time, truancy is estimated to have increased by 2.4 percentage points ($p < .10$) relative to the control group, while mathematics and reading proficiency decreased by 2.6 and 2.3 percentage points ($p < .05$), respectively.

These estimates are potentially biased given the lack of baseline equivalence in the treatment and control groups, described previously and presented in Table 7. While the DD analyses control for time-invariant differences among groups, they do not partial

out group differences that may change over time. If, for example, proficiency rates were changing irrespective of the treatment at different rates in the two groups across the years examined here, the treatment effect estimate may falsely attribute this difference to the treatment. What the significant results here show, however, go in the *opposite* direction of what might be expected: schools that saw their police staff reduced were those with fewer minority students and fewer student incidents and punishments at baseline; yet they saw lower proficiency rates and more truancy in the post-treatment year. The direction of these findings suggests that they are not artefacts of selection intervention effects, but rather provides support that they are attributable to the treatment.

In the second-most left column, the treatment estimates of the propensity score matched DD analysis are provided. No treatment effects are significant here at $p < .05$. Still, the estimates remain in the same direction as all significant effects from the purposive sample, i.e., the treatment is still estimated to have decreased incidents involving law enforcement and arrests relative to the control group, while increasing truancy and reducing mathematics and reading proficiency rates. The propensity score matched sample provides greater equivalence on baseline characteristics than the purposive sample, and thus should provide more reliable estimates. That these estimates are no longer statistically significant suggests either that the true treatment effects are null, or that the sample is underpowered to detect the effects.

The sample size for the PSM analyses is indeed smaller than that for the purposive group analyses, as shown in Tables A14 and A13, respectively. Calculating the minimum detectable effect sizes based on the observed PSM sample confirms that the

estimates in Table 13 are near or below the minimum detectable effect sizes for outcome variables of incidents involving law enforcement, arrests, truancy, mathematics proficiency rate, and reading proficiency rate. That is, the effects of the treatment, if the PSM estimates are reliable, are too small to be detected by the sample available. Given that the PSM estimates all fall in the same direction as the estimates of the purposive analysis, and that the PSM estimates, if true, are too small for their statistical significance to be detected by the sample of the PSM analyses, the study concludes that the directions of these treatment estimates are likely accurate, though the magnitudes may vary from the estimates in Table 13.

Results of the regression discontinuity analysis are provided in the two right-most columns of Table 13. First, the results of the ITT analyses—based on traditional regression discontinuity analysis—vary in direction and magnitude as compared to the estimates of the difference-in-differences analyses, and none are estimated to be statistically significant. Given the problems inherent in the treatment assignment cutoff described earlier in this chapter, the disagreement between these estimates and those of the difference-in-differences analyses and their lack of statistical significance is not surprising. Of interest is the extent to which these biased ITT estimates lead to instability in the TOT estimates. The TOT estimates—produced by the fuzzy regression discontinuity analysis—are large in magnitude as compared to all other estimates. For example, the treatment effect on the rate of suspensions per 1,000 students is 4.53 and 8.1 as estimated by the purposive and propensity-score-matched difference-in-differences analyses; this increases to 18.49 as estimated by the ITT analyses, and balloons to 200.75

as estimated by the TOT analysis. More importantly, the standard errors on the TOT estimates are even larger in comparison. Shown in Table A16, the standard errors of the treatment estimates are anywhere from 10 to 100 times larger than their respective estimates, and as much as 1000 times larger than the standard errors associated with the difference-in-differences and ITT analyses shown in Table A13, Table A14, and Table A15. The large ratio of control schools to treatment schools decreases the signal to noise, resulting in an unstable ITT effect, which is then magnified by this ratio, and thus rendered unstable. The standard errors of the TOT effects, produced by bootstrapped iterations of the model, demonstrate this instability.

Discussion

Past research on the effect of police presence in schools has found mixed evidence: some research suggests police presence increases, or is at least correlated with, outcomes such as student crime, suspensions, or expulsions. Other research finds null effects of police presence. This study contributes to this body of literature by examining the reduction of police presence and its effects on a number of student outcomes through the use of quasi-experimental methods.

The results suggest null to modest effects of a reduction in school police presence. The PSM analyses produce treatment effect estimates of small magnitude and no statistical significance. Still, the small sample included in the PSM analysis reduces the ability of that analysis to detect small effect sizes with certainty; thus the size of the estimates fall below the threshold for detecting statistical significance, though these effects may actually exist—and in a district as large the SDP, small effects could have a

meaningful impact. Reducing the arrests rate by 1 arrest per 1,000 students districtwide, for example, would theoretically equate to about 130 fewer arrests per year. Combined with the results of the DD analysis using a purposive sample, the analyses suggest potential causal relationships between reduction of police presence and fewer incidents involving law enforcement and arrests. No effects are found on student suspensions or transfers; a marginally significant, positive effect ($p < .10$) is found for student expulsions, but the magnitude of the effect is small.

The reduction in incidents involving law enforcement, paired with the reduction in arrests, suggest both that students committed fewer acts of the most severe behavioral misconduct and that they faced the most severe punishment—arrest—less often. Thus reducing police presence in schools reduced two forms of negative student interactions with law enforcement. This could support the argument that police presence in schools leads to increased criminalization and incarceration of students, at least on a very small scale; still, the evidence for these effects is weak.

On the other hand, reduced police presence may have led to more truancy and lower math and reading test scores. The causal pathway from police presence to these outcomes is beyond the scope of this study; that these negative outcomes are associated with less police presence, however, suggests that police presence may bring some stability or authority to school institutions that goes beyond simply affecting student misconduct in schools.

As the analyses only examine outcomes in the year after the policy change, however, the study cannot determine if these relationships remain over time. Student and

administrator behaviors may return to their original state after the immediate impact of the staffing change wanes, or they may gradually follow new patterns. Taken together, the results of this study paint a nuanced picture of small shifts in behavior in the year immediately following a reduction of police presence in schools. Future analyses should examine additional years after the policy treatment, ideally in a location with a larger available sample of affected schools to adequately power detection of treatment effects.

CHAPTER 8: CONCLUSION

While zero tolerance discipline policies and law enforcement in schools are two policy areas often linked to severe punishment of students and the unwarranted introduction of students to the criminal justice system, rigorous research supporting these causal links is scarce. Current literature on the topic of school discipline, law enforcement, and student outcomes often relies on correlational analyses or single-case non-experimental designs. This study contributes to this area of research by applying quasi-experimental design methods to analyses examining student behaviors and administrative responses and the extent to which they are affected as district policies change. The analyses presented in this paper provide evidence on issues of school discipline policy in a large, urban school district by contrasting the School District of Philadelphia (SDP) with multiple comparison groups and addressing the following research questions:

- (1) To what extent did the dismantling of the SDP's zero tolerance discipline policy affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?; and
- (2) To what extent did the reduction of police officer presence in particular schools in the SDP affect severe student misconduct and consequential administrative reactions? To what extent did it affect schools' academic performance and truancy rates?

The study concludes that removing zero tolerance policies and limiting police presence in schools can impact student behavior and administrative responses, but not necessarily in the directions argued by critics of such policies. Dismantling the zero tolerance policy did lead to a lagged decrease in transfers and expulsions, but it also increased student truancy and likely increased student misconduct—measured as the average school-level rate of student incidents—without affecting the rate of student incidents involving law enforcement or the rate of student arrest. Limiting the presence of police officers may have led to declines in the rates of incidents involving law enforcement and arrests, but the evidence here is weak, and it is paired with weak evidence of increases in truancy and decreases in proficiency.

The quasi-experimental methods applied here combine design, sampling, and analysis methods to produce multiple estimates of the treatment across multiple contrasts and models. Studies typically include only a single such approach. As this study indicates, the use of a single methodological approach—particularly one that does not accurately represent the relationships in the data—ignores the extent to which estimates are sensitive to different contrasts and models, even when those contrasts and models are thoughtfully crafted to approximate baseline equivalence and eliminate confounding effects. The evidence provided in this paper is at times weak as a result, as estimates of the same effect vary in magnitude and statistical significance across seemingly robust comparisons. Still, this paper errs on the conservative side of conclusions, rather than finding effects where support for such effects is inconsistent. The paper does, however, select for interpretation the least complex model for each outcome based on the statistical

significance of added complexity in predictors. The multiple approaches thus serve two purposes: to highlight the value in exploring sensitivity of effects across variations of a treatment-control contrast, and to identify the simplest but best fitting models from which to identify treatment effects.

Limitations

The study faces some limitations and some areas where further extension of the analysis would prove valuable. The foremost limitation of the study is the bounds placed on treatment estimation and the generalizability of treatment effects by the sample of treated schools. The School District of Philadelphia is a demographically unique district in Pennsylvania. Even with propensity score matching, drawing a truly equivalent control sample of schools from the rest of Pennsylvania may not be possible. Though the propensity score matching analysis did result in more similar control samples, these samples themselves demonstrated different baseline outcome trajectories, as described in the results of the CITS analysis. Further, the propensity score matched samples were small, where one-to-many matching could have otherwise resulted in large control groups where plentiful similar schools available for matching. The analyses using the propensity score matched samples were thus limited in their power to identify statistically significant treatment effects.

At the same time, the uniqueness of Philadelphia raises questions as to the generalizability of the findings to other locales. The findings are generalizable to large, low income, urban districts with a history of strict discipline policies, and thus may not apply to suburban or rural school districts facing the same disciplinary decisions. Future

replications of this study across different types of districts would expand and round out the findings of this study, as would comparisons of Philadelphia against other large, urban districts.

A second limitation of the study is its inability to examine moderating or mediating variables. The data publicly available for the study precluded such analysis in a concrete manner, as potential mediating and moderating variables—such as school climate, school leadership, school resources, etc.—are not published universally across Pennsylvania and thus could not be included in the analyses. Further, the nature of the analyses precluded the examination of time-invariant school characteristics as moderating variables, as these variables would prove redundant with the school-level fixed effects in the DD, GDD, and CITS analyses. Ideally, individual-level data could be used to examine moderating effects in place of the school-level effects; these would suggest whether certain types of students were affected more or less by the treatment. Without the estimation of moderating and mediating effects, this paper cannot provide insights as to the extent to which treatment effects are differential across different types of schools and students, or as to the causal pathways of the estimated treatment effects. This area of research would likely represent a worthwhile extension of this study.

A third limitation of the study, discussed in Chapter 4, is the reactionary nature of the research questions. Both research questions examine the removal of a certain policy, not that policy's initial introduction. Thus the paper cannot comment on the extent to which the introduction of zero tolerance policies or police in schools change student and administrative behaviors from a previous baseline; the paper can only provide evidence

on the extent to which behaviors change in the aftermath of their removal—and in the case of the second research question, this removal is only partial, as schools retained police officer presence for half of each week. This paper argues that these questions are still important; schools across the country are currently facing these same decisions with almost no hard evidence on their likely effects, no matter how minimal the decisions may seem. Still, understanding how these policies affected behavior in the first place is a key element in assessing their value and crafting long-term discipline policy.

A final limitation of this study is the inability to extend posttreatment analyses beyond two posttreatment years. As the GDD and CITS analyses suggest, the treatment effects were not stable across the two treatment years included in the study; it is likely that effects continued to change three and four years posttreatment. Estimating such long-term effects, particularly given a district-level treatment, may prove difficult outside of a well-designed experimental study which precludes other policy decisions that may interfere with long-term effects. Still, given that much of the criticism of zero tolerance policies and law enforcement in schools is centered on life outcomes of students, an examination of long-term effects is important to evaluating the full impact of such policy decisions.

Future Work

The findings and limitations of this study suggest a number of areas for future work. As described above, an important next step to complement this paper would be to extend it to other locales which made similar decisions to move away from zero tolerance and pervasive police presence in schools. A number of other districts across the

country—including Chicago, New York City, Los Angeles, San Francisco, Oakland, Miami, and Denver (Anderson, 2015; Hood, 2012)—have taken similar steps in recent years; an examination of one or more of these districts, with sufficient sample size of both treatment and control groups, could provide useful information on the generalizability of this study and potentially aid in identifying small treatment effects.

Extending the analysis to include additional years of data would also provide a useful expansion of this work. This is likely not feasible in Philadelphia, where additional district-level decisions affected school discipline policy in the 2014-2015 school year, but may be feasible in other locations.

Finally, examining moderating and mediating values is necessary to understanding why treatment effects manifest as they do, and how they may differ for different types of schools and students. This paper began with a discussion of the theories of behavior which influence thinking on why students misbehave and how best to address and curb such behavior. Examining moderating and mediating values could provide key evidence to suggest that certain theories are more applicable to students than others. Ultimately, this knowledge could lead to better informed, better crafted school discipline policies.

APPENDIX A: ADDITIONAL TABLES

Table A1. DD Regression Estimates, Purposive Phil. Charter School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	79.75 (300.08)	-137.39 (99.27)	-586.7 (1073.71)	-92.5 (73.79)	-9.42 (22.75)	-83.17* (50.63)	165.51** (67.08)	-11.26 (48.76)	-60.39 (48.07)
Treatment	-293.78 (215.26)	31.7 (40.86)	-118.18 (431.22)	70.12** (30.39)	8.65 (8.38)	8.27 (23.92)	9.3 (32.79)	-5.71 (34.78)	7.83 (25.05)
Post	-18.5*** (6.66)	-1.02 (1.06)	-23.83 (23.45)	0.08 (0.60)	-1.98*** (0.54)	0.09 (0.61)	-1.05 (2.55)	-2.9** (1.17)	-0.46 (1.17)
Treatment*Post	17.55** (7.53)	1.37 (1.30)	22.07 (32.76)	-0.23 (0.87)	1.92** (0.76)	1.13* (0.68)	9.21*** (2.49)	-1.58 (1.35)	0.34 (1.51)
Asian %	366.51 (256.31)	183.95 (113.28)	1549.51 (1050.96)	11.75 (75.46)	-6.57 (19.83)	75.73 (51.92)	-105.19 (66.81)	-45.66 (49.22)	65.88 (59.43)
Black %	187.55 (202.49)	122.39 (94.39)	1211.41 (1015.17)	54.23 (58.12)	-6.54 (22.52)	79.92* (43.36)	-93.29* (52.19)	38.1 (35.12)	87.82** (38.02)
Hispanic %	162.95 (272.48)	109.54 (103.65)	1009.02 (1129.97)	117.21 (111.24)	11.51 (20.07)	55.2 (48.76)	-173.44*** (62.71)	-16.78 (45.56)	96.1* (54.49)
White %	239.36 (224.04)	136.64 (96.81)	1130.68 (1070.29)	61.54 (57.12)	2.74 (25.65)	100.33* (54.18)	-63.17 (53.80)	57.22 (39.18)	126.66*** (44.87)
Female %	177.84 (127.67)	37.25 (35.87)	23.52 (770.52)	4.84 (38.16)	-1.17 (21.43)	15.6 (24.76)	-11.81 (32.58)	15.74 (17.54)	-1.15 (20.16)

Table A1, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	-10.48 (13.25)	2.3 (4.56)	7.58 (36.55)	-4.48 (3.23)	0.78 (1.37)	2.31 (1.90)	-5.48 (7.11)	2.84 (2.42)	1.55 (3.01)
Pupil/Teacher Ratio	-0.57 (1.10)	-0.3 (0.34)	-5.5 (6.28)	-0.17 (0.25)	-0.04 (0.11)	-0.23 (0.22)	-0.2 (0.20)	-0.03 (0.17)	-0.46** (0.19)
Primary School	119.19 (114.75)	3.96 (20.76)	-93.19 (195.09)	-3.33 (19.73)	5.02 (5.27)	0.78 (10.37)	-20.38 (21.85)	-1.38 (16.33)	-1.59 (11.17)
Middle School	144.8 (250.69)	117.35* (62.63)	386.52 (577.28)	34.91 (34.84)	10.57 (12.02)	73.29* (38.58)	7.59 (39.87)	52.37 (34.18)	-5.74 (23.14)
High School	-73.31 (176.67)	1.53 (36.19)	-350.03 (315.92)	2.01 (25.73)	0.02 (7.84)	4.13 (21.77)	-6.58 (26.81)	36.19 (28.07)	13.89 (17.52)
Enrollment	-0.03 (0.04)	0 (0.01)	0.1 (0.24)	0 (0.01)	0.01 (0.01)	0.01 (0.01)	0 (0.01)	0 (0.01)	0 (0.01)
Adj. R^2	0.81	0.81	0.70	0.51	0.24	0.84	0.80	0.95	0.93
Observations	594	594	594	594	594	594	594	594	594
Schools	297	297	297	297	297	297	297	297	297

Notes. DD = difference in differences; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A2. DD Regression Estimates, PSM Phil. Charter School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	-805.03 (505.82)	-130.66 (96.22)	-966.97 (1017.40)	-47.58 (66.20)	-1.24 (23.05)	-91.66** (46.06)	220.44*** (79.55)	-13.24 (50.80)	-40.54 (42.34)
Treatment	126.13 (80.80)	38.04 (23.42)	178.97 (234.70)	27.94** (13.69)	2.05 (3.92)	16.07* (8.56)	11.36 (22.34)	-34.78 (21.77)	-20.43 (18.03)
Post	-63.59** (30.64)	-0.41 (1.24)	-70.93** (32.05)	-0.23 (0.79)	-1.21* (0.66)	-0.12 (0.75)	-3.17 (2.97)	-4.24* (2.17)	-1.99 (1.56)
Treatment*Post	67.55* (35.18)	0.56 (1.34)	71.39* (42.20)	0.07 (0.98)	1.1 (0.86)	1.32* (0.73)	10.86*** (2.84)	0.21 (2.51)	2.05 (1.87)
Asian %	737.64 (477.27)	183.51 (116.46)	1783.3 (1153.95)	11.1 (78.80)	-10.95 (20.49)	82.58 (50.51)	-128.05* (70.80)	-2.81 (55.22)	63.66 (63.41)
Black %	493.15 (429.35)	109.26 (100.47)	1350.07 (1073.77)	54.27 (60.98)	-8.99 (24.80)	80.07* (45.37)	-123.53** (57.43)	68.17 (44.64)	95.21** (37.81)
Hispanic %	694.12 (572.41)	104.09 (112.46)	926.69 (1210.04)	127.17 (114.29)	7.47 (21.76)	60.36 (51.04)	-208.35*** (62.40)	23 (57.36)	84.76 (60.21)
White %	412.44 (359.36)	138.41 (100.61)	1397.82 (1207.18)	66.26 (60.38)	0.74 (28.85)	107.56** (54.83)	-82.86 (57.17)	79.16* (45.32)	121.36*** (45.84)
Female %	419.13* (238.73)	31.04 (38.60)	288.54 (847.05)	1.64 (40.53)	-4.67 (22.24)	11.58 (25.28)	-19.34 (36.45)	26.21 (21.01)	5.83 (21.23)

Table A2, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	-21.22 (20.81)	2.99 (4.90)	16.07 (39.53)	-3.53 (3.37)	1 (1.38)	2.61 (1.98)	-1.16 (4.90)	0.81 (2.38)	0.43 (3.38)
Pupil/Teacher Ratio	-2.71 (1.86)	-0.22 (0.36)	-6.47 (7.16)	-0.21 (0.27)	-0.04 (0.13)	-0.21 (0.22)	-0.13 (0.21)	-0.26 (0.19)	-0.52** (0.21)
Primary School	165.9 (116.48)	4.68 (18.50)	-218.33 (203.07)	-5.9 (14.12)	8.76* (4.99)	1.37 (9.10)	-47.03 (38.80)	-1.51 (10.46)	-1.22 (10.44)
Middle School	554.17** (246.74)	117.42* (64.83)	900.53 (748.47)	-17.78 (38.77)	8.58 (11.90)	75.11** (36.96)	-11.54 (56.05)	18.1 (26.00)	-20.11 (20.56)
Enrollment	0.05 (0.08)	0 (0.01)	0.02 (0.32)	0.01 (0.01)	0 (0.01)	0.01 (0.01)	-0.01 (0.01)	0 (0.01)	0 (0.01)
Adj. R^2	0.79	0.81	0.69	0.50	0.12	0.84	0.83	0.95	0.93
Observations	550	550	550	550	550	550	550	550	550
Schools	275	275	275	275	275	275	275	275	275

Notes. DD = difference in differences; PSM = propensity score matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A3. DD Regression Estimates, Purposive Pennsylvania School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	108.76** (50.24)	39.39* (22.84)	295.93* (161.47)	12.46* (7.07)	8.85*** (2.67)	33.2** (14.68)	11.72** (5.80)	67.43*** (10.10)	58.76*** (11.58)
Treatment	-102.42** (43.83)	6.47 (4.31)	-3.04 (49.51)	4.77* (2.69)	0.16 (0.83)	3.27 (2.71)	8.93*** (3.28)	15.63 (10.21)	19.49 (12.99)
Post	-5.67*** (0.96)	-0.4*** (0.13)	-3.77*** (1.23)	-0.15** (0.07)	-0.15*** (0.04)	-0.22*** (0.08)	0.45*** (0.11)	-1.91*** (0.10)	-1.42*** (0.11)
Treatment*Post	2.04 (2.10)	-0.26 (0.68)	-9.67 (10.34)	-0.81 (0.55)	0.12 (0.22)	0.58* (0.35)	8.13*** (0.38)	-3.06*** (0.31)	0.08 (0.38)
Asian %	475.66* (244.29)	16.28 (13.81)	-48.7 (226.12)	2.23 (8.43)	1.09 (4.15)	7.99 (7.71)	-9.4 (10.78)	-11.94 (12.53)	22.4 (14.82)
Black %	71.4 (105.34)	12.59 (9.01)	291.08** (144.49)	-12.68 (8.84)	3.88 (3.32)	7.1 (4.95)	0.4 (9.90)	-10.5 (8.68)	-19.91** (9.81)
Hispanic %	131.47** (60.44)	25.15** (10.87)	307.26** (135.15)	29.67*** (10.65)	-4.5 (3.16)	7.63 (5.71)	-16.64* (9.28)	-49.54*** (9.52)	-30.09*** (9.27)
White %	-43.42 (40.97)	16.44*** (5.81)	49.98 (64.56)	3.13 (3.71)	2.59 (1.78)	2.26 (3.70)	3.37 (5.01)	12.4** (6.22)	13.17** (5.84)
Female %	53.49 (67.47)	-19.86* (10.74)	-166.01* (99.57)	-5.23 (4.71)	-0.93 (2.40)	-8.72 (6.66)	-9.85* (5.28)	-0.6 (4.10)	-2.12 (4.35)

Table A3, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	6.49 (11.70)	0.29 (1.48)	1.47 (13.99)	-0.6 (0.89)	0.45 (0.38)	0.16 (0.65)	0.7 (1.14)	-0.69 (0.92)	-1.08 (1.28)
Pupil/Teacher Ratio	0.83 (0.55)	-0.27** (0.11)	1.22 (0.83)	-0.03 (0.02)	0.05 (0.03)	-0.04 (0.03)	0.01 (0.03)	-0.16*** (0.04)	-0.11** (0.05)
Primary School	5.09 (6.24)	2.8 (2.75)	41.53** (17.25)	0.99 (0.77)	0.23 (0.82)	-1 (1.85)	3.34* (1.75)	-1.7 (1.06)	1.44 (1.12)
Middle School	4.49 (11.03)	5.99 (5.22)	36.95 (26.86)	1.76 (1.18)	1.86** (0.87)	-1.09 (3.71)	2.15 (1.53)	0.57 (1.34)	-1.88* (1.02)
High School	-94.02** (38.16)	-35.27 (22.58)	-309.24** (153.71)	-13.21** (6.49)	-12.34*** (2.27)	-28.87** (14.45)	-6.96* (3.75)	-4.9*** (1.34)	12.53*** (2.88)
Enrollment	0.02 (0.02)	0 (0.00)	0.03 (0.04)	0 (0.00)	0*** (0.00)	0 (0.00)	0 (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Adj. R^2	0.75	0.75	0.83	0.43	0.41	0.73	0.82	0.93	0.91
Observations	5376	5376	5376	5376	5376	5376	5373	5376	5376
Schools	2688	2688	2688	2688	2688	2688	2686.5	2688	2688

Notes. DD = difference in differences; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A4. DD Regression Estimates, PSM Pennsylvania School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	-97.47 (580.77)	-42.47 (75.18)	34.05 (948.37)	-47.61 (55.76)	-4.37 (14.45)	-48.51 (31.96)	6.43 (63.09)	8.53 (36.06)	15.12 (35.08)
Treatment	158.2 (245.00)	98.49** (44.68)	378.48 (486.94)	39.33 (27.40)	10.92 (8.76)	66.67** (28.56)	18.1 (28.12)	0.42 (20.72)	3.71 (17.06)
Post	-16.3 (18.65)	-1.09 (1.04)	-73.68* (42.44)	-0.38 (0.90)	-0.07 (0.24)	0.36 (0.56)	-0.23 (1.43)	-3.86*** (0.88)	-0.71 (0.81)
Treatment*Post	9.65 (20.37)	0.65 (1.23)	69.76 (45.70)	0.32 (1.06)	0.07 (0.48)	0.53 (0.62)	9.31*** (1.35)	-0.82 (0.97)	0.02 (1.01)
Asian %	1950.42 (1754.63)	60.14 (74.22)	-351.13 (1074.63)	30.86 (46.23)	-5.54 (11.32)	37.28 (35.45)	-19.44 (54.60)	7.01 (39.64)	55.76 (46.29)
Black %	2.14 (494.20)	39.94 (70.48)	690.49 (920.65)	70.52 (48.27)	1.04 (16.93)	47.64 (29.96)	0.99 (54.04)	33.91 (32.75)	41.08 (28.98)
Hispanic %	208.06 (522.99)	13.01 (88.84)	-55.62 (1115.89)	131.74 (91.03)	11.76 (17.20)	35.66 (33.44)	-103.68* (56.77)	-0.68 (42.09)	35.59 (45.74)
White %	25.54 (505.03)	25.09 (72.73)	-992.26 (1072.23)	56.23 (40.91)	3.25 (17.98)	47.63 (37.12)	4.94 (57.33)	58.04* (34.23)	90.33*** (34.70)
Female %	392.97 (253.76)	11.84 (28.41)	-20.35 (669.33)	4.93 (31.36)	-0.12 (15.74)	4.75 (20.13)	4.14 (24.64)	0.15 (17.69)	-14.87 (21.10)

Table A4, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	33.31 (39.41)	3.19 (4.43)	-8.76 (42.42)	-1.79 (2.81)	0.97 (1.27)	2.53 (2.01)	-0.64 (3.17)	-0.17 (2.08)	-1.21 (2.99)
Pupil/Teacher Ratio	2.79 (3.51)	-0.31 (0.35)	-6.87 (7.37)	-0.22 (0.27)	-0.03 (0.13)	-0.28 (0.24)	-0.15 (0.22)	0.03 (0.17)	-0.28 (0.20)
Primary School	-136.13 (179.13)	-2.5 (27.60)	-395.53 (396.10)	-20.02 (29.88)	-1.15 (4.85)	-3.02 (12.53)	-12.08 (30.74)	-1.78 (15.76)	-4.05 (18.18)
Middle School	-198.77 (291.51)	18.98 (52.97)	790.88 (646.53)	-26.19 (37.66)	-1.98 (9.51)	2.42 (31.97)	12.91 (37.48)	28.54 (17.68)	-8.9 (23.02)
Enrollment	0.05 (0.13)	0.01 (0.02)	0.31 (0.40)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Adj. R^2	0.72	0.81	0.74	0.49	0.09	0.85	0.88	0.95	0.93
Observations	694	694	694	694	694	694	692	694	694
Schools	347	347	347	347	347	347	346	347	347

Notes. DD = difference in differences; PSM = propensity score matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A5. GDD Regression Estimates, Purposive Phil. Charter School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	148.64 (127.91)	31.23 (31.81)	1251.79*** (394.67)	13.69 (15.89)	11.33 (7.92)	-5.29 (12.62)	49.08 (33.24)	-165.29*** (32.13)	-125*** (27.21)
Treatment	60.54 (75.78)	27.53* (14.51)	221.75 (184.67)	23.13** (9.71)	-0.91 (2.66)	12.37** (6.07)	37.38* (20.42)	4.31 (22.34)	-5.81 (15.28)
Post1	-10.69 (7.17)	-0.39 (0.67)	-16.94 (19.44)	0.06 (0.59)	-1.33*** (0.48)	0.04 (0.37)	-1.6 (1.94)	-1.2 (1.41)	0.2 (1.21)
Treatment*Post1	10.84 (8.09)	2.89*** (1.01)	32.72 (25.56)	0.31 (0.87)	0.99* (0.55)	1.03* (0.53)	7.25*** (1.98)	-7.02*** (1.57)	-3.92*** (1.46)
Post2	-16.64 (13.91)	-0.32 (0.86)	-43.48* (22.41)	-0.84 (0.62)	-0.28 (1.16)	0.09 (0.49)	-2.91 (1.97)	-2.4 (1.53)	0.33 (1.48)
Treatment*Post2	5.59 (13.89)	0.88 (1.11)	31.92 (26.35)	-2.21*** (0.78)	-0.71 (1.19)	-0.38 (0.63)	13.6*** (2.16)	-6.46*** (1.83)	-3.6** (1.80)
Asian %	-191.43 (123.60)	-69.62** (33.76)	-1248.49*** (457.83)	-49.27 (30.55)	-20.75** (8.85)	-35.4* (20.32)	-127.44*** (39.53)	223.72*** (38.53)	174.47*** (29.82)
Black %	-181.46 (115.15)	-33.26 (27.56)	-693.25 (438.55)	-16.15 (14.92)	-14.97** (6.69)	5.17 (10.82)	-73.22*** (25.47)	188.5*** (26.99)	162.11*** (24.63)
Hispanic %	-245.8* (133.90)	-46.34 (41.93)	-1054.87** (518.60)	-24.72 (21.00)	-15.31 (10.98)	-8.32 (16.98)	-172.24*** (38.19)	181.96*** (35.91)	158.21*** (33.82)
White %	-93.36 (96.90)	-9.85 (32.37)	-641.35 (405.05)	-0.33 (15.93)	-16.9** (7.73)	2.61 (14.00)	-57.89** (29.41)	202.54*** (30.33)	171.54*** (27.95)

Table A5, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Female %	68.62 (56.55)	-1.8 (18.64)	-829.2 (528.60)	-0.42 (10.30)	3.91 (4.45)	-0.4 (10.60)	54.42*** (19.67)	35.13* (18.11)	18.91 (15.51)
FRPL %	2.68 (14.16)	3.75 (3.07)	-55.62 (34.50)	-0.57 (1.54)	1.94** (0.96)	1.05 (1.12)	-0.9 (3.36)	1.23 (3.01)	2.67 (2.31)
Pupil/Teacher Ratio	1.12 (1.09)	0.19 (0.16)	-1.93 (2.13)	-0.02 (0.09)	-0.08* (0.05)	0 (0.10)	-0.25 (0.19)	-0.04 (0.13)	-0.13 (0.11)
Primary School	6.02 (22.78)	-1.71 (2.23)	-102.59 (64.04)	-1.65 (2.29)	1.21 (1.30)	-0.14 (1.42)	-4.14 (7.47)	-2.5 (4.28)	-1.77 (3.24)
Middle School	0.78 (48.18)	-3.21 (7.53)	191.9 (129.29)	1.09 (4.95)	1.13 (2.05)	2.11 (4.30)	12.19 (13.82)	-3.82 (14.71)	-2.42 (8.02)
High School	-9.13 (14.40)	-0.41 (1.46)	-30.4 (72.60)	3.05 (2.15)	0.97 (1.32)	0.11 (0.62)	-4.52 (6.47)	1.72 (5.95)	3.37 (4.24)
Enrollment	0.01 (0.02)	0 (0.01)	0.12 (0.11)	0 (0.00)	0* (0.00)	0 (0.00)	0.02*** (0.01)	-0.01 (0.00)	0 (0.00)
Adj. R^2	0.50	0.80	0.57	0.51	0.30	0.82	0.74	0.83	0.85
Observations	1255	1255	1255	1255	1255	1255	1255	1255	1255
Schools	251	251	251	251	251	251	251	251	251

Notes. GDD = generalized difference in differences; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A6. GDD Regression Estimates, PSM Phil. Charter School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	283.46** (138.37)	24.53 (34.87)	1430.38*** (424.00)	18.61 (17.16)	9.52 (7.92)	-6.46 (13.70)	63.83 (41.53)	-159.2*** (33.05)	-121.35*** (29.30)
Treatment	115.28 (80.89)	32.02** (15.69)	344.68* (197.70)	20.73** (10.14)	0.43 (3.14)	12.6** (6.36)	42.29 (28.30)	-9.92 (21.94)	-5.5 (16.56)
Post1	-43.67 (27.75)	-0.56 (0.87)	-54.06** (24.95)	-0.68 (0.57)	-1.03* (0.56)	-0.22 (0.44)	-1.16 (2.79)	-2.34 (1.99)	-0.69 (1.53)
Treatment*Post1	43.18 (28.80)	3.01*** (1.04)	66.93** (30.99)	1.05 (0.86)	0.77 (0.64)	1.3** (0.53)	6.72** (2.92)	-5.86*** (2.07)	-3.19* (1.69)
Post2	-83.62*** (29.83)	-0.36 (1.01)	-52.24 (32.13)	-1.06 (0.70)	-0.28 (1.00)	-0.17 (0.52)	2.96 (5.08)	-2.54 (2.09)	-0.07 (2.48)
Treatment*Post2	72.52** (28.80)	0.83 (1.11)	38.04 (36.80)	-2** (0.79)	-0.6 (1.01)	-0.11 (0.59)	7.74 (5.59)	-6.43*** (2.19)	-3.47 (2.63)
Asian %	-310.5** (139.83)	-67.94* (35.06)	-1340.59*** (473.71)	-54.58* (31.60)	-19.89** (8.35)	-35.58* (19.99)	-124.33*** (43.61)	217.29*** (38.83)	168.92*** (31.52)
Black %	-327.26*** (126.83)	-31.05 (29.62)	-909.58** (451.32)	-19.4 (16.26)	-11.62* (6.08)	5.98 (11.51)	-87.68*** (30.93)	190.26*** (26.30)	155.56*** (25.85)
Hispanic %	-297.86* (154.09)	-44.13 (43.47)	-1146.63** (528.36)	-28.55 (22.23)	-11.31 (12.12)	-7.26 (18.69)	-179.42*** (43.12)	180.96*** (38.24)	153.28*** (34.19)
White %	-195.71* (117.50)	-6.69 (34.83)	-726.07* (402.29)	-3.12 (16.84)	-13.25* (7.51)	4.6 (15.18)	-48.04 (33.90)	191.03*** (30.76)	163.05*** (28.88)

Table A6, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Female %	161.7** (78.60)	0.53 (19.49)	-855.47 (549.19)	-1.12 (9.77)	2.76 (5.14)	0.19 (11.07)	51.89** (22.99)	29.19 (18.56)	13.69 (16.73)
FRPL %	-11.27 (18.35)	4.86 (3.81)	-65.75* (34.93)	-0.58 (1.88)	0.99** (0.45)	1.43 (1.42)	-4.55 (5.18)	2.22 (3.24)	3.29 (2.89)
Pupil/Teacher Ratio	0.06 (1.50)	0.23 (0.18)	-2.79 (2.24)	-0.04 (0.09)	-0.07 (0.05)	0.01 (0.10)	-0.26 (0.20)	-0.06 (0.14)	-0.13 (0.13)
Primary School	-64.47* (36.19)	-3.78 (4.39)	-150.07 (105.61)	-0.41 (2.70)	-0.82 (1.88)	-0.91 (2.44)	-5.52 (17.26)	4.17 (6.50)	2.15 (5.43)
Middle School	-25.65 (56.51)	-3.83 (9.24)	166.19 (111.07)	0.88 (5.97)	1.51 (2.46)	2.1 (3.92)	14.09 (17.25)	-2.42 (13.22)	-2.5 (9.44)
Enrollment	0.02 (0.03)	0 (0.01)	0.1 (0.12)	0 (0.00)	0** (0.00)	0 (0.00)	0.02** (0.01)	0 (0.00)	0 (0.00)
Adj. R^2	0.53	0.79	0.57	0.52	0.28	0.82	0.74	0.83	0.85
Observations	1170	1170	1170	1170	1170	1170	1170	1170	1170
Schools	234	234	234	234	234	234	234	234	234

Notes. GDD = generalized difference in differences; PSM = propensity scored matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A7. GDD Regression Estimates, Purposive Pennsylvania School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	76.92*	12.81**	137.86**	4.35	0.82	8.89**	8.53*	48.83***	59.22***
	(46.55)	(5.35)	(65.72)	(3.30)	(1.27)	(3.90)	(5.15)	(10.16)	(8.76)
Treatment	-173.9*	22.41***	-69.83	5.47	0.72	6.82	36.95***	-23.49***	-27.34***
	(101.13)	(6.98)	(98.24)	(4.92)	(2.03)	(4.67)	(6.11)	(7.12)	(7.80)
Post1	-8.68***	-0.03	-1.09	-0.05	-0.15***	-0.2**	0.67***	-2.06***	-1.42***
	(1.42)	(0.14)	(1.52)	(0.07)	(0.04)	(0.09)	(0.13)	(0.12)	(0.12)
Treatment*Post1	14.41***	3***	24.47*	0.39	-0.04	1.01***	4.67***	-8.6***	-4.77***
	(2.86)	(0.74)	(13.22)	(0.53)	(0.22)	(0.39)	(0.57)	(0.56)	(0.55)
Post2	-14.49***	-0.53***	-7.77***	-0.38***	-0.2***	-0.53***	0.44***	-2.89***	-1.27***
	(1.79)	(0.17)	(1.91)	(0.07)	(0.04)	(0.10)	(0.13)	(0.17)	(0.14)
Treatment*Post2	11.98***	1.97***	10.45	-2.62***	-0.49***	-0.01	10.25***	-9.78***	-5.72***
	(3.41)	(0.71)	(11.55)	(0.40)	(0.15)	(0.41)	(0.77)	(0.66)	(0.63)
Asian %	-124.47	-18.93***	-192.93*	-26.64***	-2.68	-12.56*	-52.83***	35.01***	24.75***
	(99.03)	(7.27)	(114.65)	(8.61)	(2.34)	(6.76)	(8.29)	(10.25)	(8.22)
Black %	171.23**	-3.22	241.31***	-2.19	0.69	0.14	-8.62	21.31***	6.67
	(85.94)	(5.64)	(87.79)	(3.40)	(1.36)	(2.98)	(5.82)	(7.41)	(5.84)
Hispanic %	163.59***	2.96	37.82	2.77	-0.56	-0.75	-40.59***	-14.31	-13.54*
	(58.03)	(5.87)	(106.42)	(2.90)	(1.73)	(3.51)	(6.86)	(9.65)	(7.14)
White %	-36.86	-1.24	-137.1**	-0.26	-0.54	-3.39	-14.64***	29.5***	19.99***
	(45.05)	(3.72)	(62.79)	(1.64)	(0.94)	(2.20)	(4.62)	(8.29)	(5.82)

Table A7, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Female %	-15.12 (48.04)	-5.72* (3.14)	1.65 (51.16)	-0.06 (1.61)	0.01 (0.92)	-0.31 (2.04)	15.69*** (2.94)	-10.52*** (3.20)	-4.38 (2.79)
FRPL %	4.65 (10.43)	2.01 (1.24)	-13.15 (18.08)	0.11 (0.57)	0.18 (0.23)	1.19* (0.65)	3.11*** (1.02)	-0.32 (1.18)	0.08 (1.20)
Pupil/Teacher Ratio	-0.02 (0.25)	0 (0.02)	-0.1 (0.25)	0 (0.00)	0 (0.00)	-0.01 (0.01)	0.04** (0.02)	-0.12*** (0.03)	-0.09*** (0.03)
Primary School	8.92 (6.17)	0.52 (1.02)	19.97 (14.45)	0.64 (0.48)	0.09 (0.27)	0.25 (0.59)	-0.39 (0.73)	-1.75** (0.79)	-0.38 (0.71)
Middle School	13.61 (10.21)	2.05 (1.80)	35.03 (24.24)	0.51 (0.56)	0.47 (0.36)	1.35 (1.11)	-0.15 (0.89)	-1.07 (0.88)	-2.85*** (0.66)
High School	2.77 (7.25)	-2.82 (3.15)	0.88 (10.06)	-1.58 (3.00)	0.18 (0.67)	-2.78 (2.63)	0.13 (1.54)	-1.72 (1.28)	1.58 (1.34)
Enrollment	0 (0.01)	0 (0.00)	0.02 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0* (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Adj. R^2	0.63	0.70	0.74	0.43	0.42	0.64	0.77	0.88	0.88
Observations	12755	12755	12755	12755	12755	12755	12751	12755	12755
Schools	2551	2551	2551	2551	2551	2551	2550.2	2551	2551

Notes. GDD = generalized difference in differences; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A8. GDD Regression Estimates, PSM Pennsylvania School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	190.24 (487.72)	21.72 (24.95)	416.12 (584.25)	-4.51 (14.08)	3.14 (5.07)	3.33 (13.61)	48.22 (30.82)	29.37 (34.84)	14.93 (27.84)
Treatment	-127.94 (208.78)	36.31* (19.90)	59.65 (415.41)	10.98 (10.28)	-0.4 (3.33)	11.25 (9.78)	24.37 (20.87)	-63.6* (37.63)	-43.26 (29.54)
Post1	-23.09 (24.47)	-0.56 (0.77)	14.37 (18.20)	-0.02 (0.57)	-0.29* (0.16)	0.38 (0.47)	0.54 (1.35)	-6.66*** (0.93)	-3.47*** (0.83)
Treatment*Post1	21.67 (24.51)	2.81** (1.14)	9.52 (25.28)	0.56 (0.75)	0.11 (0.30)	0.42 (0.67)	5.23*** (1.44)	-3.3*** (1.02)	-1.72* (1.03)
Post2	-49.72* (27.00)	0.53 (0.82)	-8.07 (23.06)	-0.17 (0.53)	-0.05 (0.30)	0.74 (0.63)	-0.7 (1.26)	-6.84*** (1.17)	-3.23*** (0.91)
Treatment*Post2	34.95 (25.08)	-0.34 (1.10)	7.24 (28.83)	-2.64*** (0.67)	-0.69** (0.30)	-1.41* (0.80)	11.38*** (1.57)	-4.7*** (1.26)	-2.22* (1.15)
Asian %	-643.8 (877.89)	-54.4*** (20.99)	-954.57 (632.27)	-63.24** (30.01)	-18.29*** (6.83)	-37.22** (17.76)	-155.69*** (26.30)	81.9* (41.95)	69.77** (29.10)
Black %	-102.34 (376.75)	-32.91 (20.10)	-240.99 (423.90)	-4.74 (12.90)	-5.94 (3.96)	0.2 (10.59)	-80.05*** (20.17)	87.17*** (24.05)	75.12*** (21.52)
Hispanic %	-109.04 (423.96)	-44.36 (31.10)	-964 (777.17)	-10.41 (17.05)	-6.8 (5.55)	-9.61 (15.94)	-166.88*** (29.99)	52.2* (29.87)	47.21* (26.43)
White %	-276.26 (415.27)	-16.59 (22.57)	-550.93 (457.16)	3.92 (13.44)	-8.56* (4.52)	-2.58 (12.21)	-63.32*** (22.84)	68.62*** (26.50)	75.17*** (22.95)

Table A8, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Female %	101.42 (224.26)	-8.82 (14.72)	-363.21 (419.45)	0.89 (8.28)	3.28 (3.37)	-4.84 (9.55)	61.69*** (15.37)	-22.75* (13.84)	-21.23* (11.96)
FRPL %	33.44 (23.29)	6.55* (3.57)	-81.66** (36.55)	-0.81 (1.66)	0.46 (0.40)	2.73** (1.38)	3.49 (2.69)	0.1 (3.07)	1.48 (2.72)
Pupil/Teacher Ratio	2.34 (1.99)	0.2 (0.15)	-1.06 (2.11)	0 (0.08)	-0.04 (0.04)	0.05 (0.09)	-0.42** (0.17)	-0.02 (0.12)	-0.09 (0.11)
Primary School	-4.08 (131.09)	2.85 (8.51)	255.89 (299.04)	5.03 (5.36)	1.98 (2.60)	1.27 (5.00)	4.02 (20.33)	4.44 (17.14)	5.19 (12.62)
Middle School	-28.6 (131.51)	2.47 (8.38)	413.99 (324.62)	3.89 (5.55)	3.28 (2.63)	0.98 (4.90)	9.61 (21.01)	1.15 (18.27)	0.87 (12.55)
Enrollment	-0.01 (0.04)	0 (0.01)	0.15 (0.13)	0 (0.00)	0** (0.00)	0 (0.00)	0.03*** (0.01)	-0.01*** (0.00)	-0.01*** (0.00)
Adj. R^2	0.70	0.79	0.66	0.53	0.22	0.81	0.77	0.84	0.86
Observations	1540	1540	1540	1540	1540	1540	1538	1540	1540
Schools	308	308	308	308	308	308	307.6	308	308

Notes. GDD = generalized difference in differences; PSM = propensity score matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A9. CITS Regression Estimates, Purposive Phil. Charter School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	131.05 (141.10)	-3.88 (37.69)	993.63** (474.49)	10.43 (16.96)	12.09 (9.24)	-13.78 (16.29)	35.99 (38.52)	-95.48*** (31.94)	-77.17*** (29.06)
Treatment	36.43 (79.10)	28.19** (13.96)	196.46 (205.01)	22.95** (9.85)	-1.11 (2.59)	12.57** (6.41)	38.06* (21.21)	-2.46 (21.22)	-11.16 (13.97)
Year	17.24* (8.84)	0.59 (0.37)	25.22** (12.31)	0.22 (0.23)	0.12 (0.36)	0.12 (0.15)	-0.08 (1.14)	2.62*** (0.99)	2.28*** (0.71)
Treatment*Year	-15.78* (9.11)	1.47** (0.66)	-9.55 (15.62)	-0.03 (0.40)	-0.16 (0.36)	0.38 (0.35)	0.84 (1.20)	-6.62*** (1.04)	-5.01*** (0.82)
Post1	-43.34** (16.90)	-1.24 (0.78)	-62.98** (28.92)	-0.34 (0.42)	-1.55** (0.74)	-0.11 (0.34)	-1.35 (2.90)	-6.71*** (1.55)	-4.5*** (1.47)
Treatment*Post1	41.73** (17.25)	0.66 (1.48)	55.74* (32.07)	0.42 (0.94)	1.28 (0.79)	0.44 (0.73)	5.84* (3.03)	4.59*** (1.72)	4.96*** (1.63)
Post2	-65.84* (34.92)	-1.75* (1.04)	-113.76*** (39.25)	-1.45** (0.70)	-0.62 (1.30)	-0.18 (0.42)	-2.58 (3.57)	-10.41*** (2.73)	-6.55*** (2.20)
Treatment*Post2	52.25 (35.32)	-2.37 (1.98)	67.56 (45.04)	-2.03 (1.31)	-0.27 (1.33)	-1.24 (0.99)	11.53*** (3.81)	10.81*** (2.92)	9.62*** (2.44)
Asian %	-169.21 (128.72)	-41.89 (36.70)	-1034.24** (525.92)	-46.61 (32.57)	-21.29** (10.06)	-28.7 (21.40)	-117.24*** (42.88)	170.42*** (40.48)	138.19*** (31.23)
Black %	-126.45 (131.47)	8.83 (34.46)	-341.88 (638.30)	-11.91 (17.73)	-15.62* (8.83)	15.3 (15.13)	-58.13* (33.63)	112.39*** (28.55)	110.93*** (27.34)

Table A9, continued

Notes. CITS = comparative interrupted time series; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A10. *CITS Regression Estimates, PSM Phil. Charter School Control*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	180.96 (152.38)	-11.43 (38.60)	1065.27** (487.73)	15.27 (17.95)	10.63 (8.78)	-15.5 (16.00)	46.65 (46.91)	-95.45*** (34.13)	-78*** (30.21)
Treatment	85.71 (73.81)	32.71** (13.96)	313.87 (229.43)	20.87** (9.91)	0.67 (3.00)	12.75** (6.30)	40.92 (28.55)	-13.5 (20.70)	-8.63 (15.10)
Year	39.4* (20.14)	0.54 (0.35)	50.89*** (19.52)	-0.04 (0.18)	-0.34 (0.29)	0.17 (0.18)	2.3** (1.14)	1.86* (1.03)	2.11*** (0.70)
Treatment*Year	-38.04* (20.76)	1.51** (0.59)	-35.4 (22.28)	0.25 (0.35)	0.31 (0.29)	0.35 (0.33)	-1.56 (1.22)	-5.85*** (1.08)	-4.92*** (0.77)
Post1	-118.77** (50.77)	-1.14 (0.75)	-147.95*** (52.10)	-0.55 (0.44)	-0.38 (0.62)	-0.42 (0.37)	-5.4 (3.47)	-6.81*** (2.57)	-5.36*** (1.83)
Treatment*Post1	118.25** (51.77)	0.54 (1.39)	140.27*** (54.23)	0.61 (0.95)	0.15 (0.67)	0.75 (0.75)	9.97*** (3.50)	4.66* (2.73)	5.78*** (2.02)
Post2	-197.22** (85.63)	-1.48* (0.77)	-195.92*** (64.79)	-0.89 (0.55)	0.7 (1.30)	-0.54 (0.43)	-3.53 (5.69)	-8.81*** (3.41)	-6.78** (3.26)
Treatment*Post2	186.19** (86.58)	-2.66 (1.75)	150.76** (68.64)	-2.63** (1.18)	-1.54 (1.30)	-0.89 (0.94)	12.75** (5.92)	9** (3.62)	9.76*** (3.51)
Asian %	-196.62 (141.83)	-39.04 (37.23)	-1009.47* (519.37)	-51.98 (32.17)	-21.06** (8.97)	-28.29 (21.58)	-108.83** (47.53)	168.43*** (40.92)	136.39*** (32.00)
Black %	-175.67 (137.70)	12.5 (35.53)	-434.51 (639.77)	-15.42 (18.17)	-13.21* (8.01)	16.96 (15.25)	-65.39* (39.20)	115.1*** (29.66)	105.06*** (28.38)

Table A10, continued

Notes. CITS = comparative interrupted time series; PSM = propensity score matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A11. CITS Regression Estimates, Purposive Pennsylvania School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	92.04*	9.45*	108.74*	4.63	0.77	8.31**	6.71	51.03***	58.93***
	(47.95)	(5.03)	(58.76)	(3.32)	(1.28)	(3.88)	(4.82)	(9.61)	(8.89)
Treatment	-181.75*	22.84***	-64.11	5.3	0.73	6.92	37.31***	-22.43***	-25.72***
	(101.70)	(6.93)	(98.72)	(4.92)	(2.04)	(4.67)	(6.11)	(7.13)	(7.79)
Year	-3.32***	0.43***	4.2***	-0.07*	0.01	0.08	0.27***	0.03	0.41***
	(0.81)	(0.09)	(1.04)	(0.04)	(0.03)	(0.06)	(0.07)	(0.07)	(0.07)
Treatment*Year	7.1***	1.42***	7.68	0.2	0.02	0.19	0.44	-3.96***	-3.49***
	(1.58)	(0.47)	(7.16)	(0.29)	(0.10)	(0.25)	(0.44)	(0.38)	(0.34)
Post1	-2.59	-0.78***	-8.52***	0.07	-0.16**	-0.35**	0.19	-2.18***	-2.21***
	(1.82)	(0.22)	(2.20)	(0.10)	(0.07)	(0.14)	(0.17)	(0.16)	(0.16)
Treatment*Post1	0.21	0.24	9.65	-0.01	-0.08	0.64	3.82***	-0.8	2.12***
	(3.42)	(1.16)	(17.65)	(0.78)	(0.30)	(0.60)	(0.92)	(0.85)	(0.79)
Post2	-5.35**	-1.66***	-18.95***	-0.19	-0.22**	-0.74***	-0.27	-3.06***	-2.45***
	(2.38)	(0.30)	(3.07)	(0.12)	(0.09)	(0.19)	(0.20)	(0.23)	(0.21)
Treatment*Post2	-9.31**	-2.19	-11.95	-3.22***	-0.55	-0.57	8.97***	1.94	4.65***
	(4.53)	(1.52)	(22.16)	(0.95)	(0.34)	(0.82)	(1.39)	(1.22)	(1.10)
Asian %	-139.34	-16.87**	-173.17	-26.94***	-2.64	-12.18*	-51.58***	34.92***	26.42***
	(99.17)	(7.21)	(112.96)	(8.62)	(2.35)	(6.77)	(8.19)	(9.92)	(8.34)
Black %	156.86*	3.8	296.26***	-2.37	0.8	1.29	-5.21	12.84*	2.66
	(85.99)	(5.50)	(89.73)	(3.38)	(1.40)	(3.05)	(5.74)	(7.17)	(5.98)

Table A11, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Hispanic %	169.32*** (59.01)	3.39 (5.59)	38.93 (103.09)	2.92 (2.93)	-0.55 (1.73)	-0.71 (3.47)	-40.53*** (6.59)	-16.31* (9.01)	-15.56** (7.19)
White %	-63.5 (47.44)	4.52 (3.49)	-86.92 (58.24)	-0.75 (1.74)	-0.45 (1.00)	-2.38 (2.28)	-11.49*** (4.39)	25.88*** (7.84)	20.67*** (6.06)
Female %	-3.59 (48.37)	-8.42*** (3.12)	-21.57 (52.12)	0.16 (1.62)	-0.03 (0.94)	-0.78 (2.01)	14.24*** (2.95)	-8.6*** (3.14)	-4.44 (2.80)
FRPL %	8.59 (10.58)	1.54 (1.24)	-17.81 (18.15)	0.2 (0.57)	0.17 (0.23)	1.1* (0.66)	2.81*** (1.03)	-0.43 (1.19)	-0.45 (1.21)
Pupil/Teacher Ratio	0.02 (0.26)	-0.01 (0.02)	-0.19 (0.22)	0 (0.00)	0 (0.00)	-0.01 (0.01)	0.03 (0.02)	-0.11*** (0.02)	-0.09*** (0.03)
Primary School	10.35* (6.15)	0.34 (1.02)	18.2 (14.59)	0.67 (0.48)	0.09 (0.27)	0.22 (0.59)	-0.5 (0.73)	-1.77** (0.79)	-0.56 (0.70)
Middle School	14.05 (10.09)	1.98 (1.79)	34.36 (24.32)	0.52 (0.56)	0.47 (0.36)	1.34 (1.11)	-0.2 (0.89)	-1.05 (0.87)	-2.88*** (0.67)
High School	2.33 (7.31)	-2.91 (3.14)	0.37 (10.25)	-1.6 (3.00)	0.18 (0.67)	-2.8 (2.63)	0.1 (1.55)	-1.46 (1.27)	1.8 (1.37)
Enrollment	0 (0.01)	0 (0.00)	0.02 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0* (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Adj. R^2	0.63	0.70	0.74	0.43	0.42	0.64	0.77	0.88	0.88
Observations	12755	12755	12755	12755	12755	12755	12751	12755	12755
Schools	2551	2551	2551	2551	2551	2551	2550.2	2551	2551

Table A11, continued

Notes. CITS = comparative interrupted time series; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A12. *CITS Regression Estimates, PSM Pennsylvania School Control*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	265.86 (491.33)	1.28 (27.49)	135.93 (591.24)	-5.83 (14.61)	4.16 (5.64)	-0.44 (14.51)	38.11 (30.20)	76.6** (33.15)	48.11* (27.64)
Treatment	-150.78 (177.02)	33.28* (17.27)	65.06 (434.51)	10.13 (9.89)	-0.76 (3.43)	10.61 (9.91)	23.34 (21.17)	-56.69* (29.26)	-38.07 (26.28)
Year	-23.26*** (8.96)	0.76* (0.46)	38.62** (16.24)	-0.35 (0.23)	-0.35*** (0.13)	0.09 (0.27)	0.69 (0.79)	-1.81*** (0.59)	-1.07*** (0.41)
Treatment*Year	22.12*** (8.15)	1.12* (0.67)	-20.86 (15.83)	0.58 (0.43)	0.34** (0.15)	0.27 (0.36)	0.15 (0.84)	-2.52*** (0.73)	-2.03*** (0.57)
Post1	19.85 (23.47)	-1.6 (1.05)	-53.75* (28.36)	0.66 (0.59)	0.35 (0.23)	0.29 (0.69)	-0.59 (1.77)	-4.16*** (1.14)	-2.1** (0.87)
Treatment*Post1	-22.07 (24.56)	0.85 (1.65)	52.87* (32.07)	-0.55 (1.05)	-0.56 (0.38)	-0.07 (0.94)	5.05*** (1.86)	1.12 (1.26)	1.89* (1.08)
Post2	15.94 (25.02)	-1.25 (1.37)	-113.86** (44.15)	0.85 (0.68)	0.93** (0.47)	0.56 (1.00)	-2.51 (1.69)	-2.59 (1.61)	-0.83 (1.26)
Treatment*Post2	-31.63 (28.12)	-3.1 (2.13)	75.12 (46.73)	-4.3*** (1.32)	-1.72*** (0.54)	-2.11 (1.30)	11.21*** (1.94)	1.52 (1.84)	2.9* (1.60)
Asian %	-635.92 (880.62)	-39.32* (22.25)	-835.55 (661.22)	-61.04** (28.98)	-18.07** (7.44)	-34.28* (17.69)	-149.22*** (27.09)	47.2 (38.86)	44.76* (26.59)
Black %	-197.03 (383.37)	-3.89 (22.90)	139.29 (531.94)	-2.62 (14.26)	-7.19 (5.19)	5.59 (12.81)	-65.8*** (24.13)	20.16 (24.55)	27.92 (20.78)

Table A12, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Hispanic %	-145.85 (412.76)	-26.55 (32.51)	-759.93 (725.73)	-8.7 (18.57)	-7.25 (6.46)	-6.25 (17.15)	-158.37*** (31.52)	11.11 (28.82)	18.06 (26.00)
White %	-412.83 (435.70)	8.33 (26.85)	-148.21 (493.71)	4.67 (15.13)	-10.48* (5.47)	1.91 (14.31)	-50.33** (25.47)	10.94 (25.68)	35.1 (23.05)
Female %	132.93 (226.00)	-13.92 (15.42)	-450.5 (449.12)	0.8 (8.07)	3.73 (3.40)	-5.75 (9.50)	58.92*** (16.11)	-10.94 (14.18)	-13.06 (12.50)
FRPL %	36.79 (24.40)	6.4* (3.54)	-87.54** (35.92)	-0.77 (1.73)	0.51 (0.41)	2.71* (1.40)	3.38 (2.90)	0.44 (3.04)	1.7 (2.62)
Pupil/Teacher Ratio	2.72 (2.01)	0.06 (0.14)	-2.8 (2.42)	-0.01 (0.07)	-0.04 (0.04)	0.02 (0.09)	-0.49*** (0.16)	0.31** (0.13)	0.14 (0.11)
Primary School	12.22 (79.64)	4.28 (8.23)	245.76 (312.88)	5.54 (5.37)	2.24 (2.51)	1.59 (4.88)	4.41 (18.51)	1.18 (16.63)	2.68 (12.54)
Middle School	-34.97 (81.57)	4.44 (8.54)	439.72 (335.36)	4.03 (5.38)	3.2 (2.66)	1.34 (4.71)	10.57 (19.66)	-3.39 (16.85)	-2.33 (13.07)
Enrollment	-0.01 (0.05)	0 (0.01)	0.17 (0.13)	0 (0.00)	0* (0.00)	0 (0.00)	0.03*** (0.01)	-0.02*** (0.00)	-0.01*** (0.00)
Adj. R^2	0.71	0.80	0.66	0.53	0.22	0.81	0.78	0.86	0.87
Observations	1540	1540	1540	1540	1540	1540	1538	1540	1540
Schools	308	308	308	308	308	308	307.6	308	308

Table A12, continued

Notes. CITS = comparative interrupted time series; PSM = propensity score matched; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Post1 indicates one year post-treatment; Post2 indicates two years post-treatment. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A13. DD Regression Estimates, Purposive SDP School Control

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	245.49 (225.80)	-85.78 (69.04)	1317.55 (984.05)	70.18 (84.60)	-4.31 (16.25)	-51.19 (45.74)	-10.75 (55.02)	37.18 (44.62)	77.47** (37.64)
Treatment	175.24 (144.37)	36.24 (46.59)	729.89 (623.35)	25.93 (33.14)	-0.5 (5.65)	40.87 (34.43)	51.77 (39.51)	-21.19 (27.35)	-20.24 (25.41)
Post	14.53*** (3.09)	2.6** (1.10)	40.01** (17.19)	4.09*** (0.95)	0.63* (0.34)	2.05*** (0.71)	-5.94*** (0.84)	2.74*** (0.58)	1.54*** (0.58)
Treatment*Post	-0.33 (7.28)	-4.1** (1.70)	4.53 (29.56)	0.53 (2.61)	-0.26 (0.36)	-2.05** (0.87)	2.38* (1.24)	-2.6** (1.22)	-2.25** (1.10)
Asian %	-575.49* (294.63)	-18.59 (98.65)	-1807.63 (1161.55)	-87.4 (93.16)	-17.98 (18.91)	-53.18 (57.85)	89.08 (65.20)	31.96 (50.06)	-22.03 (66.06)
Black %	-509.17** (236.75)	23.06 (65.59)	-2152.36** (1041.50)	-89.12 (85.69)	-17.1 (15.71)	-23.57 (35.66)	6.23 (49.23)	35.82 (43.19)	12.53 (36.14)
Hispanic %	-503.13* (271.22)	5 (68.05)	-1534.97 (1202.08)	-99.88 (103.69)	-23.83 (23.48)	-37.85 (38.04)	-61.33 (59.07)	37.55 (47.53)	2.8 (36.78)
White %	-278.03 (210.12)	73.58 (91.68)	-1588.12 (995.14)	-41.09 (69.92)	-24.62 (19.08)	41.04 (66.53)	79.63 (58.80)	10.68 (42.56)	-5.12 (40.21)
Female %	76.08 (95.75)	29.14 (43.46)	218.51 (668.05)	-26.2 (32.56)	35.41 (30.53)	39.05 (28.48)	-27.42 (31.11)	7.26 (26.66)	14.16 (25.56)

Table A13, continued

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	25.01 (17.18)	16.16 (10.03)	-4.31 (133.44)	4.07 (3.86)	1.56 (1.66)	7.2 (5.88)	4.47 (7.01)	-1.39 (4.35)	6.2 (4.94)
Pupil/Teacher Ratio	0.92 (0.77)	0.36 (0.27)	0.08 (3.36)	-0.04 (0.21)	-0.04 (0.07)	0.26 (0.19)	0 (0.15)	0.34*** (0.13)	0.15 (0.14)
Primary School	246.86** (111.33)	58.11 (49.64)	891.87* (468.85)	43.92 (32.27)	3.45 (5.55)	48.53 (36.63)	71.78** (34.35)	-54.57** (24.07)	-72.21*** (24.52)
Middle School	198.66** (99.30)	69.53 (44.78)	770.86* (451.72)	27.36 (30.60)	5.79 (6.91)	65.73* (35.18)	52.36 (35.59)	-34.14* (20.53)	-53.35** (21.31)
Enrollment	0.05** (0.03)	0.01 (0.02)	0.32* (0.18)	0.01 (0.01)	0 (0.00)	0.01 (0.01)	0 (0.01)	0 (0.01)	-0.01 (0.01)
Adj. R^2	0.77	0.86	0.82	0.46	0.03	0.81	0.87	0.96	0.95
Observations	394	394	394	394	394	394	394	394	394
Schools	197	197	197	197	197	197	197	197	197

Notes. DD = difference in differences; SDP = School District of Philadelphia; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A14. *DD Regression Estimates, PSM SDP School Control*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	425.04 (533.04)	-226.51 (140.54)	1766.3 (2267.30)	12.09 (231.60)	14.6 (12.70)	-1.8 (62.60)	63.83 (94.26)	-10.62 (106.63)	77.46 (85.75)
Treatment	109.29 (180.68)	31.22 (44.46)	306.49 (612.46)	35.51 (55.70)	-1.76 (3.55)	6.34 (18.93)	-3.46 (47.68)	57.57 (44.37)	22.43 (36.61)
Post	10.8 (11.64)	3.83 (3.48)	30.6 (35.88)	1.25 (2.64)	0.01 (0.15)	1.49 (1.18)	-4.35*** (1.53)	2.62 (1.80)	-0.88 (2.16)
Treatment*Post	11.13 (12.64)	-3.25 (2.65)	8.1 (49.01)	1.34 (4.98)	0.38* (0.21)	-1.06 (1.19)	2.54 (1.78)	-0.57 (1.91)	-1.33 (1.96)
Asian %	-1152.98* (700.32)	-41.02 (182.20)	-254.21 (2685.94)	19.28 (220.70)	-7.39 (7.63)	6.56 (68.87)	-27.24 (104.45)	-4.44 (100.40)	169.12 (114.94)
Black %	-837.57 (615.16)	38.31 (163.63)	-1715.85 (2329.78)	-1.43 (236.66)	-12.83 (10.13)	-22.7 (57.64)	-84.19 (71.61)	-42.19 (89.30)	63.23 (96.78)
Hispanic %	-464.94 (513.04)	266.43** (122.55)	-1538.59 (2228.54)	-85.81 (216.84)	1.21 (6.12)	45.45 (45.86)	-120.08 (88.06)	38.45 (77.04)	80.3 (72.23)
White %	-429.33 (580.79)	151.83 (156.11)	-1078.45 (2002.81)	-12.22 (172.85)	-8.36 (9.11)	2.74 (62.13)	7.85 (79.54)	6.05 (99.76)	68.38 (95.07)
Female %	192.88 (411.93)	152.26 (125.16)	-828.28 (1097.25)	-90.45 (106.21)	2.4 (5.72)	16.39 (37.62)	67.8 (49.45)	31.38 (65.91)	-87.07 (76.34)

Table A14, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	380.57 (239.43)	111.34 (70.97)	560.51 (828.72)	79.49 (75.44)	-4.61 (5.35)	14.91 (28.73)	6.12 (36.02)	53.08 (42.02)	-35.14 (39.51)
Pupil/Teacher Ratio	-3.29 (4.01)	0.34 (0.98)	-7.64 (15.98)	-2.13 (1.68)	0.04 (0.05)	-0.11 (0.45)	0.67 (0.66)	0.03 (0.55)	-0.42 (0.49)
Enrollment	0.03 (0.13)	-0.02 (0.02)	0.27 (0.44)	0 (0.05)	0 (0.00)	0.01 (0.02)	-0.01 (0.01)	0 (0.02)	0.01 (0.02)
Adj. R^2	0.64	0.59	0.64	0.30	0.27	0.23	0.94	0.96	0.96
Observations	96	96	96	96	96	96	96	96	96
Schools	48	48	48	48	48	48	48	48	48

Notes. School level is not included as a covariate here as the sample contains almost no variance on this measure. DD = difference in differences; PSM = propensity score matched; SDP = School District of Philadelphia; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A15. FRD Intent-to-Treat Estimates

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	65.59 (59.63)	11.66 (23.52)	-186.25 (489.85)	0.31 (10.35)	0.78 (0.80)	13.8 (17.21)	-84.72*** (30.39)	54.8 (36.32)	102.34*** (37.13)
Treatment	-3.49 (4.58)	1.44 (2.06)	18.49 (45.71)	-0.56 (0.88)	0.04 (0.09)	1.4 (1.52)	2.48 (2.81)	-0.32 (3.21)	0.56 (3.24)
Assignment Variable	1.07*** (0.14)	0.71*** (0.08)	5.08*** (1.87)	0.13*** (0.03)	0.01* (0.01)	0.3*** (0.04)	0.15*** (0.05)	-0.17** (0.07)	-0.15** (0.06)
Treatment*Assignment	0.56 (0.53)	-0.11 (0.26)	-2.09 (5.38)	-0.06 (0.10)	-0.02** (0.01)	-0.3 (0.20)	1.28*** (0.36)	-0.65 (0.41)	-0.21 (0.43)
Asian %	11.58 (61.01)	40.02* (22.19)	432.82 (456.48)	18.56* (9.67)	-0.48 (0.78)	21.58 (14.42)	72.57** (31.14)	-8.74 (34.55)	-56.15 (34.49)
Black %	15.33 (55.83)	40.2* (22.39)	555.84 (466.71)	19.21** (9.57)	-0.45 (0.78)	26.77* (14.70)	88.03*** (31.55)	-25.18 (34.42)	-62.91* (34.75)
Hispanic %	10.58 (57.17)	-44.46** (18.84)	-337.59 (219.85)	-21.56** (8.43)	-0.64 (0.58)	-41.97** (16.29)	-2.69 (12.91)	76.24*** (20.92)	88.75*** (21.15)
White %	10.24 (60.04)	-4.08 (5.77)	93.25 (57.13)	-0.14 (2.06)	0.15 (0.18)	-0.39 (4.25)	9.32** (4.66)	-27.71*** (6.38)	-32.47*** (5.13)
Female %	-64.55** (29.91)	-0.57** (0.27)	-11.87* (6.31)	-0.11 (0.11)	-0.01 (0.01)	-0.5*** (0.19)	0.2 (0.20)	0.91*** (0.34)	0.82*** (0.31)

Table A15, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	-2.7 (9.00)	-8.93** (3.51)	-18.18 (67.04)	-4.16*** (1.48)	-0.29* (0.17)	-11.82*** (3.01)	16.23*** (2.26)	3.29 (3.56)	-14.54*** (3.85)
Pupil/Teacher Ratio	-0.44 (0.58)	-7.48 (4.84)	21.96 (99.82)	-3.87** (1.96)	-0.43* (0.24)	-11.44*** (3.17)	19.62*** (3.56)	20.15*** (4.22)	7.8* (4.16)
Primary School	0.88 (6.56)	0 (0.00)	0.06* (0.03)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	-0.01** (0.00)
Middle School	-5.44 (8.48)	-0.11 (0.26)	-2.09 (5.38)	-0.06 (0.10)	-0.02** (0.01)	-0.3 (0.20)	1.28*** (0.36)	-0.65 (0.41)	-0.21 (0.43)
Enrollment	-0.01 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Adj. R^2	0.60	0.77	0.35	0.50	0.19	0.67	0.63	0.73	0.76
Observations	197	197	197	197	197	197	197	197	197
Schools	197	197	197	197	197	197	197	197	197

Notes. FRD = fuzzy regression discontinuity; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

Table A16. FRD Treatment-on-the-Treated Estimates

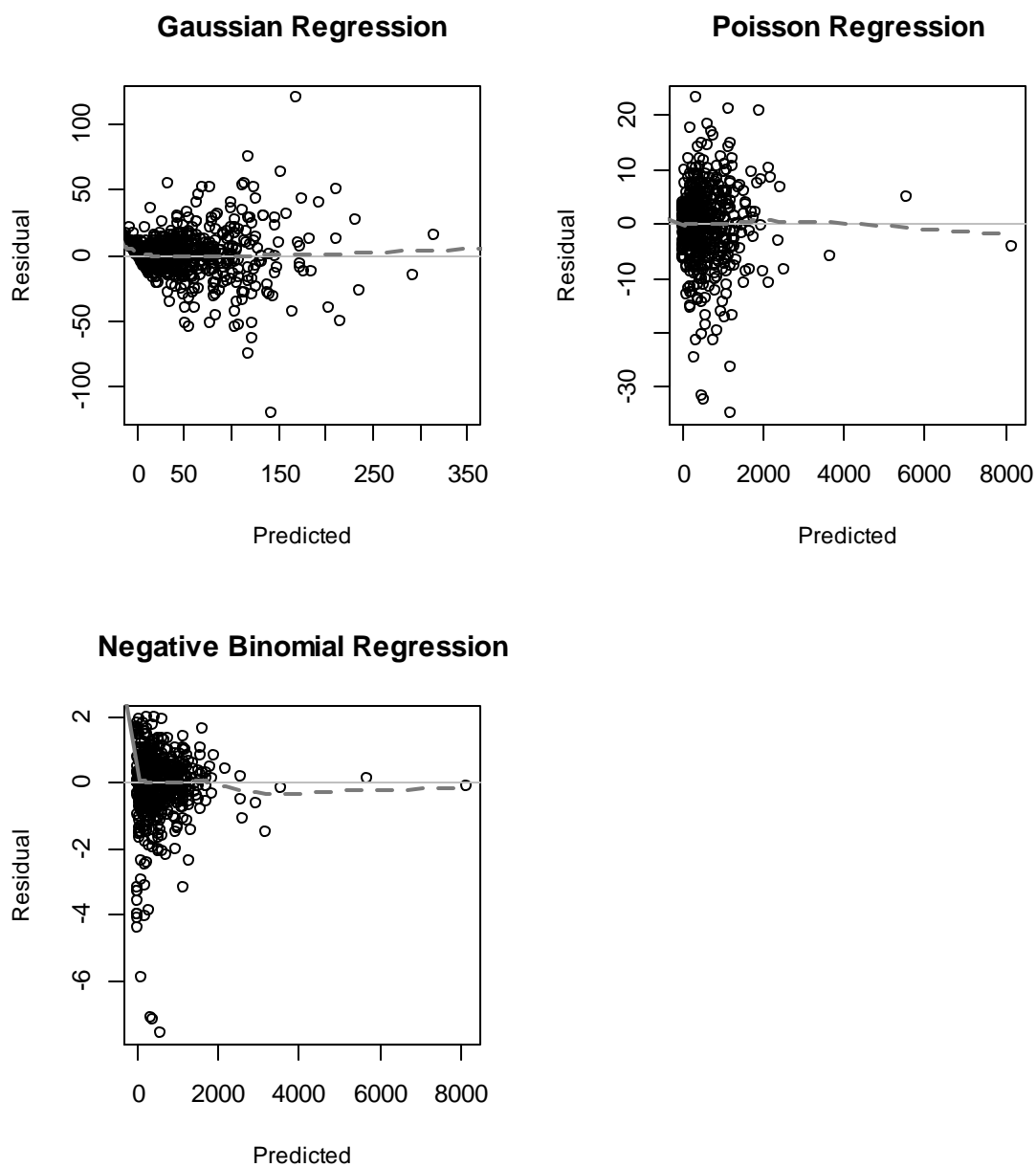
School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
Intercept	53.34 (1065.26)	16.71 (4663.10)	-121.33 (7577.38)	-1.65 (271.44)	0.94 (59.24)	18.71 (2294.93)	-75.99 (4102.07)	53.67 (738.40)	104.3 (1043.61)
Treatment	-37.88 (1230.56)	15.62 (2380.57)	200.75 (10633.46)	-6.07 (200.03)	0.47 (59.78)	15.18 (4096.94)	26.97 (1390.77)	-3.52 (744.40)	6.07 (602.64)
Assignment Variable	1 (2.58)	0.74 (2.01)	5.45 (23.20)	0.11 (0.45)	0.01 (0.11)	0.33 (11.93)	0.2 (3.48)	-0.17 (1.95)	-0.14 (1.49)
Treatment*Assignment	0.88 (19.54)	-0.25 (4.98)	-3.79 (58.73)	-0.01 (0.70)	-0.03 (0.32)	-0.43 (39.54)	1.05 (10.38)	-0.62 (6.63)	-0.26 (8.68)
Asian %	50.67 (1898.77)	23.9 (6615.18)	225.63 (7717.82)	24.82 (353.71)	-0.97 (103.88)	5.91 (7449.36)	44.74 (5503.31)	-5.11 (1646.49)	-62.42 (1187.61)
Black %	37.62 (1420.55)	27.6 (7305.44)	393.95 (9111.22)	24.1 (366.94)	-0.83 (69.83)	14.53 (5226.04)	66.28 (4841.51)	-22.34 (1235.65)	-67.81 (1056.09)
Hispanic %	32.72 (1551.92)	-43.63 (1007.53)	-326.89 (1024.68)	-21.89 (27.22)	-0.62 (11.76)	-41.16 (1552.75)	-1.26 (350.07)	76.05 (250.22)	89.07 (185.70)
White %	40.79 (1699.54)	-4.78 (553.18)	84.3 (1731.24)	0.13 (26.67)	0.13 (5.75)	-1.07 (318.50)	8.12 (48.62)	-27.56 (104.54)	-32.74 (54.62)
Female %	-66.56 (460.18)	-0.37 (22.72)	-9.39 (90.70)	-0.18 (1.51)	0 (0.57)	-0.31 (50.71)	0.53 (7.18)	0.86 (10.01)	0.89 (9.98)

Table A16, *continued*

School-Level Estimates	School-Level Dependent Variables								
	Student Misbehavior		Administrative Responses				Other Student Outcomes		
	Incidents	IILE	OSS	Transfers	Expulsions	Arrests	Truancy	Math PR	Reading PR
FRPL %	-1.01 (194.82)	-9.95 (433.94)	-31.34 (833.92)	-3.77 (12.75)	-0.33 (2.86)	-12.82 (154.86)	14.46 (244.85)	3.52 (68.43)	-14.94 (49.84)
Pupil/Teacher Ratio	-0.9 (17.52)	-8.68 (81.14)	6.53 (823.23)	-3.41 (20.70)	-0.47 (4.28)	-12.61 (97.33)	17.54 (78.58)	20.42 (87.60)	7.34 (65.97)
Primary School	3.37 (145.26)	0 (0.13)	0.07 (0.38)	0 (0.01)	0 (0.00)	0 (0.38)	0 (0.05)	-0.01 (0.03)	-0.01 (0.02)
Middle School	-2.53 (91.82)	-0.25 (4.98)	-3.79 (58.73)	-0.01 (0.70)	-0.03 (0.32)	-0.43 (39.54)	1.05 (10.38)	-0.62 (6.63)	-0.26 (8.68)
Enrollment	-0.01 (0.15)	0 (0.00)	0	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Adj. R^2	0.49	0.21	0.65	0.56	0.81	0.46	0.67	0.21	0.16
Observations	197	197	197	197	197	197	197	197	197
Schools	197	197	197	197	197	197	197	197	197

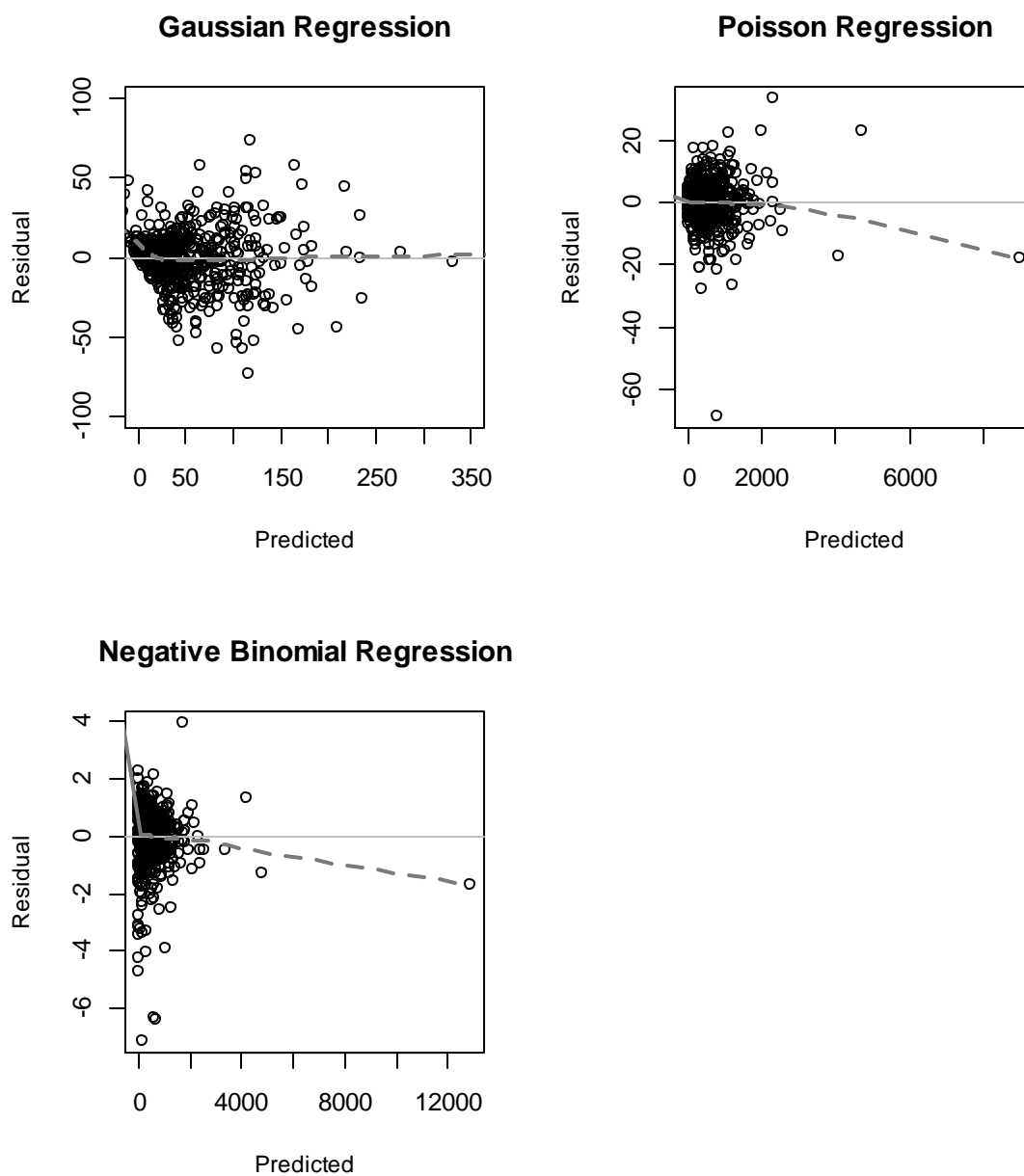
Notes. FRD = fuzzy regression discontinuity; IILE = incidents involving law enforcement; OSS = out-of-school suspensions; PR = proficiency rate; FRPL = free or reduced-price lunch. Standard errors in parentheses. Statistical significance indicated by: * $p < .10$; ** $p < .05$; *** $p < .01$.

APPENDIX B: ADDITIONAL FIGURES



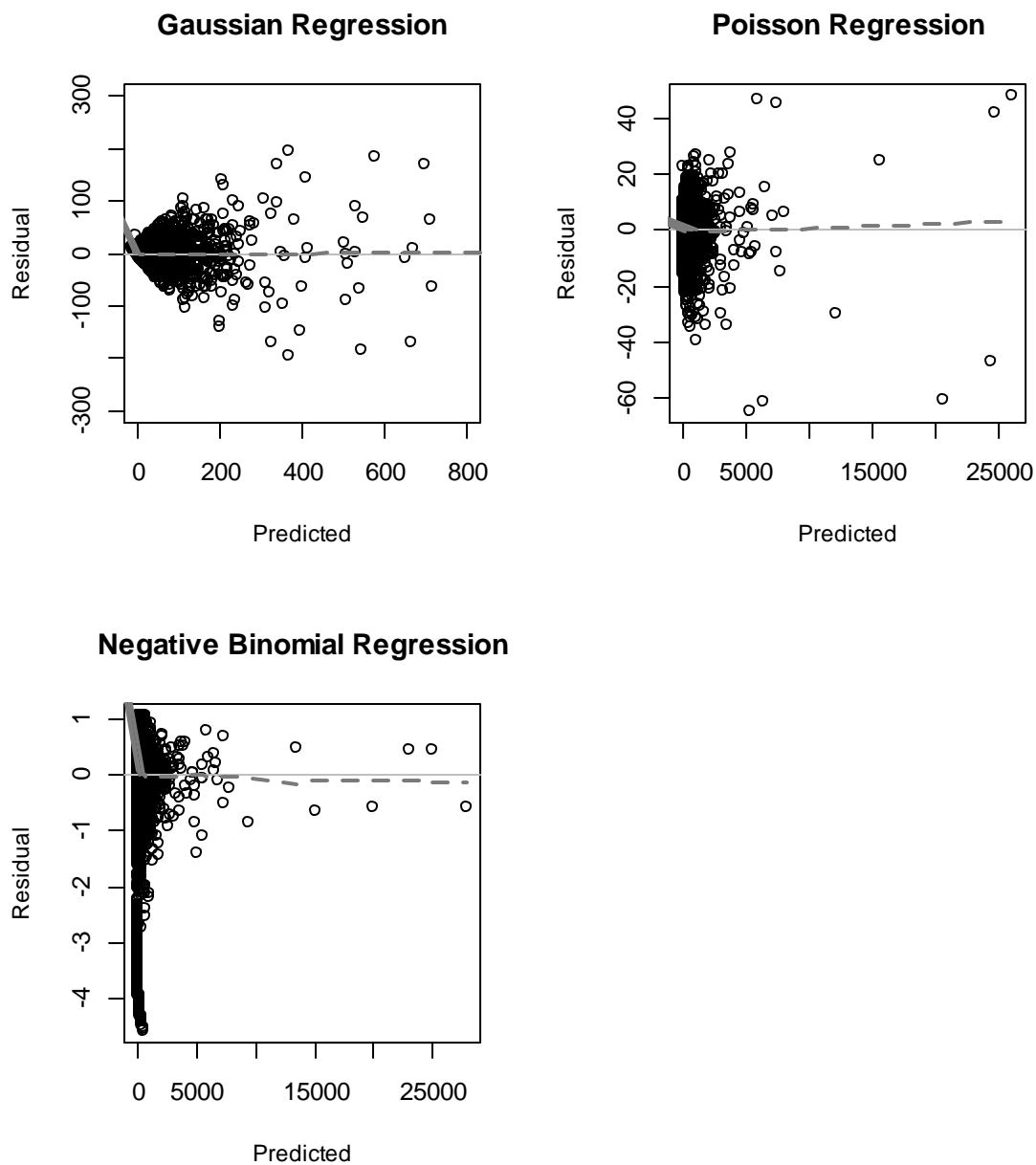
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B1. Predicted and Residual Values of RQ1 DD Analysis, Purposive Philadelphia Charter Control



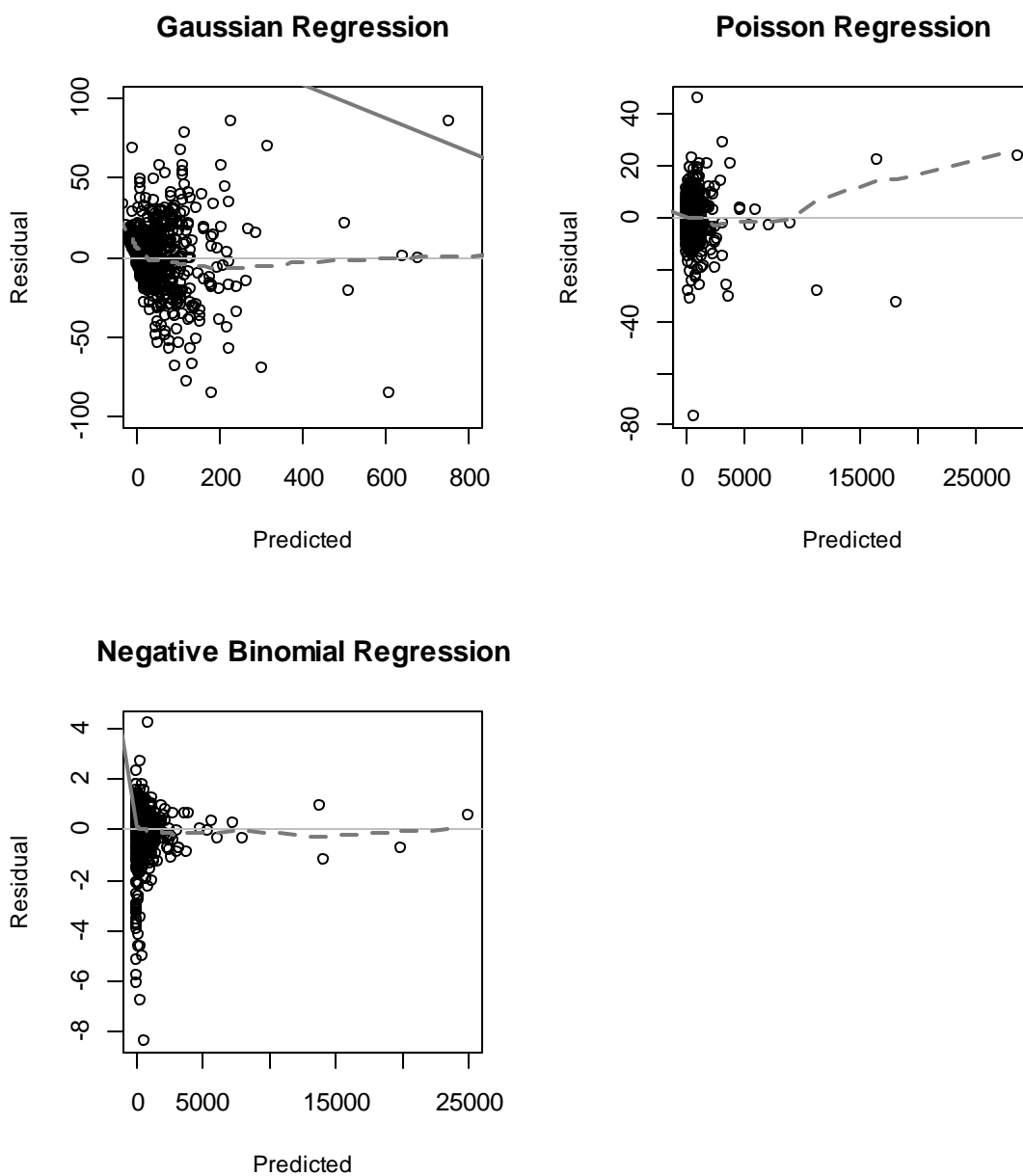
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B2. Predicted and Residual Values of RQ1 DD Analysis, PSM Philadelphia Charter Control



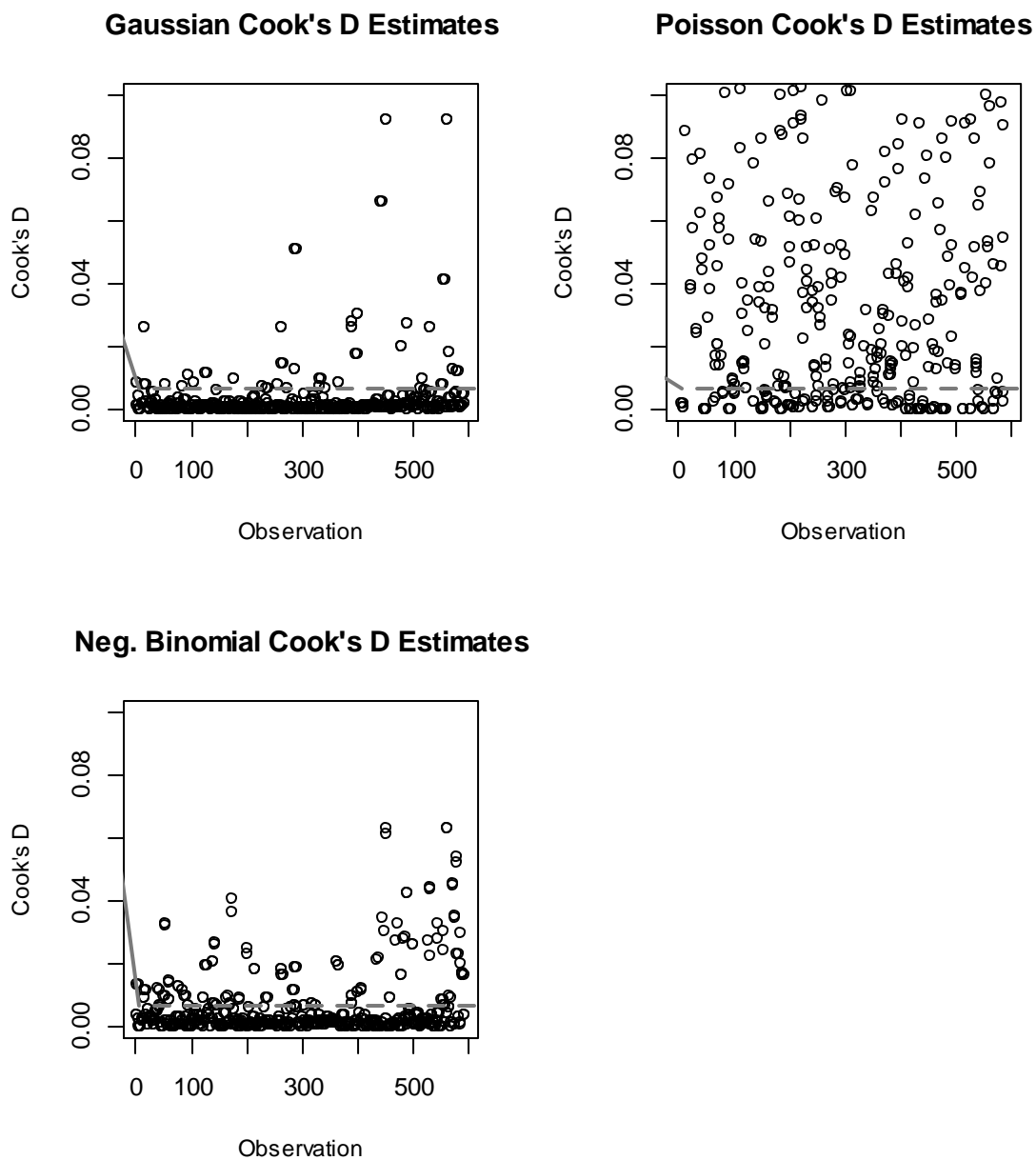
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B3. Predicted and Residual Values of RQ1 DD Analysis, Purposive Pennsylvania Schools Control



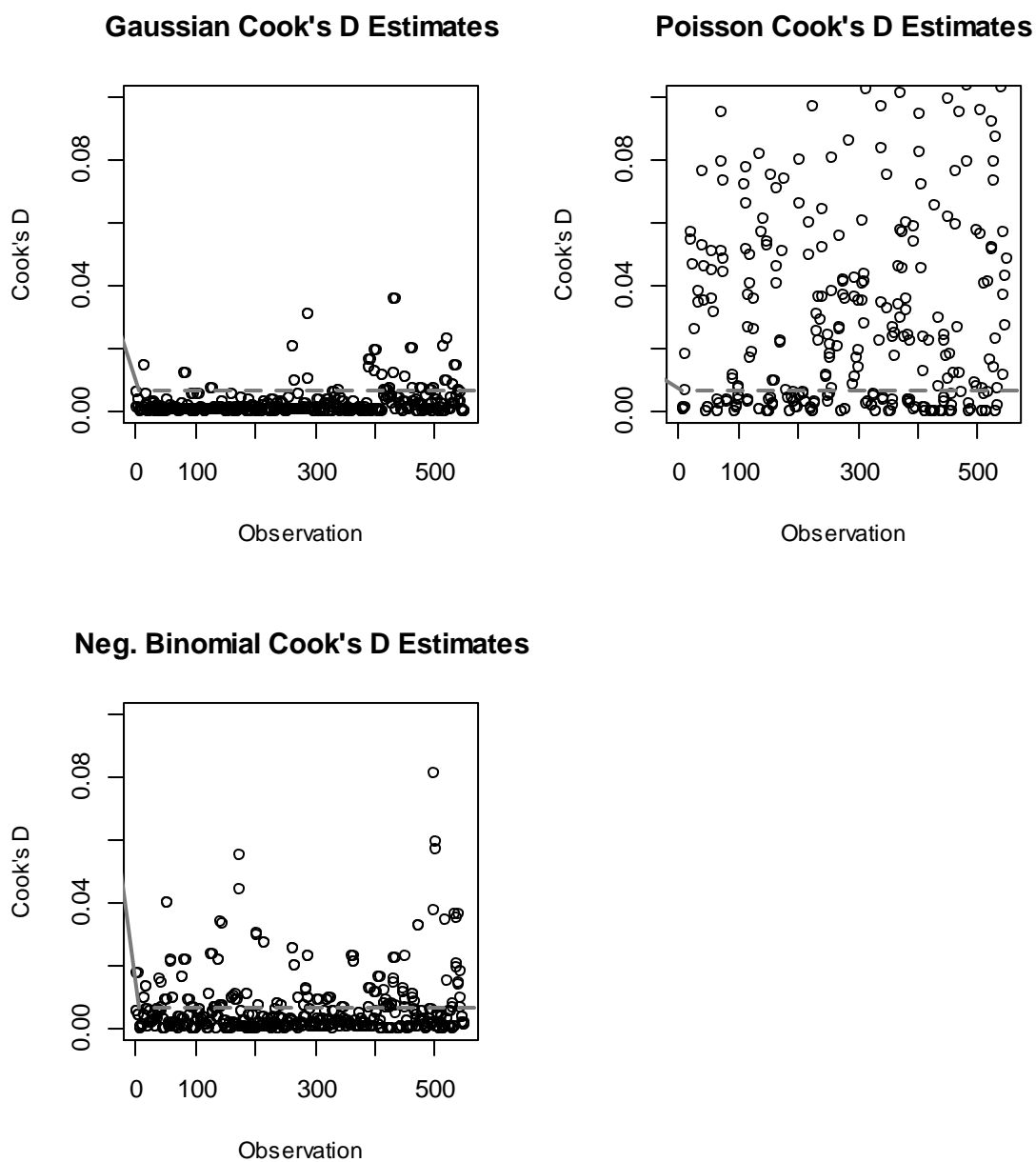
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B4. Predicted and Residual Values of RQ1 DD Analysis, PSM Pennsylvania Schools Control



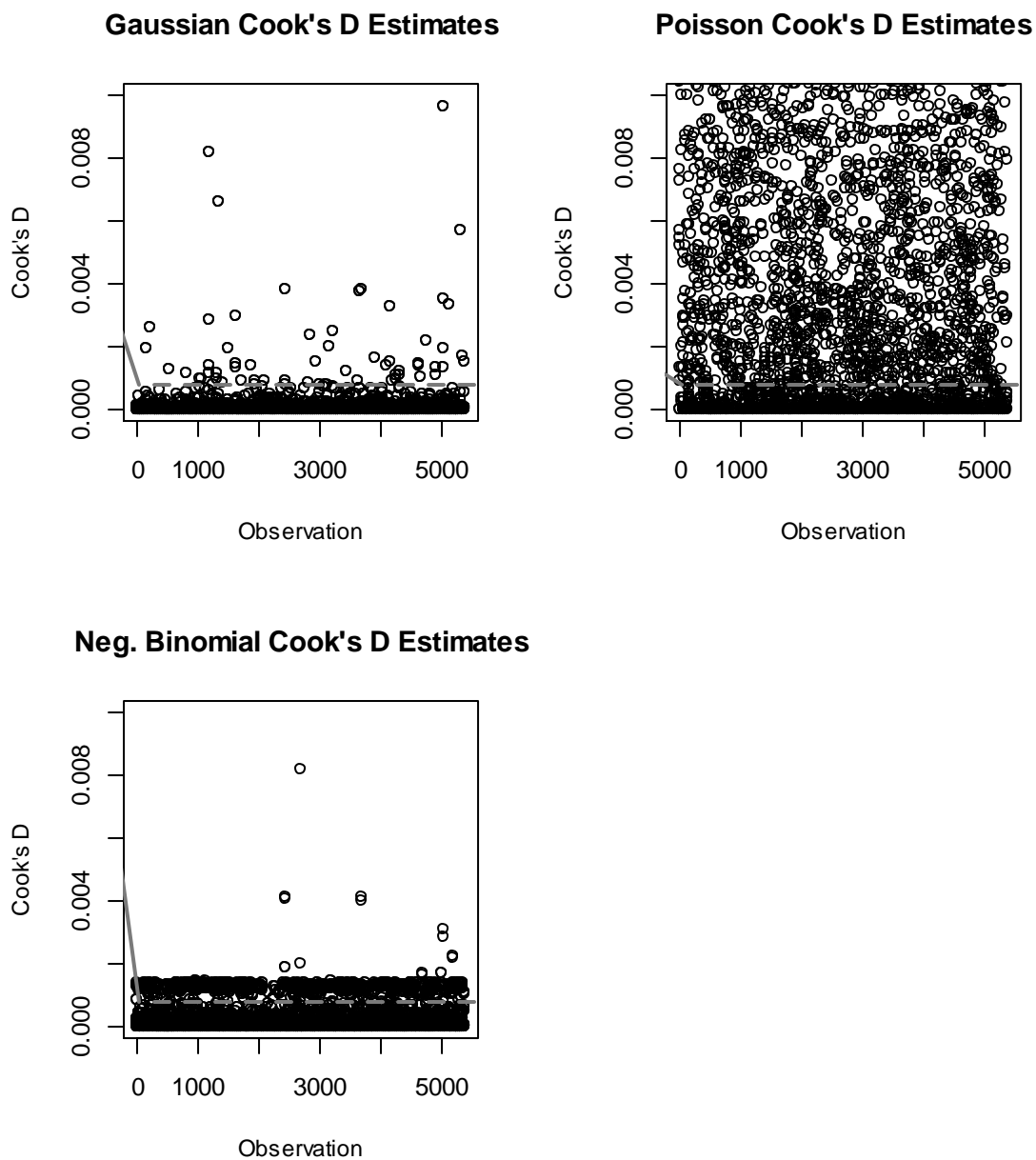
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B5. Cook's D Estimates of RQ1 DD Analysis, Purposive Philadelphia Charter Control



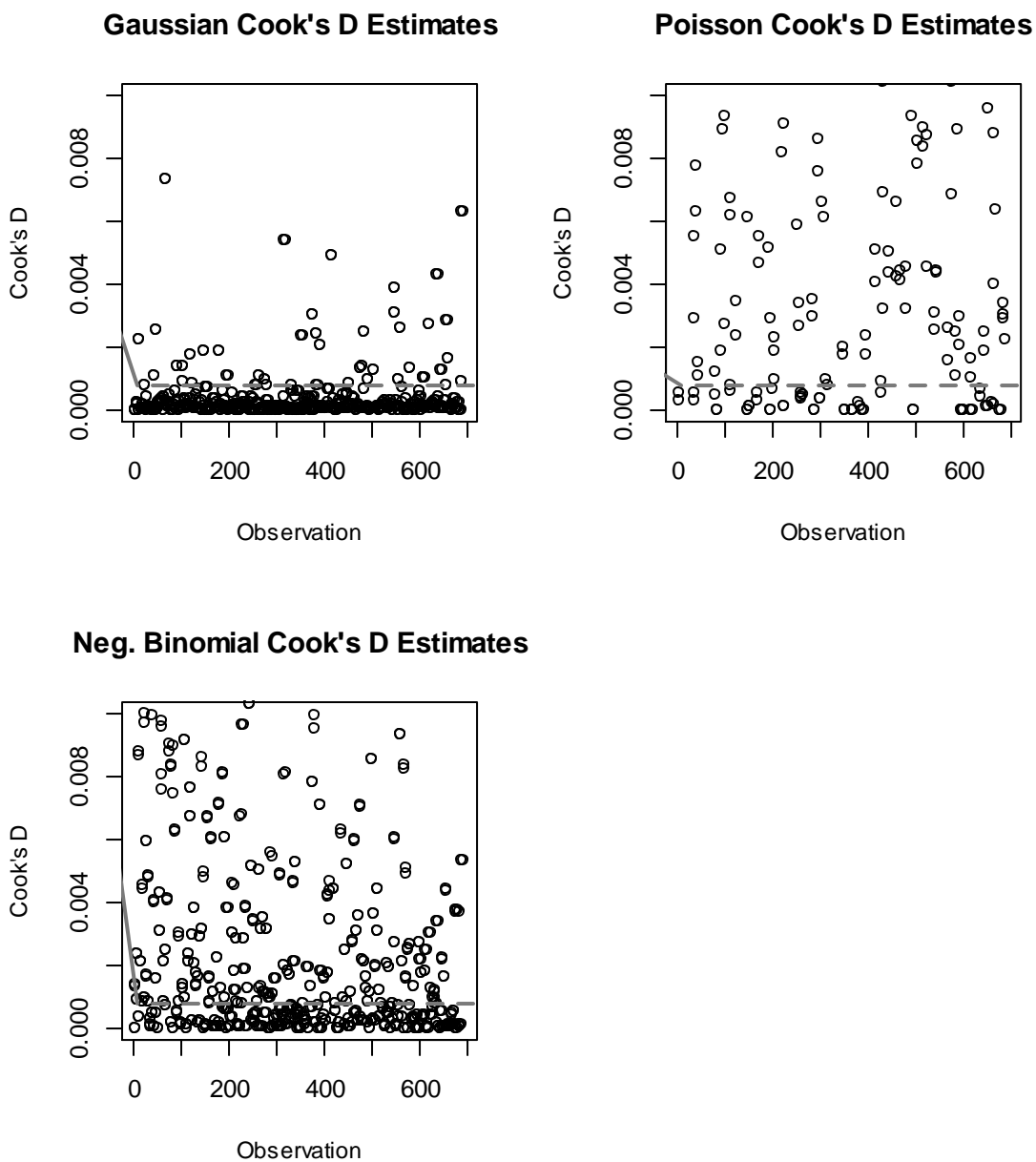
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B6. Cook's D Estimates of RQ1 DD Analysis, PSM Philadelphia Charter Control



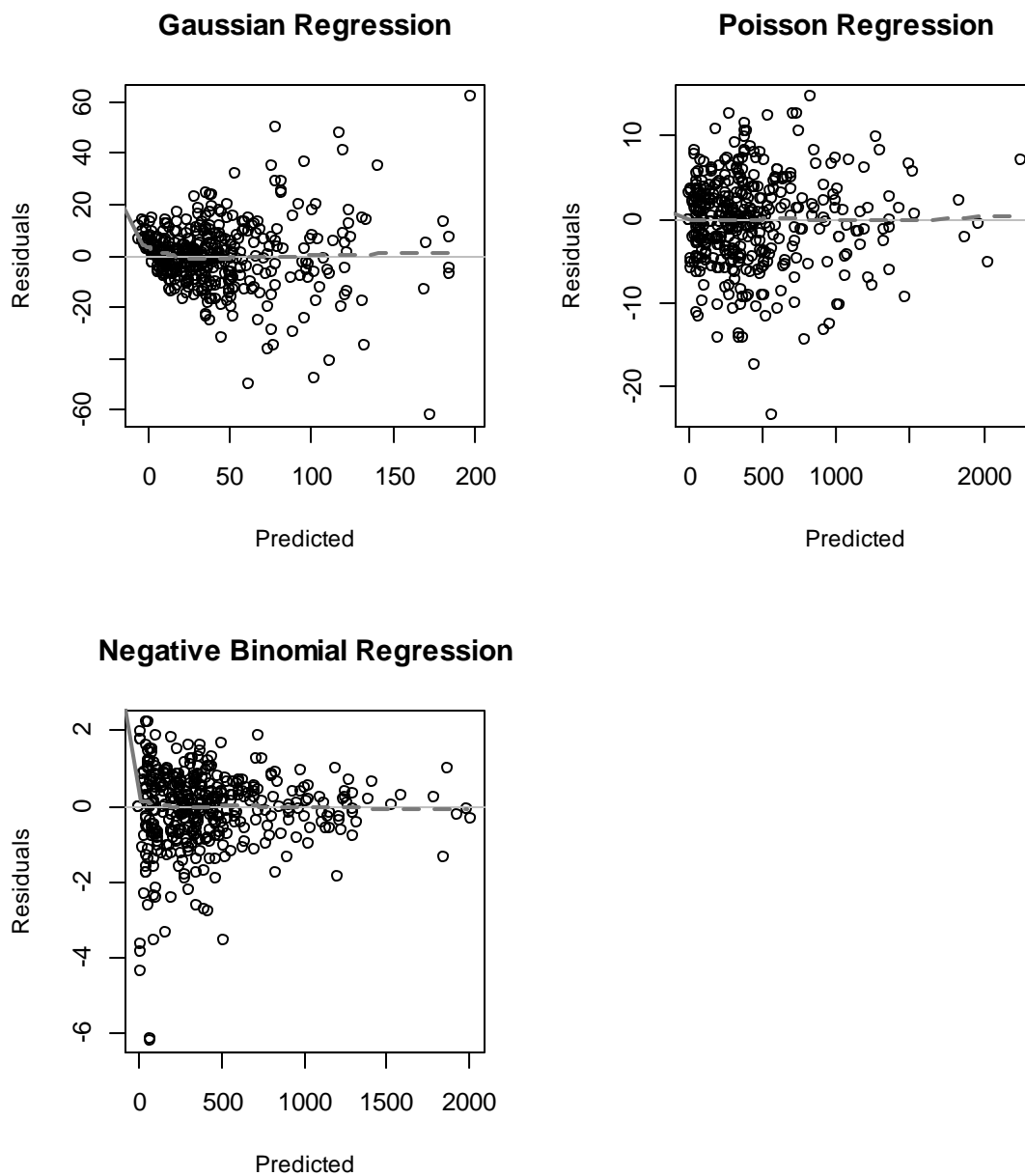
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B7. Cook's D Estimates of RQ1 DD Analysis, Purposive Pennsylvania Schools Control



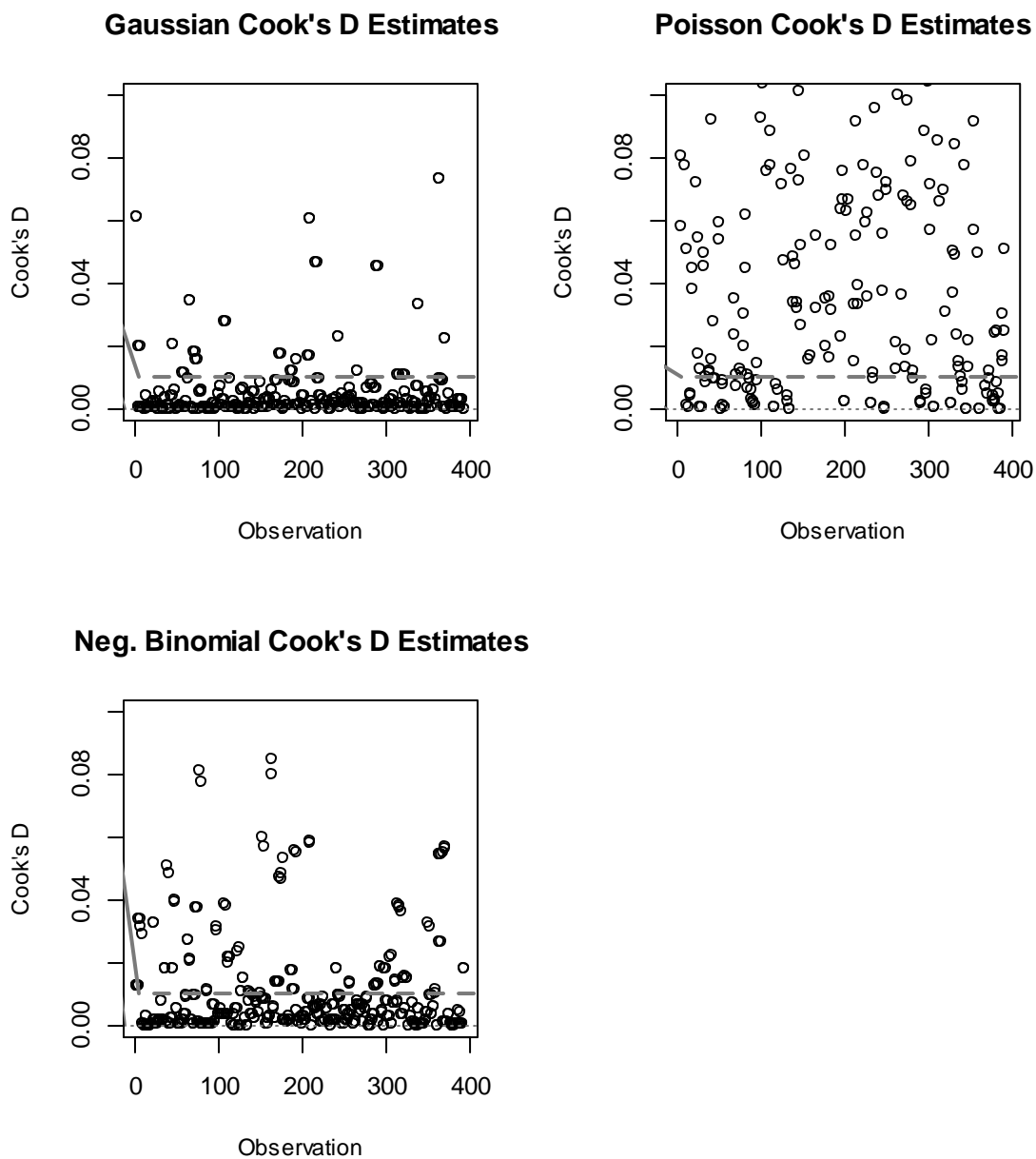
Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ1 = research question (1); DD = difference in differences.

Figure B8. Cook's D Estimates of RQ1 DD Analysis, PSM Pennsylvania Schools Control



Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ2 = research question (2); DD = difference in differences.

Figure B9. Predicted and Residual Values of RQ2 DD Analysis



Notes. Plots represent Gaussian, Poisson, and negative binomial difference-in-difference models regressing school-level incident rate onto predictor variables. RQ2 = research question (2); DD = difference in differences.

Figure B10. Cook's D Estimates of RQ2 DD Analysis

APPENDIX C: SDP DISCIPLINE POLICY, 2002-2003 CODE OF CONDUCT

Zero Tolerance Policy as per SDP 2002-2003 Code of Conduct, pages 10-19:

STUDENT MISCONDUCT

This section describes a broad range of misconduct that is prohibited in the school. Because the following sections listing acts of misconduct do not include all types of misconduct, the student who commits an act of misconduct not listed under the sections herein shall be subjected to the discretionary authority of the classroom teacher and the principal or designee.

All disciplinary actions for misconduct should include a conference between the teacher and/or principal or designee and the student, followed by notification to the parent(s) or guardian. Where schools have implemented School Peer Juries or Teen Courts approved by the Chief Education Officer or designee, a student may be given the option of appearing before a Peer Jury for determination of the disciplinary actions to be taken for Rules 1-5 only. The student, parent, or guardian who feels that the disciplinary action taken is unwarranted or excessive has the right to appeal to the principal.

The disciplinary process is intended to be instructional and corrective, not punitive. In addition to the disciplinary actions listed in this policy, schools may assign students to workshops in truancy/violence prevention or in conflict resolution. In lieu of in-school or out-of-school suspensions for acts of misconduct in Rules 1-8, a student may be referred to the Regional Office-approved school-based after-school or Saturday program, if such a program is available in the student's area. Students whose acts of misconduct most seriously disrupt the educational process may be subject to alternative school placement or expulsion.

The Uniform Discipline Code shall be followed and enforced in the same spirit and manner throughout the school system. In addition to the standards of conduct and disciplinary actions described in this Code, JROTC programs may enforce standards of conduct and disciplinary actions that are consistent with the military nature of those programs.

The range of actions is listed from the least severe to the most severe. Staff members shall consider all mitigating circumstances prior to disciplinary action and ensure due process for each student. Mitigating circumstances include, but are not limited to, the following factors:

- Age, health, maturity, and academic placement of a student
- Prior conduct
- Attitude of a student
- Cooperation of parent/guardian
- Willingness to make restitution
- Seriousness of offense
- Willingness of student to enroll in a student assistance program

The acts of misconduct on the following pages may also violate criminal laws, and the school for each violation must notify the Philadelphia Police Department. Whenever the school notifies the police concerning student misconduct, the school must also immediately attempt to contact the parent/guardian of that student. If questions occur concerning police notification, please consult the Law Department at (215) 299-7676.

APPEALS TO DISCIPLINARY ACTION

Whenever a student is subject to removal from his/her current school assignment or to expulsion from the School District the parent/guardian will be furnished with written materials describing the applicable due process and appeal rights that adhere to the proposed action.

RULES OF CONDUCT

General Statement

No student may engage in any conduct or encourage any other person to engage in conduct that jeopardizes the health, safety or welfare of any member of the school community, or in any conduct that disrupts or undermines the basic educational mission of the School District.

Note: Internal school disciplinary action shall not in any way deprive the School District or an individual victim from pursuing any legal remedies available in the Criminal, Civil or Delinquency Courts of the Commonwealth, or United States.

SCOPE OF RULES OF DISCIPLINE

These rules shall apply to any conduct:

- a) on school grounds during the school day or immediately before or after school hours;
- b) on school grounds at any other time when the school is being used by a school group;
- c) off school grounds at any school activity function or event;
- d) traveling to and from school, including actions on any school bus, van or public conveyance;
- e) off school grounds when the conduct may reasonably be expected to undermine

the proper disciplinary authority of the school, the safety of students or staff, or disruption within the school.

Level I OFFENSES

Rule 1: DISRUPTION OF SCHOOL

Students are expected to act in a courteous and peaceful manner toward staff, visitors, and each other. A student may not act in any way that disrupts or disturbs any educational or school-related program.

Violations of this rule include, but are not limited to,

1. Failure to obey directions from school staff
2. Disruption of a classroom or other school function by failing to behave in an appropriate, safe and respectful manner
3. Running and/or making excessive noise in the hall or building
4. Violation of the student dress code²⁺
5. Loitering and/or failing to provide school i.d.³⁺
6. Failing to attend class without a valid excuse⁴⁺
7. Persistent tardiness to school or class⁵⁺
8. Forgery
9. Fighting- two people, no injuries
10. Cheating and/or copying the work of another student or plagiarism from library or internet sources

Rule 2: DISRUPTIVE AND/OR OFFENSIVE USE OF LANGUAGE/GANG COLORS

Students are expected to communicate with staff, visitors and each other, as they themselves should expect to be treated.

1. A student shall not curse or use vulgar, obscene, intentionally disruptive or offensive language in any school activity.
2. A student shall not try to maliciously damage the reputation of anyone in the school community. This rule includes spoken and written messages including on any bulletin board, flyer or notice; or over the public address systems or computer networks, or displayed on a student's personal belongings.
3. Students shall not send or pass on any offensive, sexually oriented, or threatening messages, pictures or symbols from any source.
4. Students may not overtly display any indicia of gang affiliation.

Rule 3: DAMAGE, DESTRUCTION OR THEFT OF SCHOOL PROPERTY FALSE FIRE ALARM

Students are expected to respect school property. Students shall not recklessly or

intentionally cause or attempt to cause damage to, or deface school property, or steal or attempt to steal school property. Damage and defacement include graffiti, carving, tearing, cutting or otherwise marking school property.

False activation of a fire alarm seriously disrupts the educational process and is strictly prohibited.

Serious Damage, defacement or theft that interferes with the educational or safety rights of others will be treated as an aggravated offense under Rule 14.

Rule 4: DAMAGE, DESTRUCTION OR THEFT OF PRIVATE PROPERTY

Students are expected to respect the property of others. A student shall not try to steal, damage, deface or destroy the private property of others. This conduct is not allowed whether it is on school grounds, during a school activity, function or event off school grounds, or while traveling to or from school. Students may not harm or destroy data of another student or person, the Internet or other networks. This includes, but is not limited to, the creation downloading or uploading of computer viruses. Computer violations of a criminal nature are treated as an aggravated offense under Rule 14j.

Rule 5: RECKLESS ENDANGERMENT OR THREATS

Students are not only responsible for harmful actions they intend, but they are also responsible for the reasonably foreseeable consequences of their reckless actions. A student may not act in a manner which ignores the health, safety or welfare of any member of the school community by placing them in danger of injury, or pain. No student shall attempt, by physical menace, threat, or verbal taunt, to put a member of the school community in fear of injury, pain or social ridicule. The intentional posting on networks of the addresses or telephone numbers of fellow students or other members of the school community is a violation of this rule.

Threats of bodily injury or repeated acts of harassment or bullying are violations of Rule 10 a level II offense. Threats to bomb, kill, injure or use dangerous or deadly weapons will be treated as an aggravated offense under Rule 14.

Rule 6: TOBACCO PRODUCTS AND PARAPHERNALIA

Students are expected to help keep their schools a safe and healthy place. A student may not possess or use any tobacco product. Possession of cigarette lighters matches, rolling papers, pipes, or other such paraphernalia is also prohibited.

Rule 7: POSSESSION OF DRUGS OR ALCOHOL FOR PERSONAL CONSUMPTION

Students are expected to help each other to be healthy and responsible citizens. A student may not have, use or be under the influence of any unauthorized prescription or non-prescription medication, or any narcotic drug, hallucinogenic drug, steroid, growth hormone, amphetamine, barbiturate, marijuana, alcoholic beverage or intoxicant or look-alike substance of any kind.

Possession of a drug authorized by a medical prescription from a licensed physician **and** carried in the container supplied by a pharmacist, shall not be considered a violation of this rule where the supply carried by the student does not exceed the amount necessary for use during that school day. Non-prescription medications may only be possessed with a note from a parent or doctor. Non-prescription medications may only be possessed in the amount needed during the time in school for that day.

A second violation of this rule will be treated as a Level II offense which provides for transfer to another school, or transfer to an alternative school placement for substance abuse programming.

Distribution of any of the above substances, or possession under circumstances indicating an intent to distribute is an aggravated offense under Rule 14.

RULE 8: ABUSE OF COMPUTER/INTERNET PRIVILEGES

Students are expected to respect the computer privileges given to them. All students who have permission to use school computers must keep their passwords to themselves. It is against this rule to use another person's passwords or accounts. It is also against this rule to break into ("hack") other files or systems, to download copyrighted material, or to conduct a personal business enterprise using the school computer network. Students shall not go into any sites on the Internet, which contain sexually explicit material. Additional rules on computer use are listed in the School District's Acceptable Use Policy.

APPROVED SANCTIONS/CORRECTIVE ACTION - LEVEL I (Minimum to Maximum)

For any violation of a Level I offense, a student will face one or more the following corrective actions:

- a. meeting between the teacher and the student to discuss the student's behavior and expectations for improving his/her behavior;
- b. detention;
- c. meeting between the student and the counselor, small learning community coordinator;
- d. meeting between the student and the principal;
- e. notice to parents informing them of the student's behavior;
- f. temporary assignment to discipline or "cooling off" location;
- g. temporary assignment to a different class setting;

- h. supervised mediation between the students involved;
- i. placement on daily report;
- j. meeting with case worker or probation officer, where applicable and appropriate;
- k. suspension from privileges, including, but not limited to, graduation ceremonies, class trips, dances, other extracurricular activities such as athletic, club or non-instructional music programs participation), and suspension or revocation of network user accounts;
- l. temporary suspension of bus service for violations on, or while waiting for school bus or van;
- m. assignment to Saturday SMART Program.
- n. suspension for one to five days;
- o. permanent assignment to another class or section;
- p. Where indicated by the student's conduct or course of conduct, referral to CSP or SAP team;
- q. referral to counselor or substance abuse program where conduct indicates;
- r. referral to community M/H provider;

For first offenses of Rules 1 or 2 sanctions shall be limited to actions a. through f., above.

LEVEL II OFFENSES

Rule 9: REPEATED SCHOOL VIOLATIONS

A student shall not continue to break any of Rules 1 through 8. Nor may a student repeatedly fail to follow directions given by any school staff member while the student is under school supervision.

Rule 10: HARASSMENT

No student shall engage in verbal or physical activity, which he/she should reasonably expect to have the effect of harassing any student, staff member, or school visitor. Harassment, for the purpose of this rule, includes a course of conduct, or a single aggravated incident.

Harassment includes, among other things:

1. Unwelcome sexual advances, requests for sexual relations, sexual comments, sexually-oriented gestures, sounds, remarks or comments about a student, staff member or visitor's sexuality or sexual experience;
2. Offensive expressions concerning a person's race, sex, religion, disability or national origin,
3. Efforts to intimidate bully or ridicule. (See Rules 11 and 14 for offenses involving other serious forms of sexual or physical misconduct.)

Rule 11: INDECENT ASSAULT OR INDECENT EXPOSURE

All students must keep their hands and bodies to themselves at all times. No student may touch the sexual parts of other person, or encourage another person to touch him/her in sexual part of the body.

No student may show the sexual parts of his/her body to other persons in a way that would offend them or in an effort to excite them. (See Rule 14 d. for forcible or voluntary acts of sexual acts.)

Rule 12: ASSAULT ON SCHOOL PERSONNEL

A student shall not intentionally cause or attempt to cause physical injury or pain to any school employee. See Rule 5 for reckless conduct and Rule 14a for assaults that may potentially result in serious injury.

Rule 13: POSSESSION OF A WEAPON CAPABLE OF CAUSING DEATH OR SERIOUS INJURY^{6*}

Students shall not possess on their person, in their belongings, or in any storage space provided by the school, any tool, instrument, implement or weapon capable of causing serious injury or death. Such weapons, include, but are not limited to, firearms, knives, razors, stun guns, BB guns, starter pistols, harmful biological or toxic substances, explosives, fireworks with the potential to injure or devices which may cause a fire.

The use or attempted uses of a weapon or another object not traditionally viewed as a weapon in a manner which causes or risks serious injury will be treated as an aggravated offense under rule 14(a) or (f).

Rule 14: - AGGRAVATED OFFENSES

A student shall not engage in, or attempt to engage in, any conduct, which endangers the health, safety or welfare of any member of the school community, including but not limited to:

- a. Attacks on any staff member which results in injury, or places the person in danger of serious injury, or involves the use or attempted use of a weapon (including mace, pepper spray or laser pointer);
- b. Selling or distributing any of the substances set forth in Rule Seven (drugs, alcohol or look-alike substances) or possessing these items in sufficiently large quantities, or under circumstances which would indicate that they are not for personal use;
- c. Setting any fire that potentially risks injury to any person or damage to any property;
- d. Sexual acts: 1) forcible acts, as to the offender, or, 2) voluntary acts, as to each student;

- e. Rob, steal or threaten someone to give money or property;
- f. Assault on another student or other non-employee which results in serious injury or involves the use or attempted use of a weapon (including mace, pepper spray, or laser Pointer);
- g. Damage to school property which disrupts, impairs or prevents the school from carrying out any of its programs;
- h. Retaliation against a school employee, witness or hearing officer for their participation in any investigation, academic or disciplinary proceeding where the student's action takes the form of assault, threats of bodily injury or death, telephone harassment or stalking or substantial property damage.
- i. Threats relating to
 - a. the planting of bombs, or other explosive devices,
 - b. the use of any other weapon, including biological or toxic substances, or
 - c. to kill or seriously injure any member of the school community.
- j. The conduct of illegal activities via the District's network.

APPROVED CORRECTIVE ACTION LEVEL II VIOLATIONS

- A. Placement of the student in an alternative education such as CEP where the other required academic and attendance criteria are present.
- B. Placement in a School District alternative education program, including a remedial disciplinary program, or twilight program.
- C. In the case of violations of Rule 10 or Rule 11 in addition to A or B above, the counselor shall refer the student to an appropriate counseling program either within the School District or in a community mental health program. The counselor at the receiving school shall monitor the referral to assure the student's participation.
- D. Where a student has previously been assigned to a remedial disciplinary program or alternative school program, expulsion from the School District may be pursued.

Note: Level II offenses involving the possession of a firearm or explosive device, sale of drugs, possession of drugs not for personal use, forcible sexual congress, or assaults involving a weapon or serious injury shall be referred to the Central Discipline Office for the expulsion review process.

For other Level II offenses Region Office Leaders will determine whether a referral should be made to the Central Discipline Office for expulsion review. All referrals on Level II offenses shall include an EH-21, and copies of the Incident Follow-up Report (EH31), current grades and attendance, all prior disciplinary records (EH-20's) in the student's and a Behavior Performance Review. The principal may also file a statement of any aggravating or mitigating circumstances relevant to the student or the alleged incident.

APPENDIX D: SDP DISCIPLINE POLICY, 2012-2013 CODE OF CONDUCT

Revised Discipline Policy as per SDP 2012-2013 Code of Conduct page 4-9:

Purpose of the Code of Student Conduct

- To support the creation of a safe learning environment for all members of the school community.
- To provide clear and explicit expectations for social behaviors in all school settings.
- Provide administrators with in-school interventions that address students' disruptive behaviors.

District Policy Regarding Behavior

The safety and security of all school children is our highest priority, and the School District of Philadelphia is committed to providing a safe learning environment.

The Code of Student Conduct establishes policies, rules, and expectations for all school community members to learn, teach, and work together. These are common sense expectations built on respect for each other as our common core value. It follows, therefore, that students must refrain from violence against others, from bullying, harassment and carrying weapons.

Consequences for students who endanger school safety or disrupt the educational experience of others are listed in detail. The Code of Student Conduct applies during school and on the way to and from school, including, but not limited to, travel on School District vehicles, private transportation, and public transit systems.

Families and guardians are critical to our school community. We ask that they please read and understand the Code of Student Conduct and School Student Handbook, and discuss them with their child. We understand that families and guardians know best how to make sure that their child understands the expectations that will lead to a safe and orderly school community that is ready to learn.

This Code of Student Conduct provides definitions of behavioral offenses. Although some definitions include examples, the behaviors include, but are not limited to, the examples given.

Parents, guardians, and caregivers who have any questions or concerns are

encouraged to contact the school principal.

Behavioral Expectations & Responsibilities

Responsibilities of Everyone

- Respect all members of the school community.
- Maintain a positive school climate by being responsible, respectful, and cooperative.
- Communicate Code of Student Conduct expectations for students and staff.
- Motivate students to live up to the expectations through positive reinforcement.
- Use good judgment to prevent minor incidents from becoming major problems.

Responsibilities of Administrators

- Respect all members of the school community.
- Implement the Code of Student Conduct and all disciplinary procedures in a fair and consistent manner.
- Provide students and parents whose first language is not English with translation and interpretation services free of charge.
- Inform all school personnel, parents, and students of discipline policies.
- Review and act upon allegations and requests from school personnel concerning violations.
- Use professional judgment to prevent minor incidents from becoming major problems.

Responsibilities of Teachers

- Respect all members of the school community.
- Teach proper behavior and positively reinforce rules of conduct.
- Maintain a learning environment that provides for academic success.
- Hold students accountable for disorderly conduct in school and on school grounds.
- Address rule violations with multiple strategies.
- Use professional judgment to prevent minor incidents from becoming major problems.

Responsibilities of Students

- Respect all members of the school community.
- Understand and comply with school rules and climate expectations, including the Code of Student Conduct and School Student Handbook.
- Comply with the School District's attendance, dress code, unlawful harassment, and bullying policies.
- Behave in a manner that focuses on academic success.
- Be responsible and accountable for following rules.

Responsibilities of Parents / Guardians, and Advocates

- Respect all members of the school community.
- Respect, understand, and support school rules and regulations.
- Respect, understand, and support the policies of the School District of Philadelphia.
- Recognize and understand that school personnel must enforce school rules.
- Teach children to respect the rights of others and follow school rules.
- Emphasize the importance of being prepared for school and adhering to school rules to foster academic success.

Dress Code Expectations

Students are expected to follow their school's dress code so that their appearance does not cause a disturbance, distract or interfere with the instructional program, or constitute a health or safety hazard. A student's uniform should fit appropriately. School principals must determine and communicate the final decision regarding what is considered proper or improper attire. For information regarding a school's dress code, please consult the School Student Handbook.

Principals may declare spirit or club days and allow students to wear school spirit shirts, or schedule dress-up days (i.e., when school pictures are scheduled) or allow students to wear other dress code attire such as uniforms or special dress for Boy Scouts, Girl Scouts, athletes, cheerleaders, band, or chorus. Nothing in this Code of Student Conduct restricts or bans a student's wearing religious garb. Parents with concerns regarding religious dress are encouraged to discuss them with the principal.

Examples of Appropriate Dress

- **Shirt:** Solid-colored golf or polo style collared, uniform-type, short-sleeve shirts in any color. Shirts are to be worn tucked in at all times. Turtlenecks acceptable in winter.
- **Pants:** Uniform-style, in any color with belt loops. Slacks or shorts must fit at the waist and must be free from frays and holes. Belts must be worn at all times with slacks and shorts, and must be visible at the waist. Undergarments should not be visible at any time.
- **Appropriate Length:** Pants, shorts, and skirts, that measure (front and back) no shorter than three inches above the knee. Skirts, jumpers, and dresses must measure no shorter than three inches above the knee.
- **Shoes:** Students must wear shoes while on school premises.
- **Socks:** Socks (or stockings for girls) must cover the ankle and be visible.
- **Outerwear:** Sweaters, sweater vests, sweatshirts, or light jacket without hoods are acceptable in the classroom. During class time, jackets are to remain open—not zipped or buttoned.
- **Coats:** Heavy coats and jackets may be worn to and from school and/or outdoors.
- **Special:** Spirit shirts/club shirts may be worn on day(s) determined by the school administrator.

Additional questions concerning the dress code should be referred first to the principal or his/her designee.

Attendance Requirement

School Attendance Requirement

In Philadelphia County, school attendance is required for all children aged 6 to 17. “Attend school” means that a child must be enrolled and attend a public school, a charter school, a cyber charter school, or a private or religious-based school, or else participate in an approved home schooling program.

Pennsylvania Department of Education regulations state that children are considered school age from the time they are ADMITTED to the public school educational program until graduation from high school or reach the age of 21. During the time a child is of school age, he/she is entitled to attend the public schools in the School District of Philadelphia or a charter school. Students who turn 21 during the school term are entitled to finish out the school year.

Parents/guardians are expected to:

- Ensure that their children between the ages of 6 and 17 are enrolled in school and attend school regularly, on time, and for the entire school day.
- Emphasize the importance of on-time attendance in school, class, and supervised activities—and celebrate good attendance and success.
- Send their child to school every day prepared to participate and learn by providing a good night's sleep and breakfast.
- Establish reasonable, age-appropriate curfews and bedtimes.
- Make personal appointments for their child outside of the school day or during school breaks as often as possible.
- Ensure that their child receives the periodic student health examinations that are required by law.
- Schedule family vacations to coincide with school recesses.
- Call the school when their child is absent.
- Provide a written excuse for every absence when your child returns to school.
- Provide a written excuse for every late arrival and early departure.
- Provide the school with provide correct current addresses, emergency contacts, home, cell and work telephone numbers, e-mail addresses to the school at the beginning of each school year and update information whenever there are changes.
- Help develop and implement the individualized plan for their child’s improved attendance when necessary.

When Can a Child Be Absent?

Sometimes students have to miss school. These “excused absences” include

circumstances such as illness or injury, teen parent leave (delivery only), death/funeral-related absence (see below), education–related trips or activities, suspension, and religious holidays. A written notice from a parent or guardian must be submitted to the school immediately upon a student’s return to school. Such notice must include a valid telephone number or other means of contact for verification purposes. Certification of illness/injury/delivery is required if the absence extends for three (3) or more consecutive days.

An “unexcused” or “illegal absence” occurs when a student is absent without a valid excuse in writing. That means that either no written notice was submitted to the school upon the student’s return or that the reason provided in the notice was deemed invalid. Examples of invalid excuses include (but are not limited to) babysitting, waking up late, or being on vacation with family.

Responses to Non-Attendance (Truancy)

A student who is absent without a valid excuse is considered “truant” from school. A pattern of truancy can lead to a variety of interventions or penalties, depending upon the circumstances. On the school level, truancy can result in school-based interventions (home visits, truancy elimination plans, academic supports, social service referrals, etc.) or truancy court. Suspension from school or transfer to an alternative education setting is not a permissible response to truancy.

According to the Compulsory School Attendance Law, a judge or hearing officer may require students and/or his or her parent/guardian to do any or all of the following:

- Appear at a hearing regarding the student’s truancy.
- Perform reasonable school or community service for a period designated by the hearing officer.
- Complete a parenting education program.
- Obtain counseling or other supportive services, including a re-entry, or other service plan determined appropriate by authorized school officials.
- Pay a fine of up to \$300.
- Pay court costs.
- Serve up to five (5) days in county jail.

Children who are truant may be declared dependents of the state as part of the adjudication process.

Visitors to Campus

Parents and guardians are encouraged to visit schools and meet with teachers and administrators about their child’s progress. Visitors to campus must present themselves to the school office or have standing permission from the principal. To provide safe and orderly learning environments, each visitor to public school buildings or grounds must

receive authorization of the school principal or his/her designee.

Transportation Conduct

Students must comply with school rules while traveling to and from school or any school-related activity, whether on foot or by a means or modes of transportation including, but not limited to, School District vehicles, private transportation, or public transit systems. All responses to infractions, including suspension, hold for misconduct that occurs on the way to and from students' homes or school-related activities.

Referrals to the Philadelphia Police Department

Some infractions of the Code of Student Conduct, listed here on the School District's website, are classified as crimes by the city and state. Crimes must be reported to the Philadelphia Police Department by law. (You can view the Memorandum of Understanding by visiting the District's website at <http://www.philasd.org>)

In-School Intervention

Principals, teachers, and other school professionals hold high expectations for students' learning and behavior. Though there can be no excuse for behavior that harms or disrupts, there may be reasons that caring adults in school need to understand. We educate the whole child. Therefore, adults in schools will use a range of responses to help students' reconnect to their school community, and learn greater self-discipline, including:

- Get a student's attention by calling his/her name in a calm voice.
- Address the student in private.
- State the problem behavior.
- State expected behavior and why the student needs to follow the expected behavior.
- Listen to the student's response and capture opportunity to help student problem-solve new behaviors.
- Indicate the specific consequence for continuing to engage in the problem behavior—and the positive consequences of change.
- Ask student to demonstrate the expected behavior.
- Acknowledge the student's listening to you and/or exhibiting the appropriate behavior (e.g. thanks, praise, reward).
- Contact the parent/guardian to set up a conference to describe the student's problem behaviors, listen to parent/guardian, and discuss possible disciplinary consequences.
- Refer the student to appropriate intervention (e.g., individual or group counseling, or school-based programs).
- Implement a Response to Intervention (RTI) Tier II action, for continued minor infractions.
- Authorize in-school suspension when all other interventions are unproductive.

- Document the infraction on discipline referral report (EH20-pink slip).
- Behavioral Contracts.

Note: Use out-of-school suspensions as a last resort and only when in-school interventions and consequences are insufficient to address a student's inappropriate behavior.

Due Process

All students are to be treated fairly and honestly. Students have a right to be heard and tell administrators their versions of events.

If you believe your child is a victim you are encouraged to speak with the school's administrator.

If you feel you need victim services, please contact the Safe Schools Advocate at (<http://www.phillyossa.com>) or call 215-656-5381.

Student Conference Procedures

At formal conferences, students have the right to expect that school officials will:

1. Inform the student of the reason for a conference.
2. Give the student an opportunity to respond to allegation(s).
3. Discuss student's problem behavior and ways to correct it.
4. Inform the student of the corrective action and/or next steps to be taken.
5. Document the problem behavior and intervention.

Parent Conference Procedures

At formal conferences, parents, guardians, and caregivers can expect that school administrators will see to it that the following protocol is followed:

1. Notice of a conference must be provided to the parent/guardian in writing and either hand-delivered to the home, sent by certified mail, faxed, emailed, or communicated by other reasonable means.
2. When a student is suspended, a parent/guardian conference must be held within three (3) days of the suspension.
3. At the conference, the parent/guardian or caregiver may request to review the student's records and any witness statements.
4. School administrators will discuss the student's problem behavior and ways to correct it.
5. School administrators will inform the parent/guardian of any further disciplinary action.

Reproduced SDP Revised Discipline Policy

Disruptive Behaviors	Rule	Interventions*				
		Level 1: In-School Interventions (including in- school suspension)	Level 2: Out-of-School Suspension	Level 3: Contract w/Intervention or Lateral Transfer	Level 4: Disciplinary School Assignment	Level 5: Disciplinary School Assignment w/ Expulsion Referral
		Levels 3, 4, or 5 Interventions require a referral to the Office of Hearings and Expulsions				
Failure to follow classroom rules/disruption	1A	✓				
Dress code violation	1B	✓				
Failure to carry hall-pass/appropriate ID	1C	✓				
Failure to participate in class/unpreparedness	1D	✓				
Truancy/excessive tardiness/cutting class	2	✓				
Possession of beepers/pagers/cell phones/other devices	3A	✓				
Possession of other inappropriate personal items	3B	✓				
Profane/obscene language or gestures	4	✓				
Public display of affection/inappropriate touching	5A	✓	✓			
Inappropriate use of electronic devices	5B	✓	✓			
Fighting (two students engaged in mutual combat)	6	✓	✓			
Forgery of administrator, teacher, or parent's/guardian's signature	7	✓	✓			
Alteration of grade reporting/excuses/school documents	8	✓	✓	✓		
Destruction and/or theft of property (less than \$1,000)	9	✓	✓	✓		
Harassment/bullying/cyber-bullying/intimidation	10	✓	✓	✓	✓	
Sexual acts (consensual)	11		✓	✓		

* Intervention Levels are not mutually exclusive. For example, if a student is referred to the Office of Student Discipline, Hearings and Expulsions, that student will also be suspended out of school for some period between 1 – 10 days.

Reproduced SDP Revised Discipline Policy, *continued*

	Rule	Interventions*				
		Level 1: In-School Interventions (including in- school suspension)	Level 2: Out-of-School Suspension	Level 3: Contract w/Intervention or Lateral Transfer	Level 4: Disciplinary School Assignment	Level 5: Disciplinary School Assignment w/ Expulsion Referral
Disruptive Behaviors						
Threatening students/staff with aggravated assault	12		✓	✓	✓	
Destruction and/or theft of property (totaling \$1,000 or more)	13		✓	✓	✓	
Breaking and entering school property	14		✓	✓	✓	
Robbery	15A		✓	✓	✓	
Extortion	15B		✓	✓	✓	
Mutual fight (w/ documented serious bodily injury)	16		✓	✓	✓	
Simple assault (documented unprovoked attack by one student on another)	17		✓	✓	✓	
Possession of alcohol or drugs – intent to use	18A		✓	✓	✓	
Possession of alcohol or drugs - intent to distribute	18B		✓	✓	✓	✓
Possession or use of fireworks/incendiary devices/explosives	19		✓	✓	✓	✓
Assault of school personnel	20		✓	✓	✓	✓
Instigation or participation in group assaults	21		✓	✓	✓	✓
Aggravated assault (documented serious bodily injury)	22				✓	✓
Sexual acts (non-consensual)	23				✓	✓
Possession of a weapon	24				✓	✓

* Intervention Levels are not mutually exclusive. For example, if a student is referred to the Office of Student Discipline, Hearings and Expulsions, that student will also be suspended out of school for some period between 1 – 10 days.

Suspension Procedures

According to Pennsylvania law, suspension is defined as the denial to a student of the right to attend school and to take part in any school function for any period of up to 10 days. Suspensions must be preceded by notification to the student and parent/guardian in writing. For the purposes of the School District of Philadelphia's Code of Student Conduct, suspensions comprise three categories: In-school, short-term, and long-term.

In-School Suspensions

In-school Suspension is an exclusion from a classroom for disciplinary purposes that allow a student to remain under the direct supervision of school personnel. Direct supervision means school personnel are physically in the same location as students under their supervision.

Out-of-School Suspensions

Short-term Suspension is an exclusion from school and/or any school activity or function for a period of three (3) or fewer school days. Students who are suspended must meet with the principal or his/her designee prior to the suspension. During the conference, the student has the right to the procedures described under Student Conference Procedures outlined in this Code of Student Conduct.

Long-term Suspension is an exclusion from school and/or any school activity or function for a period of four (4) to 10 school days. Students who are given a long-term suspension are to participate in a student conference and a parent/guardian conference. During the parent conference, the student and parent/guardian have the right to the procedures described under Parent/Guardian Conference Procedures section in the Code of Student Conduct. In addition, prior approval by the Assistant Superintendent is required for long-term suspension. In addition, prior approval by the Deputy Chief of the Office of Student Discipline, Hearings and Expulsions is required for all long-term suspensions consisting of 10 school days.

All referrals to the Office of Student Discipline, Hearings and Expulsions for an EH21 discipline hearing, must be preceded by an out of school suspension.

Transition Procedures

Students with a pattern of disruptive behavior and/or serious violations of the Code of Student Conduct may be referred to the Office of Student Discipline, Hearings, and Expulsions. Students who are referred receive full due process, including a hearing to determine whether the student will be transferred to a disciplinary school. These schools provide high-quality alternative educational programs and more support to help students who've had behavior problems elsewhere achieve their intellectual and social potential.

Students referred for a hearing will be suspended and provided the procedures indicated

in the Suspension Procedures section of the Code of Student Conduct. The school must complete a Behavior Performance Review (BPR) for regular education students or a Manifestation Determination for special education students and share the results of the review at the parent/guardian conference. **When a hearing is not scheduled during the time of a student's suspension, students have the right to return to their school pending the outcome of the hearing.**

Students/Parents/Guardians have the right to a hearing to be conducted by an independent and impartial hearing officer. During the hearing process, parents/guardians may oppose the request for their student's removal from the school. At a hearing, students and parents/guardians have the right to:

- Present witnesses;
- Present evidence relevant to the alleged infraction
- Request and ask questions of school district personnel; and
- Review student records in advance

Note: Please refer to the Discipline for Students with Disabilities section for discipline protocols for students with disabilities.

Expulsion Procedures

According to Pennsylvania law, expulsion is defined as an exclusion from school and any school activities for more than 10 school days. Students who have committed an offense subject to expulsion will also be suspended and have the right to the procedures found in the Suspension Procedures section of the Code of Student Conduct.

In addition, students will receive a formal Expulsion Hearing conducted by an independent hearing officer appointed by the School Reform Commission (SRC).

Counsel may represent the student. In addition, the student has the right to:

- Present witnesses and evidence;
- Request and cross-examine District witnesses; and
- Review student records in advance; and
- Review evidence in advance

Readmission of Permanently Expelled Students

A student who has been permanently expelled from the School District of Philadelphia by the affirmative vote of a majority of the School Reform Commission (SRC) may apply for readmission. If a student is expelled by the SRC, the parent /guardian of the student has 30 days to provide written documentation that the student is in another educational program. If they are unable to find an alternative educational program, the School District of Philadelphia will provide for the student's education. The SRC delegates complete authority for all readmission decisions to the Chief Executive Officer/Superintendent or his/her designee. Temporarily expelled students need not apply for readmission because they are automatically readmitted to the School District of Philadelphia at the end of the

expulsion period. Readmission decisions are final and not subject to review by appeal to the SRC or the courts.

Expungement of Records for Expelled Students

A student who has been permanently or temporarily expelled by the affirmative vote of a majority of the SRC, beginning in school year 2008-2009, may apply to the School District of Philadelphia to have his/her records of expulsion erased from the students' records, or expunged.

Expungement occurs at the discretion of the Superintendent or his/her designee, and is not subject to review or appeal to the SRC or the courts.

The complete SRC Board Policy No. 233, including the criteria and process for readmission and or impingement of records, may be found on The School District of Philadelphia's website at: <http://www.philasd.org/offices/administration/policies>.

APPENDIX E: CODEBOOK AND VARIABLE DEFINITIONS

All variables included in the analyses in this paper are described below in Table A1. Full definitions of the variables collected by the Safe Schools Report are provided next, as per the Pennsylvania Information Management Volume 2 User Manual (Pennsylvania Department of Education, 2014).

Variable	Type	Description/Distribution	Definition (primary original dataset)
PA_ID	Character	4-Digit Identifier	PA School Identifier (all datasets)
Schoolyear	Character	Levels: 2009-2010 2010-2011 2011-2012 2012-2013 2013-2014	School year of data (all datasets)
Sch_Name	Character	Full School Name	Name of school (NCES)
LEA	Character	Full LEA/District Name	Name of school district/LEA (NCES)
County	Character	Full County Name	Name of school county (NCES)
Type	Character	Levels: 1-Regular school 2- Special education school 3-Vocational school 4-Alternative/other school	Indicator of school type (NCES)
Charter	Character	Levels: 1-Yes 2-No	Indicator of charter school (NCES)
Level	Character	Levels: 1-Primary 2-Middle 3-High 4-Other	Level of school enrollment (NCES)
Enrollment	Numeric	Min=31 Mean=576 Max=10559	Total school enrollment (NCES)
FRPL	Numeric	Min=0 Mean=232 Max=6674	Number of students receiving free or reduced-price lunch (NCES)
Female	Numeric	Min=0 Mean=280 Max=5721	Number of female students (NCES)
AmerInd	Numeric	Min=0 Mean=1 Max=56	Number of American Indian/Alaska Native students (NCES)
Asian	Numeric	Min=0 Mean=19 Max=746	Number of Asian or Asian/Pacific Islander students (NCES)
Hispanic	Numeric	Min=0 Mean=48 Max=2466	Number of Hispanic students (NCES)
Black	Numeric	Min=0 Mean=88 Max=2664	Number of Black students (NCES)
White	Numeric	Min=0 Mean=409 Max=8830	Number of White students (NCES)

Variable	Type	Description/Distribution	Definition (primary original dataset)
Hawaiian	Numeric	Min=0 Mean=0 Max=100	Number of Hawaiian Nat./Pacific Isl. students (NCES)
Two_Races	Numeric	Min=0 Mean=10 Max=475	Number of students of two or more races (NCES)
PTR	Numeric	Min=0 Mean=15 Max=333	Pupil/teacher ratio (NCES)
Incidents	Numeric	Min=0 Mean=21 Max=1181	Number of incidents (SSR) An incident is defined as a specific act or offense involving one or more victims and one or more offenders. A reportable incident includes one or more acts of misconduct, involving one or more offenders violating criteria defined under Pennsylvania's Act 26 of 1995. These include but are not limited to any behavior that violates a school's educational mission or climate of respect or jeopardizes the intent of the school to be free of aggression against persons or property, drugs, weapons, disruptions, and disorder.
Offenders	Numeric	Min=0 Mean=18 Max=514	Number of students committing incidents as defined above (SSR)
Inc_wLaw	Numeric	Min=0 Mean=5 Max=160	Number of incidents as defined above involving law enforcement (SSR)
Arrests	Numeric	Min=0 Mean=2 Max=102	Number of arrests that resulted from incidents as defined above (SSR)
AltEd_Assign	Numeric	Min=0 Mean=1 Max=234	Number of assignments to alternative or disciplinary education schools that resulted from incidents as defined above (SSR)
Stud_SimpAss	Numeric	Min=0 Mean=2 Max=160	Number of simple assaults on a student (SSR)
Stud_AggAss	Numeric	Min=0 Mean=0 Max=41	Number of aggressive assaults on a student (SSR)
Staff_SimpAss	Numeric	Min=0 Mean=0 Max=96	Number of simple assaults on staff members (SSR)

Variable	Type	Description/Distribution	Definition (primary original dataset)
Staff_AggAss	Numeric	Min=0 Mean=0 Max=56	Number of aggressive assaults on staff members (SSR)
RaceEth_Intimidation	Numeric	Min=0 Mean=0 Max=26	Number of incidents of racial/ethnic intimidation (SSR)
OthHarass_Intimidation	Numeric	Min=0 Mean=1 Max=138	Number of incidents of other harassment or intimidation (SSR)
Fighting	Numeric	Min=0 Mean=4 Max=212	Number of incidents of fighting (SSR)
Sex_Harass	Numeric	Min=0 Mean=0 Max=33	Number of incidents of sexual harassment (SSR)
Rape	Numeric	Min=0 Mean=0 Max=3	Number of rapes (SSR)
Invol_Sex_Deviate_Interc	Numeric	Min=0 Mean=0 Max=2	Number of incidents of involuntary sexual deviate intercourse (SSR)
Sex_Ass	Numeric	Min=0 Mean=0 Max=19	Number of sexual assaults (SSR)
Agg_Indecent_Ass	Numeric	Min=0 Mean=0 Max=17	Number of aggravated indecent assaults (SSR)
Indecent_Ass	Numeric	Min=0 Mean=0 Max=11	Number of indecent assaults (SSR)
Kidnapping	Numeric	Min=0 Mean=0 Max=2	Number of incidents of kidnapping (SSR)
Threats	Numeric	Min=0 Mean=1 Max=98	Number of incidents of threats against a school official or student (SSR)
Reckless_End	Numeric	Min=0 Mean=0 Max=177	Number of incidents of reckless endangerment (SSR)
Robbery	Numeric	Min=0 Mean=0 Max=36	Number of robberies (SSR)
Theft	Numeric	Min=0 Mean=1 Max=49	Number of thefts (SSR)
Attempt_or_Commission_ManSl_Murder	Numeric	Min=0 Mean=0 Max=1	Number of attempts or commissions of murder or manslaughter (SSR)
Bullying	Numeric	Min=0 Mean=1 Max=401	Number of incidents of bullying (SSR)
Suicide_Attempt	Numeric	Min=0 Mean=0 Max=3	Number of suicide attempts (SSR)
Suicide_Commit	Numeric	Min=0 Mean=0 Max=1	Number of suicide commissions (SSR)
Burglary	Numeric	Min=0 Mean=0 Max=18	Number of burglaries (SSR)
Arson	Numeric	Min=0 Mean=0 Max=4	Number of incidents of arson (SSR)
Vandal	Numeric	Min=0 Mean=1 Max=34	Number of incidents of vandalism (SSR)
Rioting	Numeric	Min=0 Mean=0 Max=8	Number of incidents of rioting (SSR)

Variable	Type	Description/Distribution	Definition (primary original dataset)
Terror_Thr	Numeric	Min=0 Mean=0 Max=24	Number of terroristic threats (SSR)
Disorder_Con	Numeric	Min=0 Mean=3 Max=1002	Number of incidents of disorderly conduct (SSR)
Handgun	Numeric	Min=0 Mean=0 Max=4	Number of incidents of handgun possession (SSR)
Rifle_Shotg	Numeric	Min=0 Mean=0 Max=2	Number of incidents of rifle or shotgun possession (SSR)
Oth_Firearm	Numeric	Min=0 Mean=0 Max=3	Number of incidents of other firearm possession (SSR)
Poss_Knife	Numeric	Min=0 Mean=0 Max=20	Number of incidents of knife possession (SSR)
Cutting_Instru	Numeric	Min=0 Mean=0 Max=15	Number of incidents of possession of a cutting instruments (SSR)
Explosive	Numeric	Min=0 Mean=0 Max=15	Number of incidents of possession of an explosive (SSR)
BB_Gun	Numeric	Min=0 Mean=0 Max=5	Number of incidents of possession of a BB or pellet gun (SSR)
Other_Weapon	Numeric	Min=0 Mean=0 Max=13	Number of incidents of possession of another weapon (SSR)
Poss_Control_Sub	Numeric	Min=0 Mean=1 Max=31	Number of incidents of possession or use of a controlled substance (SSR)
Sale_Dist_Control_Sub	Numeric	Min=0 Mean=0 Max=23	Number of incidents of sale or distribution of a controlled substance (SSR)
Sale_Poss_Use_Influ_Alcohol	Numeric	Min=0 Mean=0 Max=35	Number of incidents of sale, possession, use, or being under the influence of alcohol (SSR)
Poss_Use_Sale_Tobacco	Numeric	Min=0 Mean=2 Max=76	Number of incidents of possession, use or sale of tobacco (SSR)
Truancy_Rate	Numeric	Min=0 Mean=7 Max=122.69	Rate of truancy based on aggregated number of habitually truant students (SSR)
Academic_OSS	Numeric	Min=0 Mean=1 Max=592	Number of students suspended for academic violations (SSR)
Conduct_OSS	Numeric	Min=0 Mean=40 Max=3460	Number of students suspended for violation of student code of conduct (SSR)

Variable	Type	Description/Distribution	Definition (primary original dataset)
Drug_Alc_ISS	Numeric	Min=0 Mean=1 Max=58	Number of students suspended for drug or alcohol violations (SSR)
Tobacco_OSS	Numeric	Min=0 Mean=1 Max=39	Number of students suspended for tobacco violations (SSR)
Viol_OSS	Numeric	Min=0 Mean=12 Max=1844	Number of students suspended for violent incidents (SSR)
Weapon_OSS	Numeric	Min=0 Mean=1 Max=20	Number of students suspended for weapons violations (SSR)
Academic_Expul	Numeric	Min=0 Mean=0 Max=2	Number of students expelled for academic violations (SSR)
Conduct_Expul	Numeric	Min=0 Mean=0 Max=32	Number of students expelled for violation of student code of conduct (SSR)
Drug_Alc_Expul	Numeric	Min=0 Mean=0 Max=28	Number of students expelled for drug or alcohol violations (SSR)
Tobacco_Expul	Numeric	Min=0 Mean=0 Max=3	Number of students expelled for tobacco violations (SSR)
Viol_Expul	Numeric	Min=0 Mean=0 Max=31	Number of students expelled for violent incidents (SSR)
Weapon_Expul	Numeric	Min=0 Mean=0 Max=21	Number of students expelled for weapons violations (SSR)
M_Pro	Numeric	Min=0 Mean=75 Max=100	Percentage of students scoring proficient or above on state mathematics/Algebra 1 assessments (PDE)
R_Pro	Numeric	Min=0 Mean=71 Max=100	Percentage of students scoring proficient or above on state reading/Literature assessments (PDE)

Definitions (Pennsylvania Department of Education, 2014)

The following definitions are included to provide a uniform fundamental understanding of a particular criminal offense or key item as it relates to completion of the School Safety Report. Some of the definitions include actual definitions and citations from Title 18 Crime Code for further clarification.

- Incident:** A specific act or offense involving one or more victims and one or more offenders. A reportable incident includes one or more acts of misconduct, involving one or more offenders violating criteria defined under Pennsylvania's Act 26 of 1995. These include but are not limited to any behavior that violates a school's educational mission or climate of respect or jeopardizes the intent of the school to be free of aggression against persons or property, drugs, weapons, disruptions, and disorder. Examples are incidents involving acts of violence, possession of a weapon, or the possession, use or sale of a controlled substance, alcohol, or tobacco by any person on school property; at school-sponsored events; and on school transportation to and from school.

The “Misconduct Definitions” provides a listing and definition of all reportable incidents.

- **Aggravated Assault:** An unlawful attack by one person upon another in which the offender uses a weapon or displays it in a threatening manner, or the victim suffers obvious severe or aggravated bodily injury involving apparent broken bones, loss of teeth, possible internal injury, severe lacerations, or loss of consciousness. This also includes assault with disease (as in cases when the offender is aware that he/she is infected with a deadly disease) by biting, spitting, etc.
- **Assault Resulting in Serious Injury:** An intentional physical attack causing the victim obvious severe or aggravated bodily injury involving (a) broken bones, loss of teeth, possible internal injuries; severe lacerations and bleeding; or loss of consciousness; and/or (b) requiring emergency medical services by trained school personnel or other health professionals (e.g. EMS) and/or hospitalization. If the offender used a weapon in an assault resulting in serious injury, report both Assault Resulting in Serious Injury and Assault Involving Use of a Weapon. Fights or affrays, where no weapon was used, resulting in no apparent or serious injuries are required to be reported only if the incident resulted in suspensions or expulsion for the student. Local School Board policy may require reporting of fights or affrays to law enforcement.
- **Assault Involving Use of a Weapon:** An assault by one person against another where the attacker either uses a weapon or displays a weapon in a threatening manner. Weapon is defined as: any firearm or explosive device; force-impacting device; knife or sharp-edged or sharp-pointed utensil, device or tool; or any article, instrument or substance which can or is likely to produce death or great bodily harm.
- **Assault on School Employee (crime against a person):** An unlawful attack by one person upon another. By definition there can be no attempted assaults, only completed assaults. The act should be intentionally, knowingly, or recklessly causing bodily injury or serious bodily injury to an employee(s).
- **Assault on Student (crime against a person):** An unlawful attack by one student upon another. By definition there can be no attempted assaults, only completed assaults. The act should intentionally, knowingly, or recklessly cause bodily injury and/or serious bodily injury to a student(s).
- **Reckless Endangering (crime against persons):** Engaging in conduct that places or may place another person in danger of death or serious bodily injury.

Crime Code Citation: §901. Criminal Attempt: A person commits an attempt when with intent to commit a specific crime; he/she does any act which constitutes a substantial step toward the commission of that crime.

Crime Code Citation: §2702. Aggravated Assault: A person is guilty of aggravated assault if he/she:

- (1) Attempts to cause serious bodily injury to another or causes such injury intentionally. Knowingly or recklessly under circumstances manifesting extreme indifference to the value of human life;
- (2) Attempts to cause or intentionally, knowingly or recklessly causes serious bodily injury to any of the officers, agents, employees or other persons enumerated below or to an employee of an agency, company or other entity engaged in public transportation, while in the performance of duty;
- (3) Attempts to cause or intentionally or knowingly causes bodily injury to any of the officers, agents, employees or other persons enumerated below in the performance of duty;

- (4) Attempts to cause or intentionally or knowingly causes bodily injury to another with a deadly weapon;
- (5) Attempts to cause or intentionally or knowingly causes bodily injury to a teaching staff member, school board member or other employee, including a student employee, of any elementary or secondary publicly-funded educational institution, any elementary or secondary private school licensed by the Department of Education or any elementary or secondary parochial school while acting in the scope of his or her employment or because of his or her employment relationship to the school;
- (6) Attempts by physical menace to put any of the officers, agents, employees or other persons enumerated below while in the performance of duty, in fear of imminent serious bodily injury;
- (7) Uses tear or noxious gas or uses an electric or electronic incapacitation device against any officer or other person enumerated below.

The officers, agents, employees and other persons referred to above shall be as follows:

- (1) Police officer
- (2) Firefighter
- (3) County adult probation or parole officer
- (4) County juvenile probation or parole officer
- (5) An agent of the Pennsylvania Board of Probation and Parole
- (6) Sheriff
- (7) Deputy Sheriff
- (8) Liquor control enforcement agent
- (9) Officer or employee of a correctional institution, county jail or prison, juvenile detention center or any other facility to which the person has been ordered by the court pursuant to a petition alleging delinquency under 42.Pa.C.S. Ch. 63 (relating to juvenile matters).
- (10) Judge of any court in the unified judicial system
- (11) The Attorney General
- (12) A deputy attorney general
- (13) A district attorney
- (14) An assistant district attorney
- (15) A public defender
- (16) An assistant public defender
- (17) A federal law enforcement official
- (18) A state law enforcement official
- (19) A local law enforcement official
- (20) Any person employed to assist or who assists ANY federal, state or local law enforcement official
- (21) Emergency medical services personnel
- (22) Parking enforcement officer
- (23) A district justice
- (24) A constable
- (25) A deputy constable
- (26) A psychiatric aide
- (27) A teaching staff member, a school board member or other employee, including a student employee, of any elementary or secondary publicly funded educational institution, any elementary or secondary private school licensed by the Department of Education or any elementary or secondary parochial school while acting in the scope of his or her employment or because of his or her employment relationship to the school.
- (28) Governor
- (29) Lieutenant Governor
- (30) Auditor General
- (31) State Treasurer
- (32) Member of the General Assembly
- (33) An employee of the Department of Environmental Protection

- (34) An individual engaged in the private detective business as defined in section 2(a) and (b) of the act of August 21, 1953 (P.L. 1273, No. 361), known as the Private Detective Act of 1953
- (35) An employee or agent of a county children and youth social service agency.

Crime Code Citation: §2705. Recklessly Endangering another Person: A person commits a misdemeanor of the second degree if he/she recklessly engages in conduct which places or may place another person in danger of death or serious bodily injury.

Arson: The unlawful and intentional damage or attempt to damage any real or personal property by fire or incendiary device. Setting a fire (by match, lighter, fireworks, firecrackers, trashcan fires, Molotov cocktails, or any other incendiary device) providing aid, counsel or pay toward same. This category does not include a simple act of lighting a match.

Crime Code Citation: §3301 Arson and other Related Offenses:

- (1) **Arson endangering persons** – A person commits a felony of the first degree if he intentionally starts a fire or causes an explosion, or if he aids, counsels, pays or agrees to pay another to cause a fire or explosion, whether on his own property or on that of another, and if:
 - (i) he thereby recklessly places another person in danger of death or bodily injury, including but not limited to a firefighter, police officer or other person actively engaged in fighting the fire; or
 - (ii) he commits the act with the purpose of destroying or damaging an inhabited building or occupied structure of another.
- (2) **Arson endangering property** – A person commits a felony of the second degree if he intentionally starts a fire or causes an explosion whether on his own property or that of another, or if he aids, counsels, pays or agrees to pay another to cause a fire or explosion, and if:
 - (i) he commits the act with intent of destroying or damaging a building or unoccupied structure of another;
 - (ii) he thereby recklessly places an inhabited building or occupied structure of another in danger of damage or destruction; or
 - (iii) he commits the act with intent of destroying or damaging any property, whether his own or of another, to collect insurance for such loss.
- (3) **Reckless burning or exploding** – A person commits a felony of the third degree if he intentionally starts a fire or causes an explosion, or if he aids, counsels, pays or agrees to pay another to cause a fire or explosion, whether on his own property or on that of another, and thereby recklessly:
 - (i.) places an uninhabited building or unoccupied structure of another in danger of damage or destruction; or
 - (ii.) places any personal property of another having a value that exceeds \$5,000 or if the property is an automobile, airplane, motorcycle, motorboat or other motor-propelled vehicle in danger of damage or destruction.
- (4) **Possession of explosive or incendiary materials or devices** – A person commits a felony of the third degree if he possesses, manufactures or transports any incendiary or explosive material with the intent to use or to provide such device or material to commit any offense described in (1), (2) or (3).

Bodily Injury: Impairment of physical condition or inflicting substantial pain.

Bullying: By law, "bullying" shall mean an intentional electronic, written, verbal or physical act, or a series of acts:

- (1) directed at another student or students;
- (2) which occurs in a school setting;
- (3) that is severe, persistent or pervasive; and
- (4) that has the effect of doing any of the following:

- (i) substantially interfering with a student's education;
- (ii) creating a threatening environment; or
- (iii) substantially disrupting the orderly operation of the school; and "school setting" shall mean in the school, on school grounds, in school vehicles, at a designated bus stop or at any activity sponsored, supervised or sanctioned by the school.

However, it should be noted that both genders can engage in direct or indirect bullying, and it can be either physical and/or psychological in nature. Bullying can be carried out by a single individual or by a group. The target/victim of bullying can be a single individual or a group of students.

The behavior can be either overt or covert in nature utilizing various methods of communication. For example, the term cyber bullying is being used to describe bullying behavior which occurs on the Internet.

The term bullying should not be used when there is a mutual confrontation between two students or groups of students. Behavior is clearly bullying when: (1) **There is intent to harm** - the perpetrator appears to find pleasure in taunting and continues even when the target's distress is obvious. Mutual "teasing" should not be confused with bullying behavior and (2) **There is intensity and duration** - the taunting continues over a period of time, and is not welcomed by the target.

Please note: The intent of this reporting system is to document all incidents of bullying behavior. Incidents which involve law enforcement, although they may fall under the definitions provided, should not be considered the only reportable incidents of bullying behavior.

Burglary (crimes against property): The unlawful entry into a building or other structure with the intent to commit a felony or theft. It is not necessary that force be used in gaining entry, neither is it necessary that property loss occur. Attempts to unlawfully enter a structure without expressed permission are also counted in this category.

Crime Code Citation: §3502. Burglary: A person is guilty of burglary if he enters a building or occupied structure, or separately secured or occupied portion thereof, with intent to commit a crime therein, unless the premises are at the time open to the public or the actor is licensed or privileged to enter.

Dangerous incident: A weapons possession incident resulting in arrest (guns, knives, or other weapons) or a violent incident resulting in arrest (homicide, kidnapping, robbery, sexual offenses, and assaults) as reported on the Violence and Weapons Possession Report (PDE-360).

Department shall mean the Pennsylvania Department of Education.

Disorderly Conduct: A person is guilty of disorderly conduct if, with intent to cause public inconvenience, annoyance or alarm, or recklessly creating a risk thereof, he:

- (1) engages in fighting or threatening, or in violent or tumultuous behavior;
- (2) makes unreasonable noise;
- (3) uses obscene language, or makes an obscene gesture; or
- (4) creates a hazardous or physically offensive condition by any act which serves no legitimate purpose of the actor.

Crime Code Citation: §5503. Disorderly Conduct: A person is guilty of disorderly conduct if, with intent to cause public inconvenience, annoyance or alarm, or recklessly creating a risk thereof, he/she:

- (1) engages in fighting or threatening, or in violent or tumultuous behavior;
- (2) makes unreasonable noise;
- (3) uses obscene language, or makes an obscene gesture; or
- (4) creates a hazardous or physically offensive condition by any act which serves no legitimate purpose of the actor.

Fighting (Mutual Altercation): A student confrontation with another student in which the altercation is mutual, requiring physical restraint or resulting in injury or property damage. If the incident does not rise to that level, the incident should be classified as minor disruptive behavior or a minor infraction. Mutual participation in a fight involving physical violence, where there is no one main offender and no major injury. This does not include verbal confrontations, tussles or other minor confrontations. Law enforcement officers may refer this offense as simple assault. Administrators need to consider age and developmentally appropriate behavior before using this category.

Harassment: A person commits the crime of harassment when, with the intent to harass, annoy or alarm another, the person:

- (1) Strikes, shoves, kicks or otherwise subjects the other person to physical contact, or attempts or threatens to do the same;
- (2) Follows the other person in or about a public place or places; or
- (3) Engages in a course of conduct or repeatedly commits acts that serve no legitimate purpose.

Crime Code Citation: §2709. Harassment: A person commits the crime of harassment when, with intent to harass, annoy or alarm another, the person:

- (1) strikes, shoves, kicks or otherwise subjects the other person to physical contact, or attempts or threatens to do the same;
- (2) follows the other person in or about a public place or places;
- (3) engages in a course of conduct or repeatedly commits acts which serve no legitimate purpose;
- (4) Communicates to or about such other persons any lewd, lascivious, threatening or obscene words, language, drawings or caricatures;
- (5) communicates repeatedly in an anonymous manner;
- (6) communicates repeatedly at extremely inconvenient hours; or
- (7) communicates repeatedly in a manner other than specified in sections (4), (5) and (6).

Crime Code Citation: §279.1. Stalking: A person commits the crime of stalking when the person either:

- (1) Engages in a course of conduct or repeatedly commits acts towards another person, including following the person without proper authority, under circumstances which demonstrate either an intent to place such other person in reasonable fear of bodily injury or to cause substantial emotional distress to such other person; or
- (2) Engages in a course of conduct or repeatedly communicates to another person under circumstances which demonstrate or communicate either an intent to place such other person in reasonable fear of bodily injury or to cause substantial emotional distress to such other person.

Crime Code Citation: §5504. Harassment and Stalking by Communication or Address:

- (a) Harassment by communication or address – A person commits the crime of harassment by communication or address when, with intent to harass, annoy or alarm another, the person:
 - (1) Communicates to or about such other person any lewd, lascivious, threatening or obscene words, language, drawings or caricatures; or
 - (2) Communicates repeatedly in an anonymous manner;
 - (3) Communicates repeatedly at extremely inconvenient hours; or
 - (4) Communicates repeatedly in a manner not covered by paragraphs (2) and (3).
- (b) Stalking by communication or address – A person commits the crime of stalking by communication or address when the person engages in a course of conduct or repeatedly communicates to another under circumstances which demonstrate or communicate either of the following:
 - (1) An intent to place such other person in reasonable fear of bodily injury.
 - (2) An intent to cause substantial emotional distress to such other person.

Homicide (Crime against a Person) (Murder or Manslaughter): Occurs when a person intentionally,

knowingly, recklessly, or negligently causing the death of a human being. Homicide is classified as murder, voluntary manslaughter, or involuntary manslaughter.

- Suicide is not considered homicide.
- Attempted murder or assault with intent to murder should be reported as aggravated assault.
- Justifiable homicide (e.g., legitimate self defense) or the killing of a perpetrator of a serious criminal offense by a peace officer or by a private individual should be reported as other.

Crime Code Citation: §2501. Criminal homicide: A person is guilty of criminal homicide if he/she intentionally, knowingly, recklessly or negligently causes the death of another human being. Criminal homicide shall be classified as murder, voluntary manslaughter, or involuntary manslaughter.

Crime Code Citation: §2502. Murder:

- (a) Murder of the first degree – A criminal homicide constitutes murder of the first degree when it is committed by an intentional killing.
- (b) Murder of the second degree - A criminal homicide constitutes murder of the second degree when it is committed while defendant was engaged as a principal or an accomplice in the perpetration of a felony.
- (c) Murder of the third degree - All other kinds of murder shall be murder of the third degree. Murder of the third degree is a felony of the first degree.
- (d) Definitions – As used in this section the following words and phrases shall have the meanings given to them in this subsection:

“**Intentionally Killing**” – Killing by means of poison, or by lying in wait, or by any other kind of willful, deliberate and premeditated killing.

“**Perpetration of a felony**” – The act of the defendant in engaging in or being an accomplice in the commission of, or an attempt to commit, or flight after committing, or attempting to commit robbery, rape, or deviate sexual intercourse by force or threat of force, arson, burglary or kidnapping.

Crime Code Citation: §2503. Voluntary manslaughter: A person who kills an individual without lawful justification commits voluntary manslaughter if at the time of the killing he/she is acting under a sudden and intense passion resulting from serious provocation by:

- (1) the individual killed; or
- (2) another whom the actor endeavors to kill, but he negligently or accidentally causes the death of the individual killed.

Crime Code Citation: §2504. Involuntary manslaughter: A person is guilty of involuntary manslaughter when as a direct result of the doing of an unlawful act in a reckless or grossly negligent manner, or the doing of a lawful act in a reckless or grossly negligent manner, he/she causes the death of another person.

Kidnapping (Crime against Persons): The removal, restraining or confinement of an individual by another through force, threat, or deception or (if person is under 14 years) without consent of a parent, guardian or school. Kidnapping/abduction includes hostage taking. A parent taking a child in violation of a court order, although it may be a crime, is not kidnapping for this purpose.

Crime Code Citation: §2901. Kidnapping: A person is guilty of kidnapping if he unlawfully removes another a substantial distance under the circumstances from the place where he/she is found, or if he/she unlawfully confines another for a substantial period in a place of isolation, with any of the following intentions:

- (1) To hold for ransom or reward, or as a shield or hostage.
- (2) To facilitate commission of any felony or flight thereafter.
- (3) To inflict bodily injury on or to terrorize the victim or another.
- (4) To interfere with the performance by public officials of any governmental or political function.

Local Educational Agency or LEA shall include a school district, an area vocational-technical school, an intermediate unit, or a charter school.

Minor Altercation: An incident which involves a single offender who commits a minor violent act against another individual and the other individual does not respond and the incident does not elevate to a more severe type of incident such as a fight or assault (e.g., “Student A” strikes “Student B” resulting in little injury and the “Student B” does not retaliate).

Persistently Dangerous School shall mean any public elementary, secondary, or charter school that meets any of the following criteria in the most recent school year and in one additional year of the two years prior to the most recent school year:

1. For a school whose enrollment is 250 or less, at least 5 dangerous incidents;
2. For a school whose enrollment is 251 to 1000, a number of dangerous incidents that represents at least 2% of the school’s enrollment;
3. For a school whose enrollment is over 1000, 20 or more dangerous incidents.

Racial/Ethnic Intimidation (Crime against a Person): Malicious intent toward another’s person or property based on race, color, religion or national origin is a hate crime.

Crime Code Citation: §2710. Ethnic intimidation: A person commits the offense of ethnic intimidation if, with malicious intention towards the actual or perceived race, color, religion, national origin, ancestry, mental or physical disability, sexual orientation, gender or gender identity of another individual or group of individuals, he/she commits an offense under any other provision of this article or under Chapter 33 (relating to arson, criminal mischief and other property destruction) exclusive of section 3307 (relating to institutional vandalism) or under section 3503 (relating to criminal trespass) with respect to such individual or his or her property or with respect to one or more members of such group or to their property.

Robbery (Crime against Person): The taking, or attempting to take, anything of value under confrontational circumstances from the control, custody or care of another person by force or threat of force or violence and/or by putting the victim in fear of immediate harm. The type of weapon used is to be entered into question 8 (Type of weapon involved). Because some type of assault may be an element of the crime of robbery, an assault should not be reported as an assault as long as it was performed in furtherance of the robbery. However, if the injury resulted in death, the incident must be reported as a homicide. A carjacking is a robbery offense where a motor vehicle is taken into force or threat of force.

- **Robbery with a Dangerous Weapon (Armed Robbery):** Theft or attempted theft of anything of value from the person of another, or from the area under the immediate bodily control of the other, by using a dangerous weapon or by an act threatening use of a dangerous weapon. A dangerous weapon is any article, instrument or substance that is likely to produce death or great bodily harm. Forcible theft or attempted theft from a person without the use of a dangerous weapon should be reported under Robbery without a Dangerous Weapon.
- **Robbery without a Dangerous Weapon:** The taking or attempting to take anything of value from another’s person, by force, or by an act threatening force or violence, which puts a victim in fear, without the use of a weapon. The stealing of someone’s property without the use of force or from a source other than the victim’s person is not included in this offense. If the taking from the person involves use of a dangerous weapon the incident is reported under Robbery With a Dangerous Weapon.

Crime Code Citation: §3701. Robbery:

- (1) A person is guilty of robbery if, in the course of committing a theft, he:
 - (i) Inflicts serious bodily injury upon another;

- (ii) Threatens another with or intentionally puts him in fear of immediate serious bodily injury;
 - (iii) Commits or threatens immediately to commit any felony of the first or second degree;
 - (iv) Inflicts bodily injury upon another or threatens another with or intentionally puts him in fear of immediate bodily injury; or physically takes or removes property from the person of another by force however slight.
- (2) An act shall be deemed “in the course of committing a theft” if it occurs in an attempt to commit theft or in flight after the attempt or commission.

Crime Code Citation: §3702. Robbery of motor vehicle: A person commits a felony of the first degree if he steals or takes a motor vehicle from another person in the presence of that person or any other person in lawful possession of the motor vehicle.

Safe Public School shall mean a public school that has not been designated as a persistently dangerous school under the standards for identifying persistently dangerous schools, or that has had that designation removed by the Department.

Serious bodily injury is a bodily injury that creates a substantial risk of death, or that causes serious, permanent disfigurement, or protracted loss or impairment of the function of any bodily member or organ.

Sexual Related Offenses: Sexual Offense (Crime against a Person) includes sexual intercourse, sexual contact or other unlawful behavior or conduct intended to result in sexual gratification. This category includes rape, statutory sexual assault, involuntary deviate sexual assault, sexual assault, aggravated indecent assault, indecent assault, and indecent exposure. Also included are prostitution and indecent exposure of private parts to the sight of another person in a lewd or indecent manner. This category doesn't include kissing, mooning, swearing or profanity. Administrators need to consider age and developmentally appropriate behavior before using this category.

- **Rape.** Rape may be statutory or forcible.
 - **Forcible Rape** is vaginal intercourse committed by force and without the consent of the victim, regardless of age.
 - **Statutory Rape** is vaginal intercourse committed on a child under the age of 16 by a person who is at least 12 years old and at least 4 years older than the victim, regardless of whether the victim consented. Consensual vaginal intercourse between a 13, 14 or 15 year old girl or boy and a 16 year old girl or boy is not a crime; statutory rape requires at least four years between birthdays of the victim and perpetrator. Some examples of incidents which must be reported under this category are consensual intercourse between a 19-year old and a 15-year old; consensual intercourse with a person who is mentally handicapped or incapacitated, or physically helpless, regardless of whether the victim consented; or intercourse with an intoxicated or drugged victim who is too incapacitated to give consent.
- **Indecent Assault.** Committing a sexual act with or in the presence of a child under the age of 16 years, by a person at least age 16 and at least five years older than the child, for sexual gratification, regardless of whether force was used, or whether the victim consented. Examples of acts to be reported under this category are intentional exposure of genitals in front of a child; showing a child pornography, secretly or in the child's presence; or photographing girls changing clothes or using toilets, if these acts are done for sexual gratification.
- **Sexual Assault (Not Involving Rape or Sexual Offense).** An assault of a sexual nature. An unauthorized and unwanted, intentional, or forcible touching of a sex organ of a person of either sex. Sex organs are the breasts of females and genital areas of males and females. This category includes forcibly and intentionally grabbing the clothed or unclothed breast or genitals of a person, without the consent of the victim. Report **attempted rape and attempted sexual offense** under this category. **The difference between Sexual Assault and Sexual Offense** is that Sexual Assault

involves **forcible and intentional touching without penetration** of a sex organ, and Sexual Offense involves **penetration** of a sex organ or anus by any object, or touching another's mouth or anus by the male sex organ.

- **Sexual Harassment (Crime against a Person)** Is discrimination against a student based on the student's submission or rejection of sexual advances and/or requests or creating an atmosphere of harassment based on sexual issues/activity. The unwelcome sexual advances, requests for sexual favors, other physical or verbal conduct or communication of a sexual nature, and any other gender based harassment which has the purpose or effect to interfere with the individual performance, work environment, or participation in school sponsored activities, or creates an intimidating, hostile, or offensive educational environment. This includes students and staff. Examples include behaviors such as leering, pinching, grabbing, suggestive comments, gestures, jokes or pressure to engage in sexual activity.
- **Institutional Sexual Assault – §3142.2.** A person who is an employee or agent of the Department of Corrections or a county correctional authority, youth development center, youth forestry camp, state or county juvenile detention facility, other licensed residential facility serving children and youth, or mental health or mental retardation facility or institution commits a felony of the third degree when that person engages in sexual intercourse, deviate sexual intercourse or indecent contact with an inmate, detainee, patient or resident. As used in this section, the term "agent" means a person who is assigned to work in a state or county correctional or juvenile detention facility, a youth development center, youth forestry camp, other licensed residential facility serving children and youth, or mental health or mental retardation facility or institution who is employed by any State or county agency or any person employed by an entity providing contract services to the agency.

1.1.1 Crime Code Related Citations Definitions:

1.1.2

- **Deviate sexual intercourse.** Sexual intercourse per os or per anus between human beings and any form of sexual intercourse with an animal. The term also includes penetration, however slight, of the genitals or anus of another person with a foreign object for any purpose other than good faith medical, hygienic or law enforcement procedures.
- **Forcible compulsion.** Compulsion by use of physical, intellectual, moral, emotional or psychological force, either express or implied. The term includes, but is not limited to, compulsion resulting in another person's death, whether the death occurred before, during or after sexual intercourse.
- **Indecent contact.** Any touching of the sexual or other intimate parts of the person for the purposes of arousing or gratifying sexual desire, in either person.

Sexual intercourse. In addition to its ordinary meaning, includes intercourse per os and per anus, with some penetration, however slight; emission is not required.

Crime Code Citation: §3121. Rape: A person commits a felony of the first degree when the person engages in sexual intercourse with a complainant:

- (1) By forcible compulsion.
- (2) By threat of forcible compulsion that would prevent resistance by a person of reasonable resolution.
- (3) Who is unconscious or where the person knows that the complainant is unaware that the sexual intercourse is occurring.
- (4) Where the person has substantially impaired the complainant's power to appraise or control his or her conduct by administering or employing, without the knowledge of the complainant, drugs, intoxicants or other means for the purposes of preventing resistance.
- (5) Who suffers from mental disability which renders the complainant incapable of consent.

- **Rape of a child.** A person commits the offense of rape of a child, a felony of the first degree, when the person engages in sexual intercourse with a complainant who is less than 13 years old.
- **Rape of a child with serious bodily injury.** A person commits the offense of rape of a child resulting in serious bodily injury, a felony of the first degree, when the person engages in sexual intercourse with a complainant who is less than 13 years of age and the complainant suffers serious bodily injury in the course of the offense.

Crime Code Citation: §3122.1. Statutory sexual assault: A person commits a felony of the second degree when that person engages in sexual intercourse with a complainant under the age of 16 years and that person is four or more years older than the complainant and the complainant and the person are not married to each other.

Crime Code Citation: §3123. Involuntary deviate sexual intercourse: A person commits a felony of the first degree when the person engages in deviate sexual intercourse with a complainant:

- (1) By forcible compulsion;
- (2) By threat of forcible compulsion that would prevent resistance by a person of reasonable resolution;
- (3) Who is unconscious or where the person knows that the complainant is unaware that the sexual intercourse is occurring;
- (4) Where the person has substantially impaired the complainant's power to appraise or control his or her conduct by administering or employing, without the knowledge of the complainant, drugs, intoxicants or other means for the purpose of preventing resistance;
- (5) Who suffers from mental disability which renders him or her incapable of consent; or
- (6) Who is less than 16 years of age and the person is four or more years older than the complainant and the person are not married to each other.

Involuntary deviate sexual intercourse with a child. A person commits involuntary deviate sexual intercourse with a child, a felony of the first degree, when the person engages in deviate sexual intercourse with a complainant who is less than 13 years of age.

Involuntary deviate sexual intercourse with a child with serious bodily injury. A person commits an offense under this section with a child resulting in serious bodily injury, a felony of the first degree, when the person violates this section and the complainant is less than 13 years of age and the complainant suffers serious bodily injury in the course of the offense.

Crime Code Citation: §3124.1. Sexual assault. Except as provide in §3121 (relating to rape) Or 3123 (relating to involuntary deviate sexual intercourse), a person commits a felony of the second degree when that person engages in sexual intercourse or deviate sexual intercourse with a complainant without the complainant's consent.

Crime Code Citation: §3125. Aggravated indecent assault. Except as provided in §§ 3121 (relating to rape), 3122.1 (relating to statutory sexual assault), 3123 (relating to involuntary deviate sexual intercourse) and 3124.1 (relating to sexual assault), a person who engages in penetration, however slight, of the genitals or anus of a complainant with a part of his person's body for any purpose other than good faith medical, hygienic or law enforcement procedures commits aggravated indecent assault if:

- (1) The person does so without the complainant's consent;
- (2) The person does so by forcible compulsion;
- (3) The person does so by threat of forcible compulsion that would prevent resistance by a person of reasonable resolution;
- (4) The complainant is unconscious or the person knows that the complainant is unaware that the penetration is occurring;

- (5) The person has substantially impaired the complainant's power to appraise or control his or her conduct by administering or employing, without the knowledge of the complainant, drugs, intoxicants or other means for the purpose of preventing resistance;
- (6) The complainant suffers from mental disability which renders him or her incapable of consent;
- (7) The complainant is less than 13 years of age; or
- (8) The complainant is less than 16 years of age and the person is four or more years older than the complainant and the person are not married to each other.

Crime Code Citation: §3126. Indecent assault. A person who has indecent contact with the complainant or causes the complainant to have indecent contact with the person is guilty of indecent assault if:

- (1) The person does so without the complainant's consent;
- (2) The person does so by forcible compulsion;
- (3) The person does so by threat of forcible compulsion that would prevent resistance by a person of reasonable resolution;
- (4) The complaint is unconscious or the person knows that the complainant is unaware that the indecent contact is occurring;
- (5) The person has substantially impaired the complainant's power to appraise or control his or her conduct by administering or employing, without the knowledge of the complainant, drugs, intoxicants or other means for the purpose of preventing resistance;
- (6) The complainant suffers from mental disability which renders him or her incapable of consent;
- (7) The complainant is less than 13 years of age; or
- (8) The complainant is less than 16 years of age and the person is four or more years older than the complainant and the person are not married to each other.

Crime Code Citation: §3124.2. Institutional sexual assault. A person who is an employee or agent of the Department of Corrections or a county correctional authority, youth development center, youth forestry camp, state or county juvenile detention facility, other licensed residential facility serving young children and youth, or mental health or mental retardation facility or institution commits a felony of the third degree when that person engages in sexual intercourse, deviate sexual intercourse or indecent contact with an inmate, detainee, patient or resident. As used in this section, the term "agent" means a person who is assigned to work in a state or county correctional or juvenile detention facility, a youth development center, youth forestry camp, other licensed residential facility serving children and youth, or mental health or mental retardation facility or institution who is employed by any state or county agency or any person employed by an entity providing contract services to the agency.

Simple assault is the unlawful physical attack by one student upon another where neither the offender displays a weapon, nor the victim suffers obvious severe or aggravated bodily injury involving apparent broken bones, loss of teeth, possible internal injury, severe laceration or loss of consciousness.

Crime Code Citation: §2701. Simple Assault. A person is guilty if he/she:

- (1) Attempts to cause or intentionally, knowingly or recklessly causes bodily injury to another;
- (2) Negligently causes bodily injury to another with a deadly weapon;
- (3) Attempts by physical menace to put another in fear of imminent serious bodily injury; or
- (4) Conceals or attempts to conceal a hypodermic needle on his person and intentionally or knowingly penetrates a law enforcement officer or an officer or an employee of a correctional institution, county jail or prison, detention facility or mental hospital during the course of an arrest or any search of the person.

Theft by Unlawful Taking or Disposition:

- (1) Movable property – A person is guilty of theft if he unlawfully takes, or exercises unlawful control over, movable property of another with intent to deprive him thereof.
- (2) Immovable property – A person is guilty of theft if he unlawfully transfers, or exercises unlawful control over, immovable property of another or any interest therein with intent to benefit himself

or another not entitled thereto.

- (3) If amount is \$50 or more but less than \$200 it is a misdemeanor 2nd degree.
- (4) If amount is less than \$50 it is a misdemeanor 3rd degree

Threatening or Intimidating a School Official or a Student (Crime against persons) (Physical, verbal, written, or electronic threat (e.g., internet) or intimidation) is to unlawfully place another person in fear of bodily harm through verbal threats without displaying a weapon or subjecting the person to actual physical attack; stalking (i.e., secretly or stealthily pursuing another, spying on or watching another person, with or without the intent to harm, frighten, or coerce) should be included.

Vandalism is the unlawful desecration of a building or other structure with the intent to commit damage.

Crime Code Citation: §3307. Institutional vandalism. A person commits the offense of institutional vandalism if he knowingly desecrates, vandalizes, defaces or otherwise damages:

- (1) Any church, synagogue or other facility or place used for religious worship or other religious purposes;
- (2) Any cemetery, mortuary or other facility used for the purpose of burial or memorializing the dead;
- (3) Any school, educational facility, community center, municipal building, courthouse facility or juvenile detention center;
- (4) The grounds adjacent to and owned or occupied by any facility set forth in paragraph (1), (2) or (3); or
- (5) Any personal property located in any facility set forth in this subsection.

Victim shall mean the student/person against whom a criminal offense has been perpetrated while the student/person was in or on the grounds or under the supervision of their public elementary or secondary school.

ALCOHOL and CONTROLLED SUBSTANCES-RELATED CITATIONS:

Drug shall mean controlled substance as defined under the 13 act of April 14, 1972 (P.L.233, No.64), known as "The Controlled Substance, Drug, Device and Cosmetic Act."

"*Drug*" means: (i) substances recognized in the official United States Pharmacopoeia, or official National Formulary or any supplement to either of them; and (ii) substances intended for use in the diagnosis, cure, mitigation, treatment or prevention of disease in man or other animals; and (iii) substances (other than food) intended to affect the structure or any function of the human body or other animal body; and (iv) substances intended for use as a component of any article specified in clause (i), (ii) or (iii), but not including devices or their components, parts or accessories.

Drug paraphernalia shall mean paraphernalia as defined under the act of April 14, 1972 (P.L.233, No.64), known as "The Controlled Substance, Drug, Device and Cosmetic Act."

"*Drug paraphernalia*" means all equipment, products and materials of any kind which are used, intended for use or designed for use in planting, propagating, cultivating, growing, harvesting, manufacturing, compounding, converting, producing, processing, preparing, testing, analyzing, packaging, repackaging, storing, containing, concealing, injecting, ingesting, inhaling or otherwise introducing into the human body a controlled substance in violation of this act. It includes, but is not limited to:

- (1) Kits used, intended for use or designed for use in planting, propagating, cultivating, growing or harvesting of any species of plant which is a controlled substance or from which a controlled substance can be derived.
- (2) Kits used, intended for use or designed for use in manufacturing, compounding, converting, producing, processing or preparing controlled substances.
- (3) Isomerization devices used, intended for use or designed for use in increasing the potency of any

- species of plant which is a controlled substance.
- (4) Testing equipment used, intended for use or designed for use in identifying or in analyzing the strength, effectiveness or purity of controlled substances.
 - (5) Scales and balances used, intended for use or designed for use in weighing or measuring controlled substances.
 - (6) Diluents and adulterants, such as quinine hydrochloride, mannitol, mannite, dextrose and lactose, used, intended for use or designed for use in cutting controlled substances.
 - (7) Separation gins and sifters used, intended for use or designed for use in removing twigs and seeds from or in otherwise cleaning or refining marijuana.
 - (8) Blenders, bowls, containers, spoons and mixing devices used, intended for use or designed for use in compounding controlled substances.
 - (9) Capsules, balloons, envelopes and other containers used, intended for use or designed for use in packaging small quantities of controlled substances.
 - (10) Containers and other objects used, intended for use or designed for use in storing or concealing controlled substances.
 - (11) Hypodermic syringes, needles and other objects used, intended for use, or designed for use in parenterally injected controlled substances into the human body.
 - (12) Objects used, intended for use or designed for use in ingesting, inhaling or otherwise introducing marijuana, cocaine, hashish or hashish oil into the human body, such as:
 - (i) Metal, wooden, acrylic, glass, stone, plastic or ceramic pipes with or without screens, permanent screens, hashish heads or punctured metal bowls.
 - (ii) Water pipes.
 - (iii) Carburetion tubes and devices.
 - (iv) Smoking and carburetion masks.
 - (v) Roach clips; meaning objects used to hold burning material such as a marijuana cigarette, that has become too small or too short to be held in the hand.
 - (vi) Miniature cocaine spoons and cocaine vials.
 - (vii) Chamber pipes.
 - (viii) Carburetor pipes.
 - (ix) Electric pipes.
 - (x) Air-driven pipes.
 - (xi) Chillums.
 - (xii) Bongs.
 - (xiii) Ice pipes or chillers.

In determining whether an object is drug paraphernalia, a court or other authority should consider, in addition to all other logically relevant factors, statements by an owner or by anyone in control of the object concerning its use, prior convictions, if any, of an owner, or of anyone in control of the object, under any State or Federal law relating to any controlled substance, the proximity of the object, in time and space, to a direct violation of this act, the proximity of the object to controlled substances, the existence of any residue of controlled substances on the object, direct or circumstantial evidence of the intent of an owner, or of anyone in control of the object, to deliver it to persons who he knows, or should reasonably know, intend to use the object to facilitate a violation of this act, the innocence of an owner or of anyone in control of the object, as to a direct violation of this act should not prevent a finding that the object is intended for use or designed for use as drug paraphernalia, instructions, oral or written, provided with the object concerning its use, descriptive materials accompanying the object which explain or depict its use, national and local advertising concerning its use, the manner in which the object is displayed for sale, whether the owner, or anyone in control of the object, is a legitimate supplier of like or related items to the community, such as a licensed distributor or dealer of tobacco products, direct or circumstantial evidence of the ratio of sales of the objects to the total sales of the business enterprise, the existence and scope of legitimate uses for the object in the community, and expert testimony concerning its use.

The Crime Code contains too numerous citations related to alcohol to place on this site. Some of the

citations are as follows:

Crime Code Citation: §6310.3	Carrying a false identification card.
Crime Code Citation: §6310.2	Manufacture or sale of false identification cards.
Crime Code Citation: §6307	Misrepresentation of age.
Crime Code Citation: §6308	Purchase, consumption possession or transportation of alcoholic beverages.
Crime Code Citation: §6310.6	Misrepresentation of age to secure alcoholic beverages.
Crime Code Citation: §5505	Chemically influenced persons, public appearances.
Crime Code Citation: §7508	Growing marijuana.
Crime Code Citation: §6317	Driving under the influence in drug free zones, schools and school districts.
Crime Code Citation: §7509	Drug free urine, sale or use, tests.
Crime Code Citation: §308.	Intoxication.

TOBACCO RELATED CITATIONS:

Crime Code Citation: §6306. Furnishing cigarettes or cigarette papers. A person commits an offense if he furnishes to any minor, by gift, sale or otherwise, any cigarettes or cigarette paper.

Crime Code Citation: §6306.1. Use of tobacco in schools prohibited. A pupil who possesses or uses tobacco in a school building, a school bus or on school property owned by, leased by or under the control of a school district commits a summary offense. A summary offense under this section shall not be a criminal offense of record, shall not be reportable as a criminal act and shall not be placed on the criminal record of the offending school-age person if any such record exists.

Crime Code Citation: §6305. Sale of tobacco. A person is guilty of a summary offense if the person:

- (1) Sells a tobacco product to any minor;
- (2) Furnishes, by purchase, gift or other means, a tobacco product to a minor;
- (3) Locates or places a tobacco vending machine containing a tobacco product in a location accessible to minors;
- (4) Displays or offers a cigarette for sale out of a pack of cigarettes; or
- (5) Displays or offers for sale tobacco products in any manner which enables an individual other than the retailer or an employee of the retailer to physically handle tobacco products prior to purchase unless the tobacco products are located within the line of sight or under the control of a cashier or other employee during business hours.

An offense under this subsection shall not be a criminal offense of record, shall not be reportable as a criminal act and shall not be placed on the criminal record of the offender.

BOMB THREAT-RELATED CITATIONS:

Crime Code Citation: §2706. Terroristic threats. A person commits the crime of terroristic threats if the person communicates, either directly or indirectly, a threat to:

- (1) Commit any crime of violence with intent to terrorize another;
- (2) Cause evacuation of a building, place of assembly or facility of public transportation; or
- (3) Otherwise cause serious public inconvenience, or cause terror or serious public inconvenience with reckless disregard of the risk of causing such terror or inconvenience. The term "communicates" means, conveys in person or by written or electronic means, including telephone, electronic mail, Internet, facsimile, telex and similar transmissions.

Crime Code Citation: §4906. False reports to law enforcement authorities.

- (a) Falsely incriminating another – A person who knowingly gives false information to any law enforcement officer with intent to implicate another commits a misdemeanor of the second degree.
- (b) Fictitious reports – A person commits a misdemeanor of the third degree if he/she:

- (1) Reports to law enforcement authorities an offense or other incident within their concern knowing that it did not occur; or
- (2) Pretends to furnish such authorities with information relating to an offense or incident when he knows he has no information relating to such offense or incident.

Crime Code Citation: §6161. Carrying explosives on conveyances. A person is guilty of a misdemeanor of the second degree if he/she enters into or upon any railroad train, locomotive, tender or car thereof, or into or upon any automobile or other conveyance used for carrying of freight or passengers, having in his/her custody or about his person any nitroglycerine or other explosive, other than as freight regularly shipped as such.

Crime Code Citation: §5516. Facsimile weapons of mass destruction. A person commits an offense if the person intentionally, knowingly or recklessly manufactures, sells, purchases, transports or causes another to transport, delivers or causes another to deliver, possesses or uses a facsimile weapon of mass destruction and by such action, causes any of the following:

- (1) Terrifying, intimidating, threatening or harassing an individual.
- (2) Alarm or reaction on the part of any of the following:
 - (i.) A public or volunteer organization that deals with emergencies involving danger to life or property.
 - (ii.) A law enforcement organization.
- (3) Serious public inconvenience not limited to the evacuation of a building, place of assembly or facility of public transportation.

WEAPONS-RELATED CITATIONS:

Crime Code Citation: §6110.1. Possession of firearm by minor. A person under 18 years of age shall not possess or transport a firearm anywhere in this Commonwealth.

Crime Code Citation: §908. Prohibited offensive weapons. Person commits a misdemeanor of the first degree, if, except as authorized by law, he/she makes repairs, sells, or otherwise deals in, uses, or possesses any offensive weapon. As used in this section, the following words and phrases shall have the meanings given to them:

- Firearm – Any weapon which is designed to or may readily be converted to expel any projectile by the action of an explosive, or the frame or receiver of any such weapon.
- Offensive weapons – Any bomb, grenade, machine gun, sawed off shotgun with a barrel less than 18 inches, firearm specially made or specially adapted for concealment or silent discharge, any blackjack, sandbag, metal knuckles, dagger, knife, razor or cutting instrument, the blade of which is exposed in an automatic way by switch, push-button, spring mechanism, or otherwise, or other implement for the infliction of serious bodily injury which serves no common lawful purpose.

Crime Code Citation: §912. Possession of weapon on school property. A person commits a misdemeanor of the first degree if he possesses a weapon in the buildings of, on the grounds of, or in any conveyance providing transportation to and from any elementary or secondary publicly-funded educational institution, any elementary or secondary private school licensed by the Department of Education or any elementary or secondary parochial school. Notwithstanding the definition of “weapon” in section 907 (relating to possessing instruments of crime), “weapon” for purposes of this section shall include but not be limited to any knife, cutting instrument, cutting tool, nunchuck stick, firearm, shotgun, rifle and any other tool, instrument or implement capable of inflicting serious bodily harm.

CLARIFICATION: According to the Pennsylvania State Police, airsoft, pellet or replica guns are not considered a firearm as defined by the Crimes Code, Title 18. However, the possession of these types of guns in schools or on school property should be reported to local law enforcement. Furthermore, under the guidelines an incident of this nature is reportable under the "Possession of Other Weapon" category on the

Safe School Report. It is also recommended that school districts have a policy in effect that "prohibits or bans" any type of replica weapon from the campus/property and transportation to/from school that clearly states the sanction received if found in possession of such a device or replica. School districts should discuss this issue with their solicitor, board, local law enforcement and other key decision makers involved with the district's school safety planning.

RIOT RELATED CITATION:

Crime Code Citation: §5501.Riot. A person is guilty of riot, a felony of the third degree, if he/she participates with two or more others in a course of disorderly conduct:

- (1) With the intent to commit or facilitate the commission of a felony or misdemeanor;
- (2) With intent to prevent or coerce official action;
- (3) When the actor or any other participant to the knowledge of the actor uses or plans to use a firearm or other deadly weapon.

BIBLIOGRAPHY

- Adams, A. T. (2000). The status of school discipline and violence. *The ANNALS of the American Academy of Political and Social Science*, 567(1), 140–156.
- Advancement Project. (2010). *Test, Punish, and Push Out: How “Zero Tolerance” and High-Stakes Testing Funnel Youth into the School-to-Prison Pipeline*. Washington, D.C.: Advancement Project.
- Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology*, 30(1), 47–88.
- Aiken, L. S., West, S. G., Schwalm, D. E., Carroll, J. L., & Hsiung, S. (1998). Comparison of a randomized and two quasi-experimental designs in a single outcome evaluation efficacy of a university-level remedial writing program. *Evaluation Review*, 22(2), 207–244.
- Akers, R. L. (1973). *Deviant Behavior: A Social Learning Approach*. Belmont, CA: Wadsworth Publishing Co.
- Akers, R. L. (1990). Rational choice, deterrence, and social learning theory in criminology: The path not taken. *Journal of Criminal Law and Criminology*, 81(3), 653.
- Akers, R. L., & Jensen, G. F. (2011). The empirical status of social learning theory of crime and deviance: The past, present, and future. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Taking Stock: The Status of Criminological Theory* (Vol. 15, pp. 37–76). New Brunswick, NJ: Transaction Publishers.
- Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., & Radosevich, M. (1979). Social learning and deviant behavior: A specific test of a general theory. *American Sociological Review*, 44(4), 636–655.
- American Civil Liberties Union of Pennsylvania. (2015). *Beyond Zero Tolerance: Discipline and Policing in Pennsylvania Public Schools*. Philadelphia, PA: American Civil Liberties Union of Pennsylvania.
- American Psychological Association Zero Tolerance Task Force. (2006). *Are Zero Tolerance Policies Effective in the Schools? An Evidentiary Review and Recommendations*. Washington, D.C.: American Psychological Association.
- Anderson, C. S. (1982). The search for school climate: A review of the research. *Review of Educational Research*, 52(3), 368–420.

- Anderson, M. (2015, September 14). Will school-discipline reform actually change anything? *The Atlantic*. Washington, D.C.
- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton, N.J.: Princeton University Press.
- Arcia, E. (2006). Achievement and enrollment status of suspended students' outcomes in a large, multicultural school district. *Education and Urban Society*, 38(3), 359–369.
- Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, 60(1), 47–57.
- Astor, R. A., Guerra, N., & Acker, R. V. (2010). How can we improve school safety research? *Educational Researcher*, 39(1), 69–78.
- Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21(1), 1–42.
- Balfanz, R., Herzog, L., & Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist*, 42(4), 223–235.
- Baron, S. W. (2003). Self-control, social consequences, and criminal behavior: Street youth and the general theory of crime. *Journal of Research in Crime and Delinquency*, 40(4), 403–425.
- Baron, S. W. (2004). General strain, street youth and crime: A test of Agnew's revised theory. *Criminology*, 42(2), 457–484.
- Beaujean, A. A. & Morgan, G. B. (2016). Tutorial on using regression models with count outcomes using R. *Practical Assessment, Research & Evaluation*, 21(2).
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217.
- Bergdahl, M., Ehling, M., Elvers, E., Foldesi, E., Korner, T., Kron, A., ... Zilhao, M. J. (2007). *Handbook on Data Quality Assessment Methods and Tools*. Wiesbaden: Eurostat.
- Bernburg, J. G., & Krohn, M. D. (2003). Labeling, life chances, and adult crime: The direct and indirect effects of official intervention in adolescence on crime in early adulthood. *Criminology*, 41(4), 1287–1318.

- Bernburg, J. G., Krohn, M. D., & Rivera, C. J. (2006). Official labeling, criminal embeddedness, and subsequent delinquency: A longitudinal test of labeling theory. *Journal of Research in Crime and Delinquency*, 43(1), 67–88.
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1), 249–275.
- Bhabra, S., Hill, E., & Ghate, D. (2004). *National Evaluation of the Safer School Partnership Programme*. London: Youth Justice Board for England and Wales.
- Block, M. K., & Gerety, V. E. (1995). Some experimental evidence on differences between students and prisoner reactions to monetary penalties and risk. *Journal of Legal Studies*, 24(1), 123–138.
- Bloom, H. S., & Riccio, J. A. (2005). Using place-based random assignment and comparative interrupted time-series analysis to evaluate the jobs-plus employment program for public housing residents. *The ANNALS of the American Academy of Political and Social Science*, 599(1), 19–51.
- Blue Ribbon Commission on Safe Schools. (2012). *A Blueprint for Action*. Philadelphia, PA: School District of Philadelphia. Retrieved from <https://www.philasd.org/announcements/BRC-Report.pdf>.
- Boruch, R. F. (2015). *Administrative Record Quality and Integrated Data Systems*. Philadelphia, PA: Center for Research in Education and Social Policy. Retrieved from http://www.aisp.upenn.edu/wp-content/uploads/2015/06/Boruch_Record_Quality_Data_Systems.pdf
- Bowles, R., Reyes, M. G., & Pradiptyo, R. (2005). *Safer Schools Partnerships*. London: Youth Justice Board for England and Wales.
- Brady, K. P., Balmer, S., & Phenix, D. (2007). School-police partnership effectiveness in urban schools: An analysis of New York City's Impact Schools initiative. *Education and Urban Society*, 39(4), 455–478.
- Broidy, L. M. (2001). A test of general strain theory. *Criminology*, 39(1), 9–36.
- Brooks, D. (2015, September 29). The Prison Problem. *The New York Times*. Retrieved from <http://www.nytimes.com/2015/09/29/opinion/david-brooks-the-prison-problem.html>
- Brown, T. M. (2007). Lost and turned out: Academic, social, and emotional experiences of students excluded from school. *Urban Education*, 42(5), 432–455.

- Bureau of Justice Statistics. (2001). *Local Police Departments 1999*. Washington, D.C.: U.S. Department of Justice.
- Bureau of Justice Statistics. (2006). *Local Police Departments 2003*. Washington, D.C.: U.S. Department of Justice.
- Bureau of Justice Statistics. (2010). *Local Police Departments 2007*. Washington, D.C.: U.S. Department of Justice.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and Quasi-Experimental Designs for Research*. Boston, MA: Houghton Mifflin Company.
- Carrell, S., Hoekstra, M., & Kuka, E. (2016). *The long-run effects of disruptive peers* (Working Paper No. 22042). Cambridge, MA: National Bureau of Economic Research.
- Center for Disease Control and Prevention. (2014). *Youth Risk Behavior Surveillance - United States, 2013*. Washington, D.C.: U.S. Department of Health and Human Services.
- Chen, G. (2008). Communities, students, schools, and school crime: A confirmatory study of crime in U.S. High Schools. *Urban Education, 43*(3), 301–318.
- Chiang, H., & Gill, B. (2010). *Student Characteristics and Outcomes in Alternative and Neighborhood High Schools in Philadelphia: Final Report*. Washington, D.C.: Mathematica Policy Research, Inc.
- Christle, C. A., Jolivette, K., & Nelson, C. M. (2005). Breaking the school to prison pipeline: Identifying school risk and protective factors for youth delinquency. *Exceptionality, 13*(2), 69–88.
- Chute, E. (2015, June 24). Pittsburgh school board approves changes to Code of Student Conduct. *Pittsburgh Post-Gazette*. Pittsburgh, PA.
- Conger, R. D. (1976). Social control and social learning models of delinquent behavior: A synthesis. *Criminology, 14*(1), 17–40.
- Cook, P. J., Gottfredson, D. C., & Na, C. (2009). *School Crime Control and Prevention*. Rochester, NY: Social Science Research Network.
- Cornell, D. G., & Mayer, M. J. (2010). Why do school order and safety matter? *Educational Researcher, 39*(1), 7–15.
- Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology, 31*(1), 59–82.

- Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, *74*(5), 1191–1230.
- Darling-Hammond, L. (2006). Securing the right to learn: Policy and practice for powerful teaching and learning. *Educational Researcher*, *35*(7), 13–24.
- Dean, M. M. (2016, February 26). Two former Philly educators plead in cheating scandal. *Philadelphia Inquirer*. Philadelphia, PA.
- DeAngelis, C., & Wolf, P. J. (2016). *The School Choice Voucher: A “Get Out of Jail” Card?* Fayetteville, AR: University of Arkansas, Department of Education Reform.
- Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, *30*(3), 418–446.
- Diaz, J. J., & Handa, S. (2006). An assessment of propensity score matching as a nonexperimental impact estimator: Evidence from Mexico’s PROGRESA program. *Journal of Human Resources*, *41*(2), 319–345.
- Eide, E. (2000). Economics of criminal behavior. In B. Bouckaert & G. De Geest (Eds.), *Encyclopedia of Law and Economics, Volume I. The History and Methodology of Law and Economics* (pp. 345–389). Cheltenham, UK: Edward Elgar.
- Elliott, D. S., Ageton, S. S., & Canter, R. J. (1979). An integrated theoretical perspective on delinquent behavior. *Journal of Research in Crime and Delinquency*, *16*(1), 3–27.
- Ewing, P. C. (2000). Sensible zero tolerance protects students. *Harvard Education Letter*, *16*(1), 7–8.
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking Schools’ Rules: A Statewide Study of How School Discipline Relates to Students’ Success and Juvenile Justice Involvement*. New York, NY: Council of State Governments Justice Center.
- Federal Bureau of Investigation. (2009). *Crime in the United States, 2008*. Washington, D.C.: United States Department of Justice, Federal Bureau of Investigation.
- Fogg, N. P., Harrington, P. E., & Khatiwada, I. (2009). *The Tax and Transfer Fiscal Impacts of Dropping Out of High School in Philadelphia City and Suburbs*. Boston, MA: Center for Labor Market Studies, Northeastern University.

- Fowler, D. (2007). *Texas' School to Prison Pipeline: Dropout to Incarceration - The Impact of School Discipline and Zero Tolerance*. Austin, TX: Texas Appleseed.
- Fowler, D. (2011). School Discipline Feeds the "Pipeline to Prison." *Phi Delta Kappan*, 93(2), 14–19.
- Glander, M. (2015). *Documentation to the NCES Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2013-14*. Washington, D.C.: National Center for Education Statistics.
- Goldberg, E., Stillwell, R., & Little, J. (2013). *Evaluation of a New Edit Methodology for the Common Core of Data Nonfiscal Surveys*. Washington, D.C.: U.S. Census Bureau.
- Gonzalez, T. (2012). Keeping kids in schools: Restorative justice, punitive discipline, and the school to prison pipeline. *Journal of Law and Education*, 41, 281–335.
- Gottfredson, G. D., Gottfredson, D. C., Payne, A. A., & Gottfredson, N. C. (2005). School climate predictors of school disorder: Results from a national study of delinquency prevention in schools. *Journal of Research in Crime and Delinquency*, 42(4), 412–444.
- Gottfried, M. A. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis*, 31(4), 392–415.
- Graham, K. (2008, October 16). Phila. schools to expel most violent students. *Philadelphia Inquirer*, p. B01. Philadelphia, PA.
- Graham, K. (2009, February 26). Phila students challenge "Imagine 2014" plan. *Philadelphia Inquirer*. Philadelphia, PA.
- Gruber, J., & Kleiner, S. A. (2012). Do strikes kill? Evidence from New York State. *American Economic Journal: Economic Policy*, 4(1), 127–157.
- Gun-Free Schools Act, 7151 20 US (1994).
- Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201–209.
- Hahn, R., Fuqua-Whitley, D., Wethington, H., Lowy, J., Liberman, A., Crosby, A., ... Dahlberg, L. (2007). *The Effectiveness of Universal School-Based Programs for the Prevention of Violent and Aggressive Behavior*. Washington, D.C.: Task Force on Community Preventive Services.

- Hardy, D. (2014, December 9). Philadelphia's shift in discipline policy. *The Philadelphia Public School Notebook*. Philadelphia, PA.
- Harper, J. (2011). *The Effects of Prisonization on the Employability of Former Prisoners: First-hand Voices* (Doctoral Dissertation). ProQuest LLC, Ann Arbor: MI: University of Pennsylvania.
- Haut, E. R., & Pronovost, P. J. (2011). Surveillance bias in outcomes reporting. *Journal of the American Medical Association*, 305(23), 2462–2463.
- Herold, B. (2011, July 12). Did state officials bury 2009 cheating report? *Philadelphia Public School Notebook*. Philadelphia, PA.
- Herold, B., & Mezzacappa, D. (2012, August 16). New evidence: Philadelphia test cheating is likely far worse than previously revealed. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Herzog, T. N., Scheuren, F. J., & Winkler, W. E. (2007). *Data Quality and Record Linkage Techniques*. New York, NY: Springer.
- Hood, J. (2012, June 26). CPS seeks to lessen student suspensions. *Chicago Tribune*. Chicago, IL.
- Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467–475.
- Iwig, W., Berning, M., Marck, P., & Prell, M. (2013). *Data Quality Assessment Tool for Administrative Data*. Washington D.C.: National Agricultural Statistics Services, U.S. Census Bureau, and Economic Research Service.
- Jacob, R., Zhu, P., & Bloom, H. (2012). *A Practical Guide to Regression Discontinuity*. New York, NY: MDRC.
- Johnson, D. (1999a, December 1). Schools' new watchword: Zero Tolerance. *New York Times*, p. 1. New York: NY.
- Johnson, F. (1996). *Statistics in Brief: Revenues and Expenditures for Public Elementary and Secondary Education: School Year 1993-1994*. Washington, D.C.: National Center for Education Statistics.
- Johnson, I. M. (1999b). School violence: The effectiveness of a school resource officer program in a southern city. *Journal of Criminal Justice*, 27(2), 173–192.
- Jones, S. (2016, January 1). School Reform Commission is rigged against Philadelphia schoolchildren. *Philadelphia Inquirer*. Philadelphia, PA.

- Katz, C. M., Schaefer, D. R., & Uchida, C. D. (2002). *Exploring the Gang Problem in Redlands, CA: An Evaluation of a Problem Solving Partnership*. Silver Spring, MD: 21st Century Solutions.
- Kaufman, P., Chen, X., Choy, S. P., Chandler, K. A., Chapman, C. D., Rand, M. R., & Ringel, C. (1998). *Indicators of School Crime and Safety 1998*. Washington, D.C.: National Center for Education Statistics and Bureau of Justice Statistics.
- Kim, C., Losen, D., & Hewitt, D. (2010). *The School to Prison Pipeline: Structuring Legal Reform*. New York: NY: NYU Press.
- Klehr, D. G. (2009). Addressing the unintended consequences of No Child Left Behind and zero tolerance: Better strategies for safe schools and successful students. *Georgetown Journal on Poverty Law & Policy*, *16*, 585–610.
- Koren, M. (2016, March 30). Obama's clemency bucket list. *The Atlantic*. Washington, D.C.
- Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.
- Lind, A. (2015, March). *PIMS Errors: Validation, DQE, and ETL*. Presented at the 2015 Pennsylvania Department of Education Data Summit, Hershey, PA.
- Maguire, E. R., Solomon, S. E., & Uchida, C. D. (2003). *Reducing Disputes in Spartanburg, SC: An Evaluation of a Problem Solving Partnership*. Silver Spring, MD: 21st Century Solutions.
- Maltz, M. (1999). *Bridging Gaps in Police Crime Data*. Washington, D.C.: U.S. Department of Justice, Office of Justice programs.
- Mayer, M. J., & Leone, P. E. (1999). A structural analysis of school violence and disruption: implications for creating safer schools. *Education & Treatment of Children*, *22*(3), 333–356.
- McCorry, K. (2014, August 21). Dozens more Philly schools lose school police due to budget cuts. *NewsWorks*. Philadelphia, PA.
- McKay, M. D., Covell, K., & McNeil, J. (2006). *An Evaluation of Cape Breton Regional Police Service's Community Liaison Officer Program in Cape Breton-Victoria Region Schools*. Sydney, Canada: Cape Breton University.
- Mendez, L. M. R. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development*, *99*, 17–33.

- Mezzacappa, D. (2011, September 7). District pays Ackerman's \$905,000 buyout after funders back out. *Education Week*. Bethesda, MD.
- Mezzacappa, D. (2012, August 13). Hite tells safety summit: Prevention, not punishment, is key to good school climate. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Mezzacappa, D. (2015a, July 20). Pa. educators brace for fallout from plunge in PSSA scores. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Mezzacappa, D. (2015b, October 27). Mayor Nutter calls for dissolution the School Reform Commission. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Michalopoulos, C., Bloom, H. S., & Hill, C. J. (2004). Can propensity-score methods match the findings from a random assignment evaluation of mandatory welfare-to-work programs? *Review of Economics and Statistics*, 86(1), 156–179.
- Na, C., & Gottfredson, D. C. (2013). Police officers in schools: Effects on school crime and the processing of offending behaviors. *Justice Quarterly*, 30(4), 619–650.
- Nagin, D. S., & Paternoster, R. (1993). Enduring individual differences and rational choice theories of crime. *Law & Society Review*, 27(3), 467–496.
- National Center for Education Statistics. (2015). *Local Education Agency (School District) Universe Survey* (No. 2013-14 v.1a). Washington, D.C.: U.S. Department of Education.
- National Center for Education Statistics. (2016). *Grantee State - Pennsylvania*. Retrieved March 29, 2016, from <http://nces.ed.gov/programs/slds/state.asp?stateabbr=PA>.
- New York Times Editorial Board. (2016, February 6). Holding sentencing reform hostage. *The New York Times*. New York, NY.
- Nowak, J. (2015, September 9). Does Bill Hite have more time? Retrieved March 22, 2016, from <http://thephiladelphiacitizen.org/does-bill-hite-have-more-time/>.
- Office of the Surgeon General, National Center for Injury Prevention and Control, National Institute of Mental Health, & Center for Mental Health Services. (2001). *Youth Violence: A Report of the Surgeon General*. Rockville, MD: Office of the Surgeon General.
- Paternoster, R., & Iovanni, L. (1989). The Labeling perspective and delinquency: An elaboration of the theory and an assessment of the evidence. *Justice Quarterly*, 6(3), 359–394.

- Pennsylvania Department of Education. (2014). *Pennsylvania Information Management System Volume 2 User Manual*. Harrisburg, PA: Pennsylvania Department of Education.
- Pennsylvania Department of Education. (2015a). *Data for SY 2014-2015*. Harrisburg, PA: Pennsylvania Department of Education.
- Pennsylvania Department of Education. (2015b). *Technical Report for the 2015 Pennsylvania System of School Assessment*. Harrisburg, PA: Pennsylvania Department of Education.
- Pennsylvania Department of Education. (2015c). *The 2015 PSSA School Level English Language Arts, Math, and Science Proficiency Results by Grade Level and School Total*. Harrisburg, PA: Pennsylvania Department of Education.
- Petrosino, A., Guckenburg, S., & Fronius, T. (2012). "Policing schools" strategies: A review of the evaluation evidence. *Journal of Multidisciplinary Evaluation*, 8(17), 80–101.
- Petteruti, A. (2011). *Education Under Arrest: The Case Against Police in School*. Washington, D.C.: Justice Policy Institute.
- Polke, C. (2014, May 11). Fairless Elementary tackles discipline issues. *Pittsburgh Post-Gazette*. Pittsburgh, PA.
- Pope, C. (2012, October 3). District downplays punishments in new student code of conduct. *The Philadelphia Public School Notebook*. Philadelphia, PA.
- Public Citizens for Children and Youth. (2010). *A Citywide Effort to Roll Out Positive Behavior Support (PBS)*. Philadelphia, PA: Public Citizens for Children and Youth.
- Richardson, J. (Ed.). (2015). The 47th annual PDK/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, 97(1), 1–31.
- Robers, S., Zhang, A., Morgan, R., & Musu-Gillette, L. (2015). *Indicators of School Crime and Safety 2014*. Washington, D.C.: National Center for Education Statistics.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33–38.

- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach*. Thousand Oaks, CA: SAGE Publications.
- Rouff, R. (2010, March 10). District's Imagine 2014 plan includes focus on dropout prevention. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4), 169–188.
- Russo, A. (2003). Political educator. *Education Next*, 3(1), 38–43.
- Sampson, R. J., Morenoff, J. D., & Raudenbush, S. (2005). Social anatomy of racial and ethnic disparities in violence. *American Journal of Public Health*, 95(2), 224–231.
- School District of Philadelphia. (2002). *Code of Student Conduct (2002-03)*. Philadelphia, PA: School District of Philadelphia.
- School District of Philadelphia. (2008). *Code of Student Conduct (2008-09)*. Philadelphia, PA: School District of Philadelphia.
- School District of Philadelphia. (2009). *Imagine 2014: Building a System of Great Schools*. Philadelphia, PA: School District of Philadelphia.
- School District of Philadelphia. (2012). *Code of Student Conduct (2012-13)*. Philadelphia, PA: School District of Philadelphia.
- School District of Philadelphia. (2015, December). *About Us - School District of Philadelphia*. Retrieved March 21, 2016, from <http://www.philasd.org/about/#charter-schools>.
- School District of Philadelphia. (2016). *School Police Officer Job Description*. Philadelphia, PA: School District of Philadelphia.
- School District of Philadelphia & University of Pennsylvania Graduate School of Education. (2016a). *District-wide Parent/Guardian Survey 2014-2015*. Philadelphia, PA: School District of Philadelphia and University of Pennsylvania Graduate School of Education.
- School District of Philadelphia & University of Pennsylvania Graduate School of Education. (2016b). *District-wide Student Survey: 2014-2015*. Philadelphia, PA: School District of Philadelphia and University of Pennsylvania Graduate School of Education.

- Schreck, C. J., Miller, J. M., & Gibson, C. L. (2003). Trouble in the school yard: A study of the risk factors of victimization at school. *Crime & Delinquency*, 49(3), 460–484.
- Shadish, W. R. (2012). Propensity score analysis: promise, reality and irrational exuberance. *Journal of Experimental Criminology*, 9(2), 129–144.
- Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, 103(484), 1334–1344.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth Cengage Learning.
- Shaw, M. (2004). *Police, Schools, and Crime Prevention: A Preliminary Review of Current Practices*. Montreal, Canada: International Centre for the Prevention of Crime.
- Shobo, Y., Wong, J., & Bell, A. (2014, September). *Estimating the Impact of the PROMISE Scholarship Using Propensity Score Weighted Frontier Fuzzy Regression Discontinuity Design*. Presented at the Society for Research on Educational Effectiveness Fall Conference, Washington, D.C.
- Skiba, R. J., Chung, C.-G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, 51(4), 640–670.
- Skiba, R. J., & Peterson, R. (1999). The dark side of zero tolerance. *Phi Delta Kappan*, 80(5), 372–382.
- Snyder, S. (2002, September 4). Phila. promises tougher school discipline policy. *Philadelphia Inquirer*. Philadelphia, PA.
- Snyder, S. (2011a, June 19). Budget cuts pose a threat to safety in Philadelphia schools. *Philadelphia Inquirer*, p. B03. Philadelphia, PA.
- Snyder, S. (2011b, September 16). Philly schools modify “zero tolerance” policy. *Philadelphia Inquirer*. Philadelphia, PA.
- Snyder, S., Sullivan, J., Graham, K., & Purcell, D. (2011, March 27). Underreporting hides violence. *Philadelphia Inquirer*. Philadelphia, PA.

- Socular, P. (2009, June 1). Imagine 2014 gets a green light from the SRC. *Philadelphia Public School Notebook*. Philadelphia, PA.
- Somers, M.-A., Zhu, P., Jacob, R., & Bloom, H. (2013). *The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation*. New York, NY: MDRC.
- Sullivan, J., Snyder, S., Graham, K. A., & Purcell, D. (2011, March 27). Climate of violence stifles city. *Philadelphia Inquirer*. Philadelphia, PA.
- Sweeten, G. (2006). Who will graduate? Disruption of high school education by arrest and court involvement. *Justice Quarterly*, 23(4), 462–480.
- Theriot, M. T. (2009). School resource officers and the criminalization of student behavior. *Journal of Criminal Justice*, 37(3), 280–287.
- Thompkins, D. E. (2000). School violence: Gangs and a culture of fear. *The ANNALS of the American Academy of Political and Social Science*, 567(1), 54–71.
- Thurau, L. H., & Wald, J. (2010). Controlling partners: When law enforcement meets discipline in public schools. *New York Law School Law Review*, 54(4), 977–1020.
- Todd, D. M. (2010, April 22). Police called 68 times to Ben Fairless Elementary School. *Pittsburgh Post-Gazette*.
- Travers, E. (2003). Philadelphia school reform: Historical roots and reflections on the 2002-2003 school year under state takeover. *Perspectives on Urban Education*, 2(2).
- Trepanier, J. (2014). *Methodologies for an Integrated Use of Administrative Data in the Statistical Process*. Ottawa, ON: Statistics Canada.
- United States Department of Justice. (2011, July 21). *Attorney General Holder, Secretary Duncan Announce Effort to Respond to School-to-Prison Pipeline by Supporting Good Discipline Practices*. Retrieved from <http://www.justice.gov/opa/pr/attorney-general-holder-secretary-duncan-announce-effort-respond-school-prison-pipeline>
- Vera Institute of Justice. (1999). *Approaches to School Safety in America's Largest Cities*. New York, NY: Vera Institute of Justice.
- Verdugo, R. R. (2002). Race-ethnicity, social class, and zero-tolerance policies: The cultural and structural wars. *Education and Urban Society*, 35(1), 50–75.
- Vermeire, D. T., Merluzzi, N., & Ridolfi, L. J. (2013). *Balancing the Scales of Justice: An Exploration into How Lack of Education, Employment, and Housing*

Opportunities Contribute to Disparities in the Criminal Justice System. San Francisco, CA: American Civil Liberties Union of Northern California and W. Haywood Burns Institute.

- Viswanathan, M. (2005). *Measurement Error and Research Design.* Thousand Oaks, CA: SAGE Publications.
- Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development, 99,* 9–15.
- Welsh, W. N. (2000). The effects of school climate on school disorder. *The ANNALS of the American Academy of Political and Social Science, 567*(1), 88–107.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*(2 Suppl), S130–S143.
- Wing, C., & Marier, A. (2014). Effects of occupational regulations on the cost of dental services: Evidence from dental insurance claims. *Journal of Health Economics, 34,* 131–143.
- Wong, M., Cook, T. D., & Steiner, P. M. (2015). Adding design elements to improve time series designs: No Child Left Behind as an example of causal pattern-matching. *Journal of Research on Educational Effectiveness, 8*(2), 245–279.
- Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management, 27*(1), 122–154.
- Youth United for Change. (2011). *Pushed Out: Youth Voices on the Dropout Crisis in Philadelphia.* Philadelphia, PA: Youth United for Change.
- Youth United for Change, Advancement Project, & The Education Law Center. (2011). *Zero Tolerance in Philadelphia.* Philadelphia, PA.