Out of Class and Off Track: High School Suspension in New York City

Elizabeth Chu

Submitted in partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Out of Class and Off Track: High School Suspension in New York City

Elizabeth Chu

Research suggests that suspensions are detrimental to students' socio-emotional and academic development and to the likelihood that they graduate from high school. This literature often describes the types of students who are suspended or the relationships between suspension and student outcomes, either through qualitative methods or quantitative methods that fail to adequately account for the differences between students who are and are not suspended. Using longitudinal administrative data from New York City, I build upon extant research by estimating the link between suspension and short-term academic outcomes within a student fixed effects framework. This approach eliminates unmeasured differences across students that are associated with both the likelihood of being suspended and important student outcomes. I then estimate the link between suspension and long-term academic outcomes using propensity score matching. I find that suspension is associated with increased absences and latenesses, and a decreased likelihood of passing courses the term in which the suspension occurred. Furthermore, suspension is associated with a lower likelihood of graduating within four, five, or six years, and a decreased likelihood of passing Regents exams.

Table of Contents

Table of Figures

Table of Tables

Chapter 1. Introduction

Exclusionary discipline policies, including out-of-school suspension (OSS), in-school-suspension (ISS) and expulsion, are commonly part of school discipline plans (Fenning & Bohanon, 2006). In 2007, approximately one-quarter of high school students had ever been suspended, and 3 percent expelled (Aud, KewelRamani, & Frohlich, 2011). These consequences are designed to be aversive: they aim to punish students, deter them from exhibiting the same behavior in the future, and discourage their peers from engaging in similar conduct (Bear, 2000). By excluding students from the school environment, the school enacts a severe punishment that communicates intolerance for the misconduct and ostensibly protects students and staff from further disruptive behavior with the aim of creating a safe and orderly academic environment (Bear, 2000; Costenbader & Markson, 1998).

However, the fairness and efficacy of exclusionary tactics have been questioned for decades. Empirical studies paint a disquieting picture in which underrepresented groups are disproportionally suspended, suspensions are most often a result of subjective offenses, and suspensions are associated with additional negative student outcomes, including increased recidivism, school dropout and incarceration rates. The Children's Defense Fund's 1975 report, *School Suspensions: Are they Helping Children?* brought disproportional suspension rates to the nation's attention. This report revealed startling statistics about the demographics of suspension, the arbitrary nature of its use, the procedures through which suspensions are enacted, and the resulting educational deprivation (Children's Defense Fund, 1975). Following this alarming report, academics began investigating relationships between suspension and student and school demographics, student outcomes, and family and neighborhood risk factors.

We know that African American students receive disciplinary referrals and are suspended and expelled at higher rates than any other racial/ethnic group (Costenbader & Markson, 1998; Krezmien, Leone, & Achilles, 2006; Raffaele Mendez, Knoff, & Ferron, 2002; Theriot, Craun, & Dupper, 2010; Zhang, Katsiyannis, & Herbst, 2004). In 2007, approximately half of African American high school students had ever been suspended, compared to only 18 percent of white students (Aud, KewelRamani, & Frohlich, 2011). Furthermore, the black student exclusionary discipline rate increased between 1991 and 2005, even though rates for other racial/ethnic groups decreased during the same period (Wallace, Goodkind, Wallace, & Bachman, 2008).

These descriptive findings, however, raise questions themselves. Little consensus exists about why underrepresented groups are disproportionally suspended, and whether the associations between suspensions and other negative outcomes are indeed an effect of suspensions, or are instead attributable to the characteristics of students who are suspended in the first place. The black-white discipline gap cannot be explained by poverty, neighborhood characteristics, or academic achievement. Studies that control for socioeconomic status and students' academic capabilities still find evidence for a racial/ethnic contribution to the likelihood of suspension (Wehlage & Rutter, 1986; Theriot, Craun, & Dupper, 2010). Students with more cumulative risk factors for violent and antisocial behavior are more likely to have negative long-term academic outcomes; however, researchers have yet to establish that accounting for a wide range of risk factors removes racial/ethnic gaps in suspension rates (Walker & Sprague, 1999).

Extant research fails to substantiate higher rates of objectively unsafe behavior by black students (Dinkes, Cataldi, & Lin-Kelly, 2007), yet studies have found evidence that black students are more frequently punished for subjectively inappropriate behavior. These findings give rise to hypotheses of differential selection in exclusionary discipline (Gregory, Skiba, & Noguera, 2010). The black-white discipline gap has been explained using cultural mismatch theories, which maintain that conflict exists between teachers and students as a result of the tension between the teacher's majority and student's minority culture (Fenning & Rose, 2007; Gay, 2006; Irvine, 2002; Townsend, 2000). Other explanations employ bias theories, which propose that white teachers' discipline may be influenced by media stereotypes of black aggressive students. Evidence for these theories in quantitative research is not conclusive (Kinsler, 2011; Theriot, Craun, & Dupper, 2010).

Special education students, too, are suspended at higher rates than their peers (Krezmien, Leone, & Achilles, 2007; Leone, Mayer, Malmgren, & Meisel, 2000; Rafaelle Mendez, 2003; Skiba, 2002; Theriot, Craun, & Dupper, 2010; Zhang, Katsiyannis, & Herbst, 2004), as are low socioeconomic students (Christle, Nelson, & Jolivette, 2004; Rafaelle Mendez, 2003; Rafaelle Mendez, Knoff, & Ferron, 2002; Skiba, Peterson, & Williams, 1997; Theriot, Craun, & Dupper, 2010) and students with poor behavioral and academic track records (Arcia, 2006; Christle, Nelson, Jolivette, 2004; Morrison, Anthony, Storino, & Dillon, 2001; Theriot, Craun & Dupper, 2010). Suspended students also are more likely to have family or home life problems, have lower levels of self-reported personal optimism, and have weaker senses of social responsibility than students who have never been suspended (Morrison, Anthony, Storino, & Dillon, 2001).

These socio-demographic and academic discipline gaps become more troubling when considering the lack of research that establishes the efficacy of suspension. Research has not shown that suspension improves students' behavior or increases school safety (for a review, see Skiba, 2000). Students themselves do not believe that suspension decreases problematic behaviors (Costenbader & Markson, 1998), and previous suspensions are a strong predictor of future suspensions and antisocial behavior (Costenbader & Markson, 1994; Hemphill et al., 2006; Tobin, Sugai, & Colvin, 1996). In addition, suspension is negatively related to student achievement (Arcia, 2006), and is positively associated with dropping out of school, both at the student level (Arcia, 2006; Bradley & Renzulli, 2011; Raffaele Mendez, 2003; Skiba & Peterson, 1999; Suh & Suh, 2007; Suh, Suh, & Houston, 2007; Wehlage & Rutter, 1986) and at the school level (Christle, Jolivette, & Nelson, 2007; Lee, et al., 2011). While concerning, these findings raise questions about whether these negative outcomes are a result of student characteristics that influence numerous negative outcomes, including infractions leading to suspension, or a result of the use of exclusionary discipline.

Indeed, suspension may reinforce the problematic behaviors that invoked the suspension, as suspended students may prefer to be excluded from the academic environment they perceive to be irrelevant and in which they feel unwelcome (Tobin, Sugai, & Colvin, 1996). Away from school, suspended students may be surrounded by influences that increase their exposure to the negative behaviors that led to suspension and that are also associated with negative long-term academic outcomes (Walker et al., 1996). While suspended, students lose access to valuable learning opportunities (Townsend, 2000). Upon returning to school, suspended students have decreased trust in administrators and teachers and feel less engaged in the school environment (Arcia, 2006; Brown, 2007; Reyes, 2006). They may demonstrate increased rates of misbehavior, thereby escalating the likelihood of multiple suspensions (American Psychological Association Zero Tolerance Task Force, 2008; NAACP Legal Defense and Education Fund, 2005). Recidivism may ultimately result in negative long-term outcomes: students who are suspended multiple times are more likely to drop out and to be involved in the criminal justice system as adults than non-repeating offenders (ACLU & NYCLU, 2007; NYCLU, 2011).

Purpose and Outline of the Study

This evidence on the problematic use and effects of exclusionary discipline policies warrants additional research on suspensions. My work moves beyond much of the extant research in four important ways. First, many studies on characteristics associated with suspension ignore the nested nature of the data. In my work, I will use multi-level models to account for the fact that incidences of exclusionary discipline are nested within students, and students are nested within schools over the course of their high school career. Second, much research does not account for the fact that students who are suspended in high school likely differ in important but unmeasured ways from students who are never suspended. For this reason, exploring relationships between suspension and social, academic, or behavioral outcomes without accounting for these differences is misleading, and may overstate the detrimental effects of suspension. To avoid this pitfall, I will employ multivariate student fixed-effects approaches with adaptive centering that remove unmeasured differences across students, thus producing more robust estimates. I will also estimate the relationship between suspension and long-term academic outcomes using multi-level propensity score matching.

Third, I utilize outcomes that provide evidence of the relationship between suspension and high school outcomes, including credits earned, attendance, and graduation. These estimates highlight the educational deprivation that suspended students face. Educational deprivation refers to the diminished student outcomes suspended students are expected to achieve, on average, within and across semesters. Finally, unlike many previous studies that have relied on small, random student samples, I will employ data on the entire population of a cohort of New York City public high school students. Doing so eliminates cross-district policy differences that influence both suspensions and other student academic and behavioral outcomes. A focus on New York City also affords the opportunity to explore the associations between school suspensions and student outcomes among tens of thousands of students and several hundred schools in the nation's largest school system.

Specifically, my analyses are designed to address the following research questions:

1. What student academic and socio-demographic characteristics are associated with suspension in New York City?
2. What school characteristics are associated with suspension in New York City?
3. How do school characteristics influence the relationship between student characteristics and suspension?
4. What is the relationship between suspension and short-term outcomes, including attendance and likelihood of passing core classes?
5. What is the association between suspension and long-term academic outcomes, including graduation and Regents Exam performance?

I begin my work by reviewing the literature on suspensions and building a conceptual framework for my analysis. To build the conceptual framework for my empirical models, I draw upon Bronfenbrenner's (1979) *The Ecology of Human Development.* This theory posits that children's development is a result of the way children perceive and interact with their environment, which can be understood as a series of nested structures. Although my conceptual framework and review of the literature takes a broader view of the suspension process than I am able to investigate in my empirical work, this breadth informs the methodologies used to analyze the data available to me, as well as the conclusions drawn from my results. Accordingly, I then proceed to describe the methods I use to investigate my research questions. To respond to my first three research questions, in Chapter 4 I employ two-level models, with students nested within high schools. This chapter establishes the fact that students who are assigned suspensions differ from students who are never suspended in many observable ways.

I then move beyond the examination of characteristics that predict suspension to an analysis of the associations between suspension and short-term outcomes, such as course credit accumulation, attendance, and the likelihood of passing classes. Again, I use multilevel models to account for the nested nature of the data, using student fixed effects with adaptive centering to better estimate the associations between suspension and student outcomes. In my final multivariate analysis, I employ school fixed effects models and multilevel propensity score matching to estimate the effects of suspension on four-year high school outcomes, including graduation status and likelihood of passing Regents exams. I conclude by commenting on the changes made in the NYCDOE discipline policy in the years after my analysis and by reflecting on the significance and limitations of my work, both in terms of scholarly contribution and in terms of policy implications.

My findings reveal that that there are negative associations between suspension and student outcomes. However, some of the educational achievement gaps that are associated with suspension are not as large as one would expect. Students who are suspended are only slightly less likely to pass classes and somewhat less likely to pass their Regents exams, yet they are considerably less likely to attend school or graduate. When these academic deficits are combined with the statistics on the overrepresentation of poor and minority students among the suspended population, questions about the efficacy of suspension become more complex. Importantly, my results suggest that neither suspensions, nor our schools, are adequately serving these students who are at high risk for negative life outcomes. The fact that the gaps in educational achievement are large between students who are and are not suspended, but the outcomes gaps among students who are suspended are narrow, suggests that the education and services provided to these students once they are in high school are not adequate for turning around their negative trajectories.

Examining these findings through the lens of the principles of punishment for criminal law raises questions about the true intent of suspension. According to these principles, there are four main reasons for punishment: general deterrence, which means that the consequence deters non-offenders from committing the offence in the future; specific deterrence, which means that the consequence makes it less likely that the offender will commit the offence in the future; retribution, which means that the consequence is the offender's means of paying for the committed offence; and rehabilitation, which means that the consequence reforms the offender so that the offender will not commit an offence in the future (Banks, 2013). In many ways, my findings underscore an important realization that suspensions may be only designed to serve those students who are not suspended; the goals of rehabilitation and specific deterrence go unrealized. By suspending students, administrators may be attempting to preserve the learning environment for non-offending students and dissuade other students from committing similar acts, even though doing so comes at the cost of those students who are excluded from school.

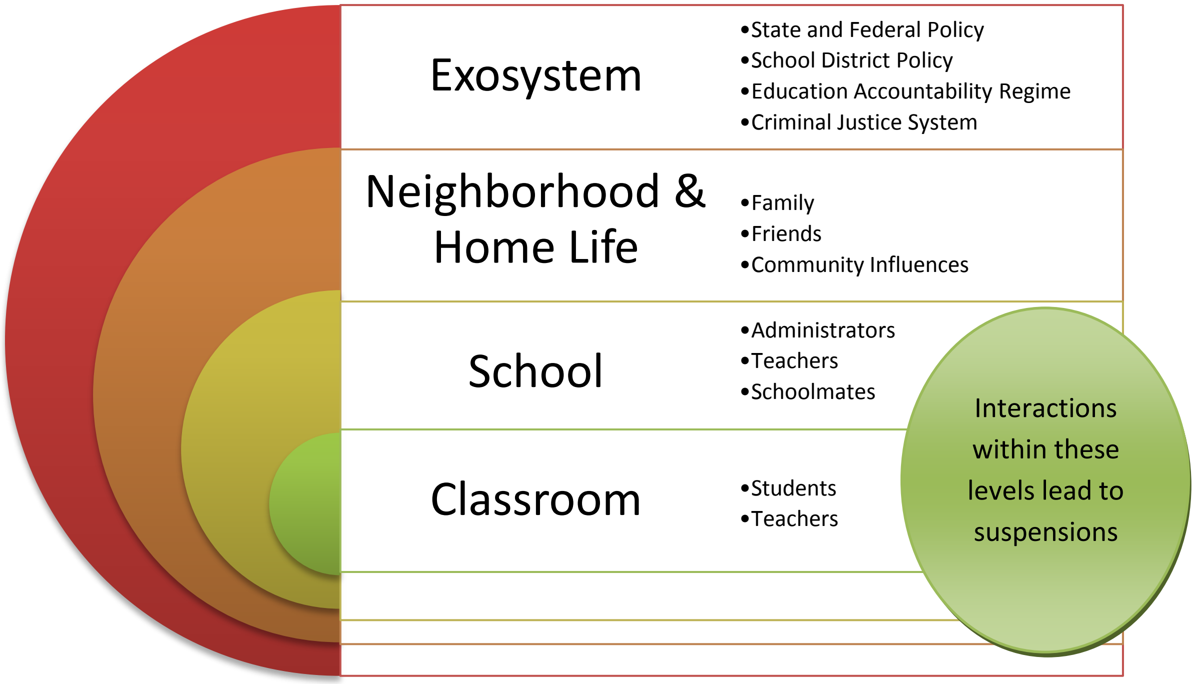
My work also suggests areas for future research, such as preventative measures schools might take to address the needs of students with histories of discipline infractions, and research on the access to and quality of education students receive once they are sentenced with a long-term suspension and sent to alternate learning centers. Furthermore, future research may also attend to the reintegration of students into their home schools following a long-term suspension, and the way that schools attempt to make suspension a policy that benefits both school safety and the offending student. This research is necessary if we are to better understand how to meet the needs of all students through high school discipline policies.

Chapter 2. Conceptual Framework and Literature Review

The process by which students are assigned suspensions is complicated, as students' behavior is influenced by their personal characteristics, family and home environments, school and work histories, and current school contexts. To be assigned a suspension, students need to interact with at least one administrator or teacher, and they often interact with fellow students as well. These interactions add further complexity to the suspension process because these individuals' behavior, like that of students, is influenced by their personal and professional histories and contexts. Moreover, these interactions, and their subsequent interpretation, are based upon perceptions of the other and the context in which the behaviors occur, thereby removing much objectivity from the exclusionary discipline process. From a researcher's perspective, the lack of objectivity is complicated by the fact that the adults involved in the disciplinary incidents have control over what discipline infractions are documented.

Because perception, interaction, and environment are central to the process by which students are assigned suspensions, I draw upon Bronfenbrenner's Ecological Theory of Human Development to construct my conceptual framework (see Figure 1). Under this theory, individuals are located within a series of nested environments. The most immediate environment surrounding the individual is the microsystem. As conceptualized in my model, the microsystem is broken down into three levels: the neighborhood, the school, and the classroom. Since the focus is on suspensions, I place home life within the neighborhood tier of the microsystem. Within each level are individuals who are central to the setting, including family members and neighbors, teachers and administrators, and friends and classmates. These immediate settings are connected through the mesosystem, which are interrelations among two or more environments, such as connections between students' home and school life.

Figure 1: Conceptual Framework for Suspensions



Within microsystems, students' personal characteristics affect how others perceive them and influence how students perceive their setting and the people with whom they interact. These personal characteristics include socio-demographic factors, such as socioeconomic status, race/ethnicity, and gender, as well as personality traits and perceptions of self, such as academic self-concept and levels of perseverance and self-control. Students' actions in school also may be a result of the mesosystem: the connections between their neighborhood and home environment and the school. For example, at home students may have learned behaviors that may conflict with behavioral norms at school, or students may be exposed to risk factors that are often associated with increased rates of suspension. In addition, students' behavior may be affected by students' perception of the school environment, such as the relative order and safety of the school setting.

Students' behavior, alone, however, cannot explain the occurrence of suspensions. Even after a problematic behavior occurs, teachers and administrators need to make sense of the action and mete out punishment, making suspensions more subjective than they may initially appear. Not only do staff members have the power to determine whether behavior warrants suspension, but they also may interact with students in such a way that makes it more likely that students exhibit negative behavior. In many ways, teachers can become "street level bureaucrats" in this setting: they have the power to interpret the discipline policy set by the school and district and enforce it as they see fit (Lipsky, 1980; Maynard-Moody & Musheno, 2003). The way teachers and administrators enact these policies, therefore, is not just a result of the objective policy, but is also a result of their subjective interpretation of the rules, the students, and the classroom and school context. Thus, like students, staff members' actions are a result of a convergence of factors from their personal and professional lives, including school-level factors, such as the socio-demographic composition of the student body and the reputation of the school.

The students' microsystem is nested within broader exosystems, which consist of settings in which students do not actively participate, but that affect the students' microsystem. Within my framework, this includes the school district and the broader policy context. District discipline policy circumscribes the actions the school administration takes in response to students' behavior. This district discipline policy is influenced by educational policy enacted on state and federal levels, as well as policy located outside of education. For example, the exosystem also includes the criminal justice system. For some students, aspects of the criminal justice system may shift to become part of their microsystem, if after being assigned suspensions they commit infractions that lead to their entry in the juvenile justice system. The persistence of the link between the criminal justice system and the school gives rise to the conceptualization of the School-to-Prison Pipeline. This theory posits that minority and disadvantaged students are pushed out of the school system through exclusionary discipline policies and a lack of services tailored toward their individualized needs, and pushed into the criminal justice system (ACLU & NYCLU, 2007; NYCLU, 2011). Once in the justice system, these students are prone to recidivism due to low levels of education, to high exposure to risk factors, and to disenfranchisement with the public education system.

Enacted suspensions can affect students' social and academic development. Under Bronfenbrenner's theory, development occurs as children shift the way they perceive, relate to, and interact with their environments. Exclusionary discipline infractions can affect students' development by inciting this shift and by changing their academic microsystems. Each time suspensions are assigned, students lose instructional time in their traditional school settings. Depending on the length of suspension, students may be provided with instruction from an outside source. Following suspension, students may perceive themselves and their teachers and administrators in a different way. Students with multiple infractions may feel unwelcome in the academic setting. The act of suspension and its subsequent effects on academic outcomes then becomes folded into students' and teachers'/administrators' histories, often increasing the likelihood of recidivism, as students may become disenfranchised with the school system and teachers and administrators may be more likely to view students' behavior in a harsh light.

In the review that follows, I use my framework to summarize literature relevant to the suspension process. I begin with a discussion of the exosystem, as this allows me to review the socio-political trends on state and federal levels that influence district and school discipline policy. I then narrow in on the student- and school-level predictors of suspension, as these are characteristics that are relevant to the inner levels of the suspension process. Finally, I turn my attention to the student-level effects of suspension.

The Exosystem: Theories of District, State, and Federal Policy that Shape School Discipline

The occurrence of suspensions and the rise in the use of exclusionary discipline is a result of the exosystem, the district or state level where discipline policy is outlined. The school discipline exosystem is not restricted to education policy alone. Rather, policies that affect the classroom may be situated within the broader political landscape, especially with regard to the politics of crime control. David Garland (2001) argues that social crises in the 1960s and 1970s transformed the political order of the United States to one shaped by fear and crime. Civil disorder and inflation in these decades threatened the dominant governance structure at the time, one rooted in the policies of the New Deal and that was based upon principles of collective risk spreading and welfarism. In response to the social crises, a new governance model emerged that emphasizes personal responsibility and strictly enforced criminal law, along with increased rhetoric surrounding the necessity of harsh crime control measures. Garland posits that politicians from both sides of the aisle used crime to further their agenda, whether it be to enforce legal segregation or to create social welfare programs that invested in poor communities. The resultant "war on crime" used law enforcement to crack down on illegal activity and created a crime prevention and security sector that included public schools. This third sector is necessary to the "culture of control" - the need to manage crime risk - that still pervades society.

Jonathan Simon also places school discipline within the context of crime control. He argues that since the 1960s, political leaders have governed through crime. Governing through crime occurs when policymakers use rhetoric and the media to heighten the awareness that all citizens are potential victims of crime, and that the government's role is to protect its constituents from their fellow citizens, each of whom is conceivably a criminal (Simon, 2007). In his view, this shift began in earnest under Kennedy's administration with Robert F. Kennedy at the helm of the Department of Justice. As Attorney General, RFK became "America's prosecutor:" he made criminal law central to his appointed position. He both waged a war on organized crime and used crime to address social issues central to his and JFK's political agenda. For example, his Mobilization for Youth program attempted to reduce juvenile delinquency, and his pursuit of bail reform aimed to make one's release on bail less dependent on wealth. This shift crystallized with the passage of the federal Omnibus Crime Control and Safe Streets Act of 1968, which was voted on the day after RFK's assassination and received broad bipartisan support. The justification of the law depends upon viewing Americans as potential crime victims who can unite despite differences and take political action: it "imagine[s] the needs of the citizenry as framed by the problem of crime, the purposes and means of intervention, and the means of achieving a higher level of success against crime" (p. 75). This bill and its accompanying rhetoric solidified the perception that policymakers should advocate on behalf of victims and law enforcement. Representing the interests of criminals or prisoners, individually or collectively, is perceived as soft on crime and as harmful to potential victims or to law enforcement.

Simon applies his theory to public schooling. He argues that schools have been criminalized by links between youth culture and drugs, by references to youth violence from the 60s and 80s, and by rhetoric from right-wing conservatives who are eager to call attention to the flaws in the prevailing school governance model. Since the 80s, schools have increasingly addressed discipline using measures frequently associated with crime control, such as metal detectors, uniformed school safety officers, security guards, and id scanning machines. Simon theorizes that Goals 2000 - through the sixth goal that links drugs, violence and school discipline - and the Safe Schools Act of 1994 helped establish the current regime of governing the schools through crime, much like the Omnibus Crime Control and Safe Streets Act did on the broader policy scale. The Safe Schools Act incentivizes community support for school discipline programs and data collection on school discipline because schools need to demonstrate a school violence problem in order to qualify for competitive grants. Additionally, to qualify for these federal funds, districts are required to have formal discipline codes that detail exclusionary polices and that provide evidence of the working relationship between the school and police and juvenile justice agencies.

Since the passage of Goals 2000, governing bodies have amassed increasing amounts of information on school discipline, which allows school community members, policymakers, and the media to draw attention to rates of school violence. Importantly, this coverage appears unrelated to the frequency and objectivity of violent offenses. Rather, media coverage appeals to readers' emotional response, raising alarms about schools' current inability to proactively prevent major disciplinary problems (Kupchik & Bracy, 2009). This attention increases fears that dangerous incidents can occur in all schools, including those attended by primarily white, middle class families (Herda-Rapp, 2003; Kupchik & Bracy, 2009; Lawrence & Mueller, 2003), which, in turn, facilitates governing through crime, as parents support measures that promise to ensure the safety of their children.

These policies and data are to be used to inform school safety plans that outline measures taken to reduce disciplinary infractions and that articulate school-specific goals. The federal government suggests that schools receiving federal funds use specific programs to achieve their violence-reduction goals. Simon sees many of these programs as actually bolstering the connection between crime and schools because they encourage the use of criminal enforcement measures – e.g. installing metal detectors – and call attention to the fact that drugs or weapons could be found within schools – e.g. establishing “Drug and Weapon Free School Zones” (p. 219). Many states, in turn, have created analogous Safe Schools Acts to reinforce the commitment to decreasing school violence, largely through data collection and increased security measures.

When schools are governed through crime, practices typically ascribed to the criminal justice system pervade schools. These practices move beyond the obvious parallels of metal detectors, x-ray machines, and school security agents, to policies such as school uniform codes, which the federal Department of Education explicitly states is a means of decreasing violence, theft and gang activity. Even the No Child Left Behind Act of 2001 (NCLB), the latest reauthorization of the Elementary and Secondary Education Act of 1965, can be examined through Simon’s lens of governing through crime. In promoting this bill, President George W. Bush rhetorically joined crime and school: he characterizes students in failing schools as victims and principals and teachers as perpetrators who deserve punishment. The power of the legislation relies upon consequences: schools that do not improve face punishment for their crime of failing to educate students. This legislation also explicitly addresses school safety through its provision regarding persistently dangerous schools, which enables students in schools labeled as “persistently dangerous” to exercise school choice and move to a safer school (Lindle, 2008).

In *Homeroom Security*, Kupchik (2010) builds upon Garland’s and Simon’s theories to take a more expansive view of the current state of school discipline. He sees the rise of punishment and security as a result of broader insecurities regarding the state of U.S. public education. He argues that the public uses schooling as a forum to debate and act upon social conflicts and anxieties, all of which lead society to question the efficacy of our school system.

For example, racial conflicts that pervade society impact education policy and individuals' schooling choices. This relationship is obviously seen in landmark events, such as *Brown v. Board of Education*, but is also apparent in white flight away from urban public schools and the continued prevalence of de facto segregation. Like Garland and Simon, Kupchik also notes that the public's fear of crime and anxiety over youth morality relates to our disquiet over the state of schooling. As with racism, major events, such as the 1999 Columbine tragedy, contribute to the perception that schools are in crisis, as do commonly held concerns by older generations that youths are less moral today than they were a few decades earlier. Finally, Kupchik posits that shifts in family structures, such as an increasing number of women in the workforce, and concentrated poverty, affect the public's view of public education. Together, these concerns lead to low confidence in public education, which, in turn, facilitate increased accountability measures, such as those mandated in NCLB, and increased use of punishment and security in school discipline policy.

Zero Tolerance Discipline. One manifestation of governing through crime is the increasing adoption of zero tolerance discipline policies, which are policies that mandate the use of exclusionary discipline upon first offense, regardless of how severe the infraction is or the context surrounding the noncompliance (Johnson, 1999). Districts and schools increasingly adopted these polices after the 1994 Safe Schools Act. They were put in place based on five key assumptions: (1) school violence is uncurbed and increasing, (2) zero tolerance would create more consistent and clear discipline, (3) removing offending students would create a better learning environment for non-offending students, (4) the enactment of zero tolerance consequences would deter future misbehavior, and (5) parents support zero tolerance (APA Zero Tolerance Task Force, 2008).

The Zero Tolerance Task Force (2008) concluded that all five of these assumptions were inconsistent with extant discipline data and research on the efficacy of zero tolerance policy. School violence is not on the rise; rates of discipline vary across schools as a result of many characteristics; suspension rates are positively related to the amount of time schools spend on discipline matters; and suspension rates are negatively related to academic achievement (APA Zero Tolerance Task Force, 2008; DeVoe et al., 2004; Skiba, Peterson, & Williams, 1997, National Center for Education Statistics, 2006; Scott & Barrett, 2004). Furthermore, suspension is negatively related to the likelihood of another exclusionary discipline incident, of school dropout, and of delayed graduation (Bowditch, 1993; Costenbader & Markson, 1998; Raffaele Mendez, 2003; Tobin, Sugai, & Colvin, 1996; Wehlage & Rutter, 1986). Zero tolerance has also been implicated in the black-white discipline gap, as increased rates of suspension associated with zero tolerance coincide with disproportional discipline rates of African American students (APA Zero Tolerance Task Force; Costenbader & Markson, 1998; Raffael Mendez & Knoff, 2003; Skiba et al., 2002).

Within the past five years, the merits and legality of zero tolerance policies have increasingly been called into question. For example, in 2008, as a result of North Carolina zero tolerance discipline policies, two female students were suspended for five months for a fist fight that did not involve weapons or injuries (Eckholm, 2010a). The case went to the North Carolina Supreme Court, which ruled that schools must provide strong justification for applying exclusionary discipline consequences and denying alternative schooling or tutoring to students once they are suspended (Eckholm, 2010b). In New York City, the use of zero tolerance discipline rose between 1998 and 2007, with only seven zero tolerance infractions included in the 1998 discipline code, but 29 in the 2007 code. In 2011, New York City, under growing pressure to reduce its suspension rate, decreased the number of zero tolerance infractions to 21 (Santos, 2011).

School Resource Officers. Another expression of this punitive school discipline era is increasing police presence in schools. Commonly known as school resource officers (SROs), these authority figures represent a marked shift in the role that the justice system plays in school discipline, as SROs are typically trained in law enforcement rather than education and are accountable to local police agencies, not to school boards or departments of education (Brown, 2006). The duties of these officers vary across districts and schools, but often include helping control disruptive students, upholding school rules and regulations, providing educational programs designed to reduce juvenile delinquency, acting as an intermediary between schools and local law enforcement agencies, and making police presence visible to students (Brown, 2006). Proponents of a police presence in schools argue that officers increase safety, serve as role models that can make educational and disciplinary contributions to students' education, and improve students' perceptions of police. Opponents, however, see SROs as negatively contributing to the school environment by making school culture unnecessarily punitive and by imposing limits on students' legal rights that are unique to the schooling situation (Bracy, 2010). For example students in school have less protection from search and seizure than they would outside of school or as an adult. SROs are able to search a student or his/her possessions without a warrant based on suspicion that students have committed a disciplinary infraction, even one that would be legal outside of school grounds, such as possessing a cell phone (Brown, 2006).

The degree to which SROs improve school safety is yet to be established. Evaluating the efficacy of SROs is difficult because it is challenging to disentangle the effects of SROs on school discipline from other contributing factors to school violence and disorder. Furthermore, discipline-related statistics can be used both to argue for and against an increasing police presence in schools. For example, statistics indicating increasing crime could be viewed as evidence that more SROs are needed or it could be taken as proof that SROs are ineffective at decreasing disciplinary infractions (Brown, 2006). Moreover, increasing rates of juvenile disorder may be a result of increased reporting of infractions due to the police presence in schools: since SROs are charged with upholding rules and regulations, infractions that went unrecorded or unaddressed before the police presence may now be acted upon and included in school discipline and juvenile crime statistics (Brown, 2006). Finally, juvenile court records confidentiality laws complicate efforts to obtain accurate counts of arrests made by SROs (Theriot, 2009).

Empirical work presents a mixed analysis of school resource officers (e.g. Beger, 2003; Jennings, et al., 2011; May, Fessel, & Means, 2004; Mukherjee, 2007). For example, while Theriot (2009) found that SROs are positively related to arrests for disorderly conduct and negatively related to arrests for assault and weapons charges, he found no relationship between SROs and total arrests in schools. Using structural equation modeling, Mayer and Leone (1999) concluded that the use of discipline personnel, such as security guards, is associated with higher levels of disorder within a school, and suggest that the relationship between increased use of security personnel and heightened disorder may be cyclical, meaning that increased disorder may lead to increased security personnel, which may, in turn, lead to yet another increase in disorder. Johnson (1999), on the other hand, concludes that school resource officers are related to decreases in school violence and disruption based on interviews and questionnaires given at middle and high schools in a southern urban district. Results are inconsistent with regard to educational programs implemented by school resource officers as well. For instance, program evaluations of the Charlotte School Safety Program, which utilizes school resource officers to deliver classes aimed at reducing students' fear of crime, produce conflicting results, with early evaluations suggesting a positive impact on reduction of fear of crime and on students' sense of safety and later evaluations, which use more sophisticated methodology, determining the efficacy of the program is inconclusive (Kenney & Watson, 1998; Miller, et al., 2005).

Critics argue that in the absence of evidence substantiating their positive contribution to school climate, the potential danger to school culture, students' perception of the criminal justice system, and students' psychology, and the cost of implementing SROs outweigh the potential benefit of maintaining the police presence in schools. Opponents argue that relying on punitive policing measures may take attention away from preventative measures that are more developmentally appropriate for youths struggling with violence or substance abuse issues, thereby inadvertently increasing misbehavior (Hyman & Perone, 1998).

School-to-Prison Pipeline and Social Reproduction Theory. The links among the disproportionate use of exclusionary discipline and implementation of zero tolerance discipline policy and increased school security measures have given rise to the School-to-Prison-Pipeline, the theory that underrepresented students are systematically pushed out of their school environments and into the juvenile justice system, followed by the criminal justice system (Advancement Project, 2006; Children's Defense Fund, 2007; NAACP Legal Defense and Educational Fund, 2005; Wald & Losen, 2003). Disadvantaged students enter the pipeline when they enter their failing, public schools, which may suffer from dysfunctional school governance, underperforming teachers, inadequate structural and academic resources, negative peer effects, and a history of poor performance. Proponents of this theory argue that as standards and accountability began to dominate school reform and the culture of control took over school governance, schools needed a method to remove at-risk students from the classroom to boost the performance statistics and to maintain the rhetoric of safe schools with high behavioral expectations.

Given these pressures, it became easier to push students out through exclusionary discipline into alternative learning centers or to the juvenile justice system than provide them with the services they need to overcome systemic inequities (NAACP Legal Defense and Educational Fund, 2005). Indeed, exclusionary discipline predicts juvenile court referrals (Nicholson-Crotty, Birchmeier, & Valentine, 2009). Proponents of this theory argue that the conditions in which urban students learn actually prepare underrepresented students for incarceration by creating prison-like environments within schools. This happens through increased security measures, such as surveillance cameras and metal detectors, as well as by staffing schools with police officers and security guards (Beger, 2002). Transferring school discipline to police officers and security guards coincides with more severe consequences for behavioral misconduct, including zero tolerance policies that shifted the responsibility of minor infractions from teachers and administrators to police and school safety officers. This transference increases the likelihood that minor incidents are handled in juvenile or adult court rather than within the school system (Dohrn, 2001). School officers may be empowered to engage in other preemptive measures that resemble actions taken in prison. For example, they may be empowered to conduct random canine searches of students' personal property, and undercover officers may pose as students in order to investigate drug dealing within schools (Beger, 2002; Biskupic, 2000; Lait, 1999; Ransom, 1999). Indeed, Bracy (2010) writes: "Discipline and punishment dominate in these schools, and students' rights are treated as an obstacle, suggesting the logics of a carceral regime do influence how SROs and disciplinary staff work with students. Even students who are merely suspected to have done something wrong are subject to criminal treatment" (p. 309).

Just as African American students are disproportionately suspended, they are over-represented in juvenile arrests. In 2003, 45 percent of juvenile arrests were of African-American youths, even though African-American youths comprised just 16 percent of the national juvenile population (NAACP Legal Defense and Educational Fund, 2005). African Americans are also over-represented in the incarcerated population: while 30 percent of prisoners are white, fifty percent are African American (Bureau of Justice Statistics, 2007). These high incarceration rates, along with high drop-out rates among African American adults, place African American students at additional risk of being placed in foster care and of being subject to exclusionary discipline or entrance into the juvenile or criminal justice system (Bernstein, 2005; Brewer, 2007; Heitzeg, 2009; Roberts, 2004; Wildeman, 2009).

Advocates of this theory also argue the school-to-prison pipeline is fueled by the media, which perpetuates stereotypes that foment fear and biases that facilitate the school-to-prison pipeline. In the media, underrepresented minorities are overrepresented as perpetrators and inter-racial crime is over-reported, African American offenders are portrayed more negatively than are white offenders, the majority of violent crime covered is of youth under age 25, and the majority of gang violence is of African American and Latino gangs (Entman & Rojecki, 2000; Hancock, 2001; McCorkle & Miethe, 2000). Numerous studies have shown that people are biased toward believing that criminals are African American. For example, Gilliam & Iyengar (2000) showed viewers a newscast in which a photograph of the perpetrator was not shown. After watching the coverage, 70 percent of study participants recalled that the offender was black.

The school-to-prison pipeline theory could be seen as a manifestation of social reproduction theory. According to social reproduction theory, schools do not aim to provide the means through which students can overcome their backgrounds; rather, schools themselves – intentionally or otherwise – perpetuate inequity by socializing students to attain a similar social position as the one in which they were raised. Bowles and Gintis saw this reproduction mainly through economic terms. In their seminal work, *Schooling in Capitalist America,* they argue that students are treated differently in school based on the class from which they came, with students from lower classes being socialized to assume the mindsets and skills of laborers, while those of the upper classes are socialized to acquire those of positions of power and authority. Their theory extends to the degree to which students are afforded behavioral freedom, with students from upper classes being granted more freedom to roam, and students from lower classes being granted limited permission to leave class with hallway passes. Critics of school discipline and school-to-prison pipeline theorists see the punitive discipline as playing a role in perpetuating inequalities through social reproduction. Not only are the students in these underperforming schools socialized to accept the use of security and surveillance measures used outside of schools to control crime and apprehend criminals, but they are also put at greater risk of arrest due to increased exposure to law enforcement and to heightened emphasis on discipline and security within their schools.

Even without these additional security measures, suspension is connected to the cyclical nature of poverty. Children who are suspended are more likely to be socio-economically disadvantaged and are more likely to have poor academic outcomes, both with regard to the academic skills they ultimately attain and with regard to graduation and post-secondary success. Furthermore, these students are less likely to experience labor market success and are more likely to go to prison. Given these adult outcomes, these people are then more likely to have children who are also socioeconomically disadvantaged and have lower academic outcomes, thus perpetuating intergenerational inequality (Brooks-Gunn & Duncan, 1997; Jencks, et al., 1972; Mayer, 1997).

The school-to-prison pipeline and the cycle of poverty underscore the interconnections between schools and social institutions, such as child welfare and family court. Children in foster homes have increased likelihood of being assigned suspensions and more severe punishments than are their more advantaged peers (Gallagos & White, 2013; Scherr, 2007). Similarly, child welfare recipients have a higher probability of becoming involved in the justice system; those involved in both systems are known as crossover youths. Research suggests that collaboration between institutions is critical to improve the outcomes for these children, (Leone & Weinberg, 2010; Havalchek, 2009). Advocates for collaboration promote having liaisons that communicate with child welfare institutions and schools in order to keep track of student outcomes and advocate on behalf of students' differentiated needs. Increasing the degree to which juvenile defenders advocate for educational services and leverage students' education histories may also help offending students receive the educational and support services they need to avoid recidivism (Langberg & Fedders, 2013).

Student- and School-Level Predictors of Suspension

These trends in district, state, and federal discipline policy have contributed to relationships between student and school characteristics and the likelihood of suspension. These associations are the focus of my work, and represent the interactions within the microsystem and mesosystem of the suspension process.

Racial/Ethnic Gaps. The Children's Defense Fund first brought the black-white discipline gap to the nation's attention in 1975: their seminal report showed that the black student suspension rate was twice as high as that of any other group, and three times as high in some localities (Children's Defense Fund, 1975). Subsequently, numerous researches have corroborated the existence of the black-white discipline gap at elementary and secondary levels (for a review, see Gregory, Skiba, & Noguera, 2010), and have shown that African American students receive disciplinary referrals and are suspended and expelled at higher rates than any other racial/ethnic group (Costenbader & Markson, 1998; Krezmien, Leone, & Achilles, 2006; Raffaele Mendez, Knoff, & Ferron, 2002; Theriot & Craun, 2010; Zhang, Katsiyannis, & Herbst, 2004).

The National Center for Education Statistic's *America's Youth: Transitions to Adulthood* (2011) report suggests that these racial/ethnic gap findings are nationally representative. Approximately half of African American high school students had ever been suspended, compared to only 18 percent of white students in 2007 (Aud, KewelRamani, & Frohlich, 2011). The 2007 black-white expulsion gap was even larger: 10 percent of black students had ever been expelled, compared to 1 percent of white students (Aud, KewelRamani, & Frohlich, 2011). Importantly, this black-white discipline gap differs by gender (Aud, KewelRamani, & Frohlich, 2011; Costenbader & Markson, 1998; Skiba, Peterson, & Williams, 1997) with black males suspended at higher rates than white males, and black females suspended at higher rates than white females (Raffaele Mendez, Knoff, & Ferron, 2002). The suspension rate of black males is increasing: whereas 41 percent of black males had ever been suspended in 1999, 57 percent had ever been in 2007 (Aud, KewelRamani, & Frohlich, 2011).

This gap varies by grade level as well. Raffaele Mendez, Knoff, and Ferron (2002) found that within a large, diverse school district, approximately 12 percent of black males, compared to 3 percent of white males, were suspended in elementary school. In middle school, almost 50 percent of black males, compared with 25 percent of white males, were suspended. These numbers decrease at the high school level, where about 37 percent of black, versus 19 percent of white, students had been suspended at least once.

Extant research has reached less consensus surrounding exclusionary discipline gaps between American Indian and white students and Hispanic and white students (for a review, see Gregory, Skiba, & Noguera, 2010). Some studies suggest that American Indian students are suspended at disproportionate rates (Devoe & Darling-Churchill, 2008; Zhang, Katsiyannis, & Herbst, 2004), yet others suggest that this difference is non-existent or inconclusive, depending on the context or time frame for the analysis or on the small suspended-student sample size (Aud, KewelRamani, & Frohlich, 2011; Krezmien, Leone, & Achilles, 2006; Wallace et al., 2008). Similarly, national statistics show that 26 percent of Hispanic students had ever been suspended in 2007, versus 18 percent of white students (Aud, KewelRamani, & Frohlich, 2011), but multiple researchers have failed to find discrepancies between white and Hispanic exclusionary discipline rates (Costenbader & Markson, 1998; Gordon, Della Piana, & Keleher, 2000; Zhang, Katsiyannis, & Herbst, 2004). Even though findings regarding the disparities between Hispanic and white students are inconsistent, extant research does suggest that Hispanic students are excluded at lower rates than their African American peers (Aud, KewelRamini, & Frohlich, 2011; Raffaele Mendez & Knoff, 2003).

Researchers continue to struggle to explain the black-white discipline gap, much as they debate the explanation for black-white achievement gaps. Researchers are unable to explain away the black-white discipline gap using poverty, neighborhood characteristics, or academic achievement in quantitative models (for a review, see Gregory, Skiba, & Noguera, 2010). Studies that control for socio-economic status and students' academic capabilities still find evidence for a racial contribution to the likelihood of suspension (Wehlage & Rutter, 1986; Theriot, Craun, & Dupper, 2010). Students with more cumulative risk factors for violent and anti-social behavior – such as poverty; abuse and neglect; harsh parenting; exposure to drugs and alcohol; and physical, sexual, and emotional abuse – are more likely to have negative long-term academic outcomes, including increased incidences of school failure and disciplinary infractions (Walker & Sprague, 1999). However, researchers have yet to establish that accounting for a wide range of risk factors removes racial/ethnic gaps in suspension rates.

There also is little evidence for differential behavior hypotheses, which posit that black students actually do misbehave more often than their peers (Skiba, et al, 2002; for a review, see Gregory, Skiba, & Noguera, 2010). Research fails to substantiate higher rates of objectively unsafe behavior by black students (Dinkes, Cataldi, & Lin-Kelly, 2007). However, there is very limited evidence that black students may be more frequently punished for objectively unsafe or disruptive behavior. Data allowing the examination of use of exclusionary discipline after taking into consideration students' behavior is extremely rare, as behaviors deemed undeserving of punishment are not recorded. In an effort to address this question, Gastic (2013) utilized behavior-adjusted relative risk ratios (BAR) to estimate disproportionality in exclusionary discipline as a result of physical altercations. After adjusting for self-reported rates of fighting, Gastic estimates that black students were 1.7 times more likely than their white peers to be punished for fighting at school. In contrast, Latino students were found to be punished at similar rates as white students after accounting for self-reported behavior.

Other studies have found evidence that black students are more frequently punished for subjectively inappropriate behavior. These findings give rise to hypotheses of differential selection in exclusionary discipline (Gregory, Skiba, & Noguera, 2010). One study of school district disciplinary records discovered that white students were more likely to be referred to the office for objective behavioral infractions, such as vandalism or cutting class, whereas black students were more likely to be referred for subjective infractions, such as exhibiting disrespectful or threatening behavior (Skiba, et al., 2002). In addition, these authors found the gap in suspension rates between black and white students is largely explained by the disproportional office referral rate of black students, raising questions about differential selection hypotheses at the classroom level.

An ethnographic and discourse analysis of school discipline practices also found that suspension often resulted from nonviolent classroom disruptions that are highly subjective in nature (Vavrus & Cole, 2002). The authors' observations reveal that referrals and suspension often occur as a result of a sequence of events, rather than a single incident. The decision to suspend happens in a "disciplinary moment" when the teacher determines that a student's behavior has crossed a line that warrants suspension. In these moments, rather than displaying unsafe or objectively destructive behavior, students often violate the "normal order of turn taking and classroom management found in most U.S. high schools" (p. 108). This ritualized classroom management routine, however, is one established by the majority culture and may be in conflict with or less accessible to black students. Thus, Vavrus and Cole conclude that "disruptions that are interpreted by teachers as events worthy of suspension are often violations of these unspoken and unwritten rules of linguistic conduct that cannot be neatly delineated in discipline policy" (p. 91). This code violation is problematic if it privileges the behavior of white students while simultaneously putting black students at increased risk of exclusionary discipline.

Vavrus and Cole's (2002) findings are consistent with cultural mismatch theories, which are often employed to explain differential treatment based on race/ethnicity. Cultural mismatch theories look to discrepancies between students' and teachers' culture when students are from underrepresented minority groups.   These discrepancies can affect the way students learn and interact with their school environment. In the case of the black-white discipline gap, cultural mismatch theories posit that the teacher's white culture is in conflict with students' black culture, thereby creating communicative tension and conflict between teacher and students (Fenning & Rose, 2007; Gay, 2006; Irvine, 2002; Townsend, 2000). Delpit (1988, 2006) discusses this mismatch through her theory on the culture of power within schools. She makes explicit the educational power dynamic: not only do teachers have power over students, but the mainstream culture has power over students through the worldview presented in curriculum and by the dominant classes that dictate the way schools operate. As a result, the classroom is governed by explicit and implicit behavioral norms that are set by mainstream culture. Making these codes of conduct explicit to all students makes it easier for students of color to acquire power in the educational setting. What is crucial in doing so is that white teachers acknowledge the culture of power that runs throughout school.

Students who have less power are often more aware of this power dynamic than student of privileged classes (Delpit, 1988). In school, students try to use their schema to understand the presented material, and they try to use behaviors and strategies that they have deemed effective or appropriate in other setting, including their home setting (Anderson, 1988; Jenkins, 1982). If these schemata and background conflict with the mainstream culture of learning and discipline at the school, students may encounter increased difficulty in learning the material and conforming to behavioral expectations (Anderson, 1988; Franklin, 1998; McIntyre, 1993; Rueda & Forness, 1994). On its own, the struggle to effectively participate in classroom instruction may lead students to act in such a way that allows them to be excused or excluded from class (Gable, 2004; Gable, Bullock, & Evans, 2006; Shores, Gunter, & Jacks, 1993). Students who are aware of their subordinate place in the educational power structure also may knowingly rebel against the mainstream class, developing a counterculture that resists the school's authority and attempts to liberate students' from the dominant culture to which they have limited access and provide students with an alternative source of power (Willis, 1977).

Importantly, the struggle to bridge the cultural boundary between students' home life and school norms varies within racial/ethnic groups. Carter (2006) shows that breaking minority students into those who assimilate to the school environment and those who oppose it is an overgeneralization of underrepresented students' public school experience. In her work, she identifies three groups of low-income African-American and Latino students based upon their method of handling the conflict between their non-dominant culture and the white dominant culture of public schooling: cultural mainstreamers, who assimilate into the dominant culture while maintaining their racial/ethnic identity; noncompliant believers, who value education but choose not to assimilate to the mainstream culture; and cultural straddlers, who move seamlessly between the two cultures that they inhabit while maintaining a strong racial/ethnic identity. Although cultural straddlers are able to effectively function within the mainstream school environment, the noncompliant believers and cultural mainstreamers may experience disciplinary incidents because they act in opposition to school authorities or because they feel culturally marginalized.

Researchers also look to bias theories to explain differential selection given that teachers' perceptions and expectations of students inform the way in which they discipline them (Bennett & Harris, 1982). Under bias theories, white teachers' discipline may be influenced by media stereotypes of black aggressive students; by believing that black students are more inclined to dangerous behavior, white teachers may be more likely to interpret black students' behavior as threatening or noncompliant than white students (Devine & Elliot, 2000; Noguera & Akom, 2000). Closely related are theories of teacher perceptual or expectancy effects, which hypothesize that teachers' academic and behavioral expectations for students vary based on race and class, and that students' performance ultimately conforms to these expectations (Burkam, et al., 2007; Farkas, 2003; Farkas et al., 1990; Lareau & Horvat, 1999; Rist, 1970). These expectations and perceptual effects not only may be present within classrooms, but they also may be a result of school culture. Indeed Diamond, Randolph, and Spillane (2004) found that in urban schools that serve predominantly low-socioeconomic status African-American children, teachers had lower expectations for students performance, on average, and had decreased sense of ownership for their students' learning.

Special Education. Special education students are also suspended at higher rates than their peers (Krezmien, Leone, & Achilles, 2007; Leone, Mayer, Malmgren, & Meisel, 2000; Rafaelle Mendez, 2003; Skiba, 2002; Theriot, Craun, & Dupper, 2010; Zhang, Katsiyannis, & Herbst, 2004). The rate at which special education students are suspended varies both by race/ethnicity and by disability type: disabled African American students have greater odds of being suspended than their white special education peers, and students with emotional and behavior disorders have the highest odds of suspension compared to all other disability categories (Krezmien, Leone, & Achilles, 2007; Zhang, Katsiyannis, & Herbst, 2004). Emotionally and behaviorally disabled students are approximately twice as likely to be excluded as their special needs peers (Zhang, Katsiyannis, & Herbst, 2004).

Exclusionary discipline for special education students is incredibly complicated, as special education students' right to a free and appropriate public education is protected in the

Individuals with Disabilities Education Act (IDEA) (see Skiba, 2002). The 1997 Amendments to IDEA try to ensure safe school environments while protecting students with special needs from being unfairly excluded from the classroom. If students are excluded for longer than ten days, either consecutively or summed over the course of the school year, special education students must be provided with services that allow them to achieve their Individualized Education Program (IEP) goals. If special education students are to be removed from the classroom for a longer period of time, the school must establish that the student either is dangerous to himself or to others in his current placement; parents are guaranteed the right to appeal all long-term changes of placement, including suspension and expulsion. During the appeal, the student must stay in his original educational placement.

Manifestation Determination, codified in Public Law 94-142 and upheld by the United States Supreme Court in *Honig v. Doe*, makes the process for excluding special education students even more complex. Manifestation Determination is a process by which the school-level IEP team assesses whether the student's disability is responsible for the infraction in question, as students cannot be excluded from school as a result of their documented disabilities. In conducting this review, the IEP team must first confirm that the student was receiving appropriate services and interventions. If so, the team must establish that the student's disability did not prevent the student from understanding the ramifications of and consequences for his behavior. Finally, the team must affirm the student's ability to control his behavior. All three criteria must be met if the special education student is to be exacted an exclusionary discipline sentence through standard procedures. Once excluded, however, the student must still be provided with services necessary to meet his IEP goals.

In some ways, Manifestation Determination makes findings about the disproportional representation of special education students in exclusionary discipline more surprising, as the students with special needs who are most often subjected to exclusionary discipline – those with emotional or behavioral disorders – are precisely the students whose misbehavior seems most likely to be related to his or her disability (Zhang, Katsyannis, & Herbst, 2004). These students are more likely to have underdeveloped social skills and judgment, and are less likely to be able to hide their behavior should they commit a transgression (Leone, et al., 2000). Furthermore, students with emotional and behavioral disorders are more likely to be in more restrictive environments than their peers in other disability categories, which suggests that these students should be receiving more extensive services to help them deal with their emotional and behavioral challenges (Zhang, Katsiyannis, & Herbst, 2004).

Additional Student Characteristics. In addition to race/ethnicity and disability status, numerous student characteristics are related to subjection to exclusionary discipline. Much research demonstrates that males are suspended more often than females, and numerous findings support a negative relationship between socio-economic status and suspension, with low socio-economic students suspended at higher rates than their more affluent peers (Christle, Nelson, & Jolivette, 2004; Rafaelle Mendez, 2003; Rafaelle Mendez, Knoff, & Ferron, 2002; Skiba, Peterson, & Williams, 1997; Theriot, Craun, & Dupper, 2010). Academic and behavioral history is also related to exclusionary discipline, with underperforming students and students with higher numbers of previous infractions suspended more often than their peers with stronger academic and behavioral track records (Arcia, 2006; Christle, Nelson, Jolivette, 2004; Morrison, et al., 2001; Theriot, Craun & Dupper, 2010). Suspended students also are more likely to have family or home life problems, have lower levels of self-reported personal optimism, and have weaker senses of social responsibility than students who have never been suspended (Morrison, et al., 2001).

School-Level Predictors of Suspension. Researchers generally agree that school-level characteristics that relate to suspension deserve attention. Research suggests that suspension is negatively related to school-level attendance, academic achievement, socioeconomic status, and percent white enrollment, and is positively related with school-level dropout rates and behavioral violations (Bruns, et al., 2005; Christle, Nelson, & Jolivette, 2004; Rafaelle Mendez, Knoff, & Ferron, 2002; Theriot, Craun, & Dupper, 2010). Research on the relationship between school size and suspension rates and staff experience and suspension rates is inconclusive, with some research suggesting positive relationships and other research failing to find statistically significant associations (Christle, Nelson, & Jolivette, 2004; Rafaelle Mendez, Knoff, & Ferron, 2002; Theriot, Craun & Dupper, 2010).

Qualitative research reveals that school leadership characteristics and school learning environment are also likely related to suspension rates. Schools with high suspension rates may be more likely to have inexperienced school leaders, less clear behavioral expectations, and less orderly conditions than schools with low suspension rates (Christle, Nelson, & Jolivette, 2004). Schools with low suspension rates may have higher academic and behavioral expectations for their students and employ teachers who utilize a wider variety of instructional procedures than do those at schools with high suspension rates (Christle, Nelson, & Jolivette, 2004). Furthermore, schools with low suspension rates may have greater parental involvement, both in the creation of discipline plans and discussion of academic outcomes, but also in recreational activities unrelated to student achievement (Rafaelle Mendez, Knoff, & Ferron, 2002). At schools with lower suspension rates, staff members may be more likely to use positive behavior reinforcements and place greater emphasis on the teacher's responsibility as a classroom manager (Rafaelle Mendez, Knoff, & Ferron, 2002).

Suspension and Student Outcomes

Although the intent of suspensions is to deter negative behavior and preserve safe and orderly school environments, little research establishes the effectiveness of suspension. Research has not shown that suspension improves students' behavior or increases school safety (for a review, see Skiba, 2000). From the student perspective, the degree to which suspension succeeds in its aversive role is uncertain. Students do not believe that suspension decreases problematic behaviors (Costenbader & Markson, 1998). In this study of 620 middle and high school students, a third of suspended students believed that suspension was not helpful and that they would likely be suspended again (Costenbader & Markson, 1998). Another third of students believed that they learned little from their suspension. Students' beliefs about the ineffectiveness of suspension in deterring problematic behavior are supported by research that suggests that previous suspensions are a strong predictor of future suspensions and antisocial behavior (Costenbader & Markson, 1994; Hemphill et al., 2006; Tobin et al., 1996).

In addition to predicting future infractions, suspension is associated with negative student academic outcomes. These negative associations make the systematic relationships between student and school characteristics and the likelihood of suspension even more troubling. More specifically, racial/ethnic disproportionality in suspension, when combined with negative relationships between suspension and academic success, raises questions about the role of suspensions in establishing and perpetuating racial/ethnic achievement gaps. Suspension is negatively related to student achievement (Arcia, 2006) and is positively associated with dropping out of school, both at the student level (Arcia, 2006; Bradley & Renzulli, 2011; Raffaele Mendez, 2003; Skiba & Peterson, 1999; Suh & Suh, 2007; Suh, Suh, & Houston, 2007;

Wehlage & Rutter, 1986) and at the school level (Christle, Jolivette, & Nelson, 2007; Lee, et al., 2011). The relationship between suspension and dropout status varies by gender, with males who have ever been suspended being more likely to dropout than females (Bradley & Renzulli, 2011; Suh, Suh, & Houston, 2007).

These negative relationships may be a result of many factors. First, suspension and dropout status may be related because school staff use suspension as a way to push troublesome students out of their schools (Fine, 1986; Skiba, 2000). When suspension excludes students from school, this lack of instruction may negatively impact student achievement (Bear, 2000; Townsend, 2000). Suspended students may also feel stigmatized or may rebel against a school environment that is perceived to be against them (Hemphill et al., 2006). Suspension also may weaken the bond between the student and the school, which is detrimental given that feeling socially connected to school decreases the likelihood of juvenile delinquency (Hawkins, Smith, & Catalano, 2004; Jenkins, 1997).

Past suspension may predict future suspension because suspension does not address the cause of students' misbehavior (Hyman, 1997; Martens & Meller, 1990). Once suspended, suspension may reinforce the problematic behavior that invoked the suspension, as suspended students may prefer to be excluded from the academic environment they perceive to be irrelevant and in which they feel unwelcome (Tobin, Sugai, & Colvin, 1996). Away from school, suspended students, who are more likely to be from neighborhoods with increased risk factors, may be surrounded by influences that increase their exposure to the negative behaviors that led to suspension and that are also associated with negative long-term academic outcomes (Walker et al., 1996). Indeed, research suggests that suspended students with more positive outlooks on their futures and healthy influences, such as peers who value education and plan to attend college and constructive relationships with teachers, are less likely to drop out than their suspended peers surrounded by negative influences (Suh, Suh, & Houston, 2007).

Methodological Concerns with Extant Research

Much debate surrounds the validity of the methods used to calculate disproportional representation in suspension. The most common means of establishing disproportionality are questioned due to their simplicity. The baseline racial/ethnic distribution method and the absolute proportion method are comparisons between the representation of the concerned group as measured by their representation in disciplinary sentences (e.g. percent of suspension earned by the ethnic group) or in the exclusionary discipline population (e.g. percent of suspended students who are in that ethnic group) compared to their representation in the school population (Fenning & Rose, 2007). If the proportion in the exclusionary discipline population is at least 10 percent greater than their proportion of the overall population, then this group would be considered to be overrepresented in exclusionary discipline (Reschly, 1997). One concern with these methods is that they fail to control for confounding factors, such as socio-economic status (Fenning & Rose, 2007). Furthermore, estimates of disproportional representation vary depending on the unit of analysis (Westat, 2004). Because student populations vary across schools, districts, regions and states, estimates of disproportionality also vary based on the chosen locality (Fenning & Rose, 2007).

This research on suspension rates also fails to address whether the racial/ethnic disproportionality is a result of bias, as doing so requires more advanced methodology and comprehensive information than is frequently employed in discipline research. For example, examining bias would necessitate information on students' behavior and office referral information. Studies that do attempt to address bias often produce inconclusive evidence. For example, McCarthy and Hoge (1987) found that students' race/ethnicity predicted the likelihood of suspension even after controlling for teachers' rating of students' behavior and students' self-reported behavior information. Yet their study failed to account for the nested nature of students within schools, thereby not eliminating the possibility that differential suspension rates by race/ethnicity could have been a result of school assignment.

Roque (2010) improved upon this research by using school fixed effects to explore racial/ethnic differences in office referrals after controlling for teachers' behavioral rating of students. Although racial/ethnic coefficients are smaller using this fixed effects approach, Roque still finds that African American students are more likely receive office referrals than are white students within the same school. Kinsler (2011), however, found little evidence for differential treatment of students according to race/ethnicity. Unlike Roque, Kinsler does not control for ratings of student behavior but also employs school fixed effects. He finds that once school fixed effects are entered in the model, racial/ethnic suspension gaps between black and white students disappear, suggesting that race/ethnicity is related to variation in suspension across schools, not within schools. To measure racial/ethnic bias, Kinsler includes interaction terms for student and teacher race/ethnicity and student and principal race/ethnicity. None of the teacher or administrator race/ethnicity effects are significant.

Literature on the effects of suspension is also open to critique, as researchers often fail to account for the nested nature of students within classrooms, within schools, and within district policy contexts. Ignoring the nested nature of discipline data not only means that researchers may make incorrect inferences based off of simplistic modeling, but also that researchers fail to explore the complex interactions between student and school level characteristics (Theriot, Craun & Dupper, 2010). Moreover, much literature fails to adequately control for the unmeasured differences between students who are and are not assigned suspensions. Thus, many associations uncovered in the literature may be due to unmeasured attributes of students, rather than of suspensions themselves.

Summary

Taken as a whole, the literature on exclusionary discipline policy suggests that these punishments are used unevenly within and across schools based on student characteristics, such as race, gender, and special education status. Reviewing the literature also highlights a need for additional research on the reasons for the unequal suspension assignment, as well as for the necessity of rigorous methods when examining the characteristics that predict suspension and the associations between suspension and student outcomes. Furthermore, the relative scarcity of work on the effects of suspension indicates a need for more research that investigates the links between exclusionary discipline and students' academic achievement. Therefore, I employ methodology that helps reduce differences between suspended and non-suspended students based on measured characteristics, accounts for the nested nature of students within schools, and examines a range of academic outcomes in an effort to help academics, policymakers, and educators better understand the efficacy of exclusionary discipline policies.

Chapter 3. Methodology

My study focuses on five questions central to suspension policy. These questions arise from my conceptual framework, but take a more focused view of the suspension process. First, I analyze which student social and academic characteristics are associated with suspension in New York City. I then mobilize my conceptual frame, which sees suspension as an act within a series of nested environments, by examining the school characteristics that are associated with suspensions while simultaneously accounting for student-level influences. Because the relationship between student characteristics and suspension may depend upon school characteristics, I also investigate whether these school characteristics mediate the relationship between student characteristics and suspension. For example, are traditionally disadvantaged students more likely to be suspended in particular school contexts? Responding to these first three questions will allow me to describe the characteristics that are associated with suspension in a sophisticated manner. Doing so allows me to establish that suspended students differ from their non-suspended peers in ways that likely make endogeneity problematic when trying to use basic quantitative techniques to estimate relationships between suspension and academic outcomes. Next, I move on to an assessment of the relationship between suspension and short-term outcomes, including attendance, and the likelihood of passing core classes. Finally, I estimate the relationship between suspension and long-term academic outcomes, including likelihood of graduation and Regents Exam performance.

To design this study, I have drawn upon the existing literature on the relationship between student and school characteristics and exclusionary discipline, upon literature on the discipline policy context in New York City, and upon a conceptual framework based upon Bronfenbrenner's Ecological Theory of Human Development, which acknowledges the nested nature of students within schools within districts. In this chapter, I briefly explain the policy context for suspensions in New York City, and then proceed to describing the data and methods I will employ to address my research questions.

Suspensions in New York City

New York City schools use the annual Citywide Standards of Discipline and Intervention Measures, known as the Discipline Code, as the foundation for their discipline system. This code articulates standards of behavior, types of infractions and corresponding consequences, and preventative measures that can be taken to help students meet the district's behavioral expectations. In doing so, the Discipline Code "ensures both consistency and equitable treatment for all students and enables principals and superintendents to exercise discretion and educational judgment" (NYCDOE, 2005) and instructs school personnel to be "sensitive to issues that may be impacting upon the behavior of students and respond in a manner that is most supportive to their needs" (NYCDOE, 2005).

In the Discipline Codes for years 2005-2008, behavioral infractions and consequences are listed separately for kindergarten-fifth grade students and sixth-twelfth grade students. For each set of grades, infractions are broken into five categories: Level 1 – Insubordinate Behaviors, Level 2 – Disorderly Disruptive Behaviors, Level 3 – Seriously Disruptive Behaviors, Level 4 – Dangerous or Violent Behavior, and Level 5 – Seriously Dangerous or Violent Behavior. For each level, the Discipline Code outlines a range of possible consequences that the administration may employ; for some infractions, the range of consequences is restricted to a subset of suggested punishments. In 2005-2006, there were 57 infractions listed in the Discipline Code, 45 of which could warrant a suspension and 18 of which were zero tolerance. In 2006-2007, there were 62 infractions listed in the Discipline Code, 50 of which could warrant a suspension and 27 of which were zero tolerance. In 2007-2008 and 2008-2009, the number of infractions increased to 63, and the number of zero tolerance infractions went up to 29 (NYCLU, 2011).

Beginning in 2006, the Discipline Code also includes a list of possible guidance interventions that can be used for each infraction level, and a note about supporting students transitioning back to the classroom following a suspension.[1]

There are two main categories of suspension in New York City: Principal's Suspension and Superintendent's Suspension. Principal's Suspensions last between one and five days and do not appear on students' permanent records. To assign these suspensions, principals must provide parents/guardians with written notice of the suspension, and meet with the parents/guardians within five days of the infraction. Ultimately, it is up to the principal's discretion to determine whether the suspension is warranted (The Legal Aid Society, 2013). Schools are responsible for providing alternative instruction to students during the length of their suspension.

Students who are suspended for more than five days are assigned Superintendent's Suspensions. These suspensions may last up to one year based on the severity of the infraction, and, if the student is over 17 years of age, may also involve expulsion. These suspensions also dictate whether the student can petition for early reinstatement to their school and whether students are assigned to a Second Opportunity school following their sentence. In 2008-2009, twenty-two percent of students earning a long-term suspension were suspended for more than ten days, a 66 percent increase since the 1999-2000 school year (NYCLU, 2011). On average, these long-term suspensions lasted twenty-five days, which is over two times the length of the maximum suspension period in most school districts (NYCLU, 2011).

After students are notified of a Superintendent's Suspension, students have a constitutional right to due process before being excluded from school. In New York City, this due process is

[1] This note reads, "Students returning from suspension should be provided with supportive services to maximize their ability to meet social and academic standards within the school community. Support services may include any of the range of guidance interventions or a combination of services as best meets the needs of the individual student."

outlined in the Chancellor's Regulations; however, students are often unaware of their legal rights largely due to the inaccessibility of due process documentation (NYCLU, 2011). Students who are suspended for longer than six days are granted a suspension hearing. Before the suspension hearing, the school must provide parents with written notification of the suspension, the date of the suspension hearing, and the student's alternate school assignment. At this time, parents may request a packet of all evidence that may be presented at the suspension hearing. Next, the parents and student attend a pre-hearing conference to decide whether to request an adjournment, perhaps to allow the student to acquire legal representation; to enter a "no content plea;" or to demand a full hearing. Based on these proceedings, the hearing officer determines the student's sentence (The Legal Aid Society, 2013).

Until the 2006-2007 school year, students earning Superintendent's Suspensions were removed from their traditional schools and classrooms and placed in Borough Suspension sites. In 2007-2008, Chancellor Klein reorganized these Suspension Sites into 38 Alternate Learning Centers (ALCs) assigned to District 88 in response to a class action lawsuit, *E.B. v. Board of Education.* This suit, filed by Advocates for Children, argued that students with special needs had been improperly excluded from the educational setting to which they were entitled and placed within settings that did not meet the requirements of their IEPs (Gootman, 2004). These ALCs are intended to be an improvement on Suspension Sites, which did not have any curriculum and had attendance rates of approximately 38 percent (Harball, 2012). The 2007 reforms attempted to improve the quality of education provided to students by requiring that all centers use the same curricular materials and are staffed by a site supervisor, four content area teachers, one special education instructor, one counselor, one paraprofessional, and one school aide (NYCDOE, 2014a). However, attendance at these centers is still almost half of the city average and hovers around 50 percent (NYCLU, 2013), and students attend shorter school days. Students under 16 are provided with five hours of instruction; those older are provided with two hours of instruction (NYCLU, 2013).

From 1999 to 2009, the New York City Department of Education's (NYCDOE) use of an exclusionary punishment system increased (NYCLU, 2011). In the 2008-2009 school year alone, 73,943 New York City students were suspended, compared to 42,937 students in 1999-2000. Many factors relate to the sharp rise in suspension rates. One contributing factor is NYCDOE's increasing reliance on zero tolerance discipline policies, which mandate that students be suspended upon their first infraction: in 2009-2010, there were two times as many zero tolerance discipline infractions listed as there were in 1999-2000 (NYCLU, 2011).

A second factor that relates to the rise in suspension rates over the past decade is NYCDOE's growing use of police in schools. Since 1998, when the Board of Education transferred responsibility over school safety to the New York Police Department (NYPD), the school safety division has increased by 64% (NYCLU, 2011). School safety officers are not trained in positive discipline techniques or in child development. Instead, their 14 week training relates only to penal law and to the Discipline Code (NYCLU, 2011). Whereas school safety officers may maintain a positive presence in schools with safe and welcoming school environments, they often are seen as aggressive and antagonistic in schools with higher rates of discipline incidents (ALCU & NYCLU, 2007). In these schools, such as the Impact Schools which are purposefully assigned higher numbers of school safety officers, police presence may mean that minor disciplinary infractions lead to arrest rather than to positive disciplinary intervention (ACLU & NYCLU, 2007; NYCLU, 2011).

Data and Methods

Data. My analyses employ longitudinal administrative data on public school students in

New York City, the nation's largest public school district, which serves approximately 1.1 million students in just under 1,700 schools. The city serves a diverse student clientele: 31% are black, 40% Hispanic, 14% white, and 15% other minorities, including Asian and multi-racial students (NYSED, 2011), and almost half of the student population (42%) comes from a non-English speaking household (NYC DOE, 2007). My research follows a single cohort of students from the Fall, 2005 semester—when they were first-time ninth graders—through Spring, 2009, when they should be preparing to graduate (*n*=70,130).

In essence, focusing on New York City data fixes the estimated effects for district policy, as all suspensions occurred under the same discipline code. These data include middle and high school socio-demographic, academic and behavioral measures, including eighth-grade New York State English Language Arts and Mathematics scores, middle and high school attendance and suspension records, and high school transcript information, Regents Exam scores, and four-, five-, and six-year graduation status. I supplement these data with New York City Department of Education (NYCDOE) School Survey data, and New York State accountability data.

Descriptive Sample. My descriptive sample includes all New York City public high school students who were identified as members of the 2005-2006 ninth-grade cohort by the NYCDOE, were enrolled in a high school for the first time in Fall 2005 and were enrolled for credits in at least one semester between the 2005 and 2008 school years (n=70,130). Notably, this sample includes students in both Districts 75 and 79.

Analytic Sample. When possible, I utilize my full descriptive sample for my analytic models. However, by necessity, at times this sample must be limited. For research questions 2 and 3, when I use the NYCDOE School Surveys, I exclude students in District 75 and 79 schools, as well as transfer schools. These schools are excluded both because there is not survey information for the majority of these institutions, and because these schools are, by definition, very different from those serving the majority of students in New York City. Since these models are focused on school characteristics related to suspension, I limit my sample to traditional schools to provide more generalizable results.

For research question 4 – which focuses on the relationship between suspension and short-term outcomes within students – the semesters for which students lack transcript and attendance information are excluded from the relevant analyses. I assume the information for these semesters is omitted due to lack of enrollment, not due to missing information that should have been included in the data file.

Suspension Indicators. To explore the relationships between student and school characteristics and suspension, and suspension and short- and long-term outcomes, I utilize dichotomous high school suspension measures indicating whether a student earned at least one Principal's Suspension that semester, whether a student earned at least one Superintendent's Suspension that semester, and whether a student earned either a Principal's or Superintendent's Suspension that semester.[2] Approximately 16 percent (n=11,163) of this cohort earned at least one suspension during their first four years of high school. My analyses also include an indicator of whether students were assigned a Principal's suspension (1=yes, 0=no), a Superintendent's suspension (1=yes, 0=no) or any type of suspension (either Principal's or Superintendent's) in seventh or eighth grade (1=yes, 0=no).

Student-Level Measures. To account for students' socio-demographic characteristics, I employ dummy-coded measures that identify whether the student is black, Asian, Hispanic, or white. In the multivariate models, white students serve as the comparison group. I also utilize

[2] Suspension does not, on its own, prevent students from taking Regents examinations, making it unlikely that spikes in the suspension rate occur near the administration of state exams.

an indicator of gender (1=female, 0=male), whether the student received free or reduced-price lunch (1=yes, 0=no), had an Individualized Educational Program ([IEP] 1=yes, 0=no), and was an English Language Learner (1=yes, 0=no). Finally, I employ a continuous measure of age at ninth grade entry, measured in years, as well as an overage indicator, with overage students being at least 15 months older than the youngest age for that cohort (1=yes, 0=no).

To control for students' academic histories prior to high school, I include four measures of middle school characteristics: students' eighth grade New York State English Language Arts and Mathematics scale scores, and eighth grade absences and latenesses. These exams include content that is aligned to the New York State learning standards. I use the scale scores, as test items are scaled to reflect their difficulty. I standardize these scores to facilitate the interpretation of results.

I also employ multiple measures related to students' high school academic careers. First, I include semester-level measures of days absent and late.[3] It is important to note that the school district does not count students as absent the days that they are suspended if they show up to their alternate learning site. I also utilize a continuous measure of the percent of all credits passed per semester, the percent of math credits passed per semester, and the percent of English credits passed per semester. Additionally, I employ measures of the total number of credits, math credits, and English credits earned per semester.

To investigate relationships between suspensions and Regents Exam scores, I include dichotomous measures indicating whether the student passed Regents Exams, including the English, at least one Math, the Global History and Geography, the United States History, and at least one Science exam. These exams include content aligned with the New York State learning

[3] Lateness is used in lieu of tardiness, as tardy often implies being late for class, whereas late implies showing up late for school.

standards. On these exams, passing is considered earning a score of at least 55, as students of the 2005 cohort were able to earn Local diplomas if they had grades of at least 55 on 3 Regents and 65 on the others. I also utilize students' scores on these exams, which range from 0-100. The exams are offered in January, June, and August, and students are able to take these exams until they pass. Finally, I use dichotomous measures of whether the student graduated in four, five, or six years.

School-Level Measures. I utilize two data sources to account for school-level factors that may influence the likelihood of suspension. The first includes aggregate measures of the student characteristics described above. Because I have access to the population of first-time ninth graders within the 2005-2006 cohort and because aggregate statistics are not publicly available through the NYCDOE, the New York State Education Department, or the Common Core of Data managed by the National Center for Education Statistics for all schools within my study,[4] using these aggregate socio-demographic and academic measures allows me to employ information on all schools within my sample. These aggregate statistics are used in the fifth chapter, which answers my fourth question, as school characteristics are not the focus of these models and are just used for additional controls.

From the New York State Report Card Data, I have access to the percent of the student body that is female; is an English Language Learner; has an Individualized Education Program; qualified for free- or reduced-price lunch; is white, black, Hispanic, Asian or Other. When adjusting for these characteristics with aggregate statistics, I include the percent of ninth graders are female, are English Language Learners, have IEPs, qualified for free- or reduced-price lunch, or are black/Hispanic.

[4] These school-level statistics are not available for all District 75 and 79 schools in the New York State Report Card database or in the National Center for Education Statistics Common Core of Data. Eliminating these schools would change the population about which I am able to make inferences.

From the state, I include a three-level enrollment measure. Schools are classified as small if they have fewer than 550 students in ninth through twelfth grade, midsize if they have 551-1,400 students, and large if they have greater than 1,400 students. I use midsize as my comparison group. I use a measure of school selectivity based on the New York City High School Directory. Schools are classified as selective if over half of first-time ninth graders participate in an academically selective program at that school.

I also use a measure of the average age of ninth graders. To account for the entering characteristics of students, I employ average measures of eighth grade English Language Arts and Mathematics scores and of the percent of students who earned a Principal's or a Superintendent's Suspension during seventh or eighth grade. In addition, I include eighth grade attendance data, include number of absences and latenesses over the course of the year.

In order to investigate the relationship between school characteristics and the likelihood of suspension, I employ the NYCDOE School Survey data and New York State Report Card data. Beginning in the 2006-2007 school year, high school students, parents, and teachers were asked to fill out a school learning environment survey to assess the degree to which their schools maintained an engaging environment of safety and respect, upheld high academic expectations, and facilitated useful communication among school community members. From these surveys, I employ four factors created by the NYCDOE: the Safety and Respect score, the Engagement score, the Communications score and the Academic Expectations score, which are calculated from responses on four-point likert-scaled items. These scores are available as school composites, as well as factors based separately off of student, parent, and teacher responses. These factors are highly reliable, but are also highly correlated with each other, leading to some question as to whether they are best combined into one learning environment metric (Nathanson, McCormick, & Kemple, 2013).

Missing Data. Table 1 presents missing data rates for key measures included in my analyses. I am fortunate to have complete data on all background characteristics except age at ninth grade entry, for which I am missing data for approximately 20 percent of cases. I am also missing data on 15-20 percent of cases for middle school academic characteristics. For roughly 15 percent of cases, this middle school information is missing because students transferred into the NYCDOE for their high school years. With regard to high school academic data, I am not missing any information on Regents pass rates or on whether the student graduated. I have the complete transcript files from students' first four years in high school. I make the assumption that these files are complete. Therefore, if a student is missing transcript information for a particular semester, it is assumed that this missing information is due to non-enrollment. Furthermore, even if these measures were not complete, imputing the missing data would not be appropriate given that they are used as outcomes in my analytic models (Von Hippel, 2007).

Table 1 also presents the missing data rates for the school-level information used to answer my second and third research questions in Chapter 4. This sample includes information on 322 schools. Five schools were missing 2006-2007 survey information; for these schools, I used the 2007-2008 survey measures rather than multiply imputing the data. There is no missing information on school-level socio-demographic measures; however, there I do not have complete information about school staffing.

It is likely that the missing data mechanism for these data is not Missing Completely at Random (MCAR). If it were, using listwise deletion in my research would not pose a threat to the validity of my conclusions. Unfortunately, it is likely that missingness arises due to student and school characteristics. For example, it is likely that students with high rates of missing data are more mobile and less advantaged than are students with lower rates of missing data. Students with higher absence rates may be less likely to take their eighth grade state exams, and students with more disadvantaged backgrounds are likely to have higher absence rates than their more advantaged peers. School characteristics also could give rise to missing data, as higher functioning schools or schools with more competent administrative personnel may be better at meticulously maintaining student records and submitting reports. Therefore, using listwise deletion would not only change the population about which I am able to make inferences, but it would also force me to lose valuable information about the cases for which I have partial data.

As the preferred alternative, I use multiple imputation to address my missing data through the MI package in Stata. More specifically, I employ an Imputation by Chained Equations (ICE) approach, which allows me to make fewer assumptions about the distribution of my measures, as the imputations are calculated based on a series of univariate models. This approach allows me to maintain my entire dataset, use all available information to generate replacement values, reflect both sampling and model uncertainty, and preserve relationships among all variables. Critically, this approach affords me the ability to make less biased inferences about my dataset than had I employed more simplistic methods of handling missingness (Allison, 2002; Rubin, 1996).[5]

[5] As previously mentioned, I only utilize ICE for variables where missingness is indicative of a lack of information. Where missingness is purposeful (e.g. missing on percent of credits passed because the student was not enrolled in any credits), ICE is not used.

Table 1. Missing Data Rates for Key Variables

	Number Missing	Percent Missing
Background Characteristics (n=70,130)		
Special Education	0	0
Free/Reduced-Price Lunch	0	0
Poverty Status		
Female	0	0
Race/Ethnicity	0	0
Limited English Proficiency	0	0
Age at 9th Grade Entry	14,440	20.6
Middle School Academic Characteristics (n=70,130)		
Eighth Grade ELA Scale Score	16,025	22.8
Eighth Grade Math Scale Score	12,124	17.3
Eighth Grade Absences	10,926	15.6
Eighth Grade Lateness	10,926	15.6
Received Suspension in Middle School	11,368	16.2
High School Academic Characteristics (n=70,130)		
English Regents Scores	0	0
Math Regents Scores	0	0
Science Regents Scores	0	0
Global History Regents Scores	0	0
United States History Regents Scores	0	0
Graduated in 4 Years	0	0
Graduated in 5 Years	0	0
Graduated in 6 Years	0	0
School Characteristics (n=322)		
School Survey Factors	0	0
Enrollment	0	0
Percent Free/Reduced-Price Lunch	0	0
Percent Female	0	0
Percent White/Black/Hispanic/Asian/Other	0	0
Percent Overage	0	0
Percent English Language Learner	0	0
Percent Special Education	0	0
Selective High School	0	0
Percent Teaching out of Certification Area	57	17.7
Percent Teaching fewer than 3 Years	57	17.7
Percent without Appropriate Certification	57	17.7
Percent Turnover in the last 5 Years	59	18.3
Student-to-Teacher Ratio	1	0.31

Descriptive Approach. My first set of analyses details the suspension rates for the 2005-2006 cohort and describes the social and academic differences that distinguish students who were and were not suspended using t-tests for relationships between dichotomous and continuous student characteristics, chi-square analyses for categorical characteristics, and one-way Analysis of Variances (ANOVAs) for relationships between categorical and continuous characteristics.

Analytic Approach: Suspension and Student and School Characteristics. When estimating the relationships between suspension and student and school characteristics, I build models that take into account the nested nature of students within schools by using multi-level modeling techniques. Doing so allows me to connect to the nesting central to my conceptual framework. This way, I can examine both the student and school characteristics that may contribute to the likelihood that a student is assigned a suspension. Within these models, I am able to use cross-level interactions – a slopes-as-outcomes approach – to determine whether school factors help explain the relationship between student characteristics and the likelihood of suspension. Furthermore, I address the fact that students are mobile – they may attend a number of schools during their high school career – by modeling the relationship between student and school characteristics and suspension separately by 9th and 10th grade schools.

In the interest of model parsimony, I eliminate many non-significant terms from the models presented here. These include school-level measures of teacher turnover; the percent of teachers teaching fewer than three years; the percent of teachers teaching without appropriate certification; the pupil-to-teacher ratio; the standard deviation of the students' eighth grade exam scores; the school safety factors separated by parent, student, and teacher. Furthermore, the final results when using the school-level racial/ethnic composition as percent black, Hispanic, Asian, or Other were analogous to those in models using percent black/Hispanic, the latter was included. I also excluded interaction terms between students' race and school-level entering math abilities, school survey factors, and school-level racial/ethnic composition.

I also examined whether the relationship between middle school suspension and high school suspension varies across schools by utilizing random slopes within a linear probability model framework. I found that this relationship does vary and that including additional student and school characteristics does not explain this variance. This relationship warrants further investigation.

Analytic Approach: Suspension and Short-Term Outcomes. To examine the extent to which short-term student outcomes are a function of school suspensions, I again take into account the nested nature of the disciplinary process that is at the heart of my conceptual framework. I employ multivariate student fixed-effects models within a multilevel analytic framework. This approach reduces bias in the estimates by eliminating unmeasured differences across students in terms of characteristics that are likely associated with both the likelihood of suspension and my outcomes of interest. More specifically, I use multilevel modeling with adaptive centering and random effects, which holds several advantages over traditional fixed-effects approaches (see Raudenbush, 2009). In particular, this technique allows me to estimate both the measurement level characteristics associated with student outcomes (such as attendance), as well as the time-invariant student characteristics that may be related to short-term outcomes (such as race/ethnicity). Moreover, this approach results in the measurement level characteristics being orthogonal with any time-invariant student characteristic and time-invariant student characteristics being orthogonal with school-level characteristics, which satisfies the random effects assumption of independence.

Finally, these techniques provide better estimates than standard Ordinary Least Squares (OLS) regression, as they allow me to estimate within-student effects. If one were to interpret these results in a causal framework, the counterfactual for the semester in which a student was suspended would be the other semesters in which the student was not suspended. Thus, these models remove unmeasured differences between students when estimating the effects of suspension. In other words, students serve as their own control or counter-factual. However, causal interpretation should still be used with caution, as there are other unmeasured events within a semester that a student earned suspension – which may or may not be related to the suspension – that could affect the short-term outcomes considered in the model. For example, students may have increased chaos in their home environments during the semester of suspension, thereby being related both to the suspension assignment and the end-of-semester outcomes. Thus, without accounting for all time-variant characteristics that may relate both to suspension and the student outcomes employed, I am unable to confidently make causal statements about these results.

As with the models presented in Chapter 4, the models presented in Chapter 5 exclude interactions that were removed in the interest of model parsimony. Since suspension is our variable of interest, I tested and removed interactions between the suspension indicators and all student-level measures. I did the same with school-level aggregate measures of age, gender, race, eighth grade test scores, and special education status.

Suspension and Long-term Outcomes. To estimate the relationship between suspension and graduation and Regents performance, I will employ school-fixed effects and multi-level propensity score matching. With these techniques, I try to create a more plausible counterfactual for suspended students than all students who were never suspended. With the school-fixed effects, I estimate the average treatment effect for students within schools, thereby eliminating any school-level differences that may be related to students' end-of-high-school outcomes. With multi-level propensity score matching, I use pre-suspension student and school characteristics to calculate the predicted probability of being assigned a suspension, and then match non-suspended students with suspended students based on similar probabilities (Gelman & Hill, 2007; Rubin, 1997; Rubin & Thomas, 1996). In this technique, the predicted probabilities for suspension assignment are calculated using probit models, with the predictors being the confounding covariates that are pre-treatment characteristics related both to the likelihood of suspension and to the long-term outcome in question. For both sets of analyses, I consider the treatment to be any suspension within the first three terms of high school. I also use seven outcomes: the likelihood of graduation, the type of diploma received, and Regents Exam performance on five exams (English, Global History and Geography, United States History, a Mathematics exam, and a Science exam). I conduct a separate set of analyses regarding the likelihood of graduation with suspension in 9th grade as the treatment.

I use propensity score matching with regression adjustment to estimate the average effect of the treatment on the treated. In other words, I will estimate how suspended students would have performed, on average, had they not been suspended. This technique allows me to provide more valid estimates over traditional regression techniques, as it allows me to be sure that my treatment and control group are balanced and have substantial overlap (Gellman & Hill, 2007; Rosenbaum & Rubin, 1983). Said a different way, I am only comparing students who were assigned suspensions to non-suspended students who are similar on all measured characteristics. While doing so, I ensure that I have a sufficient number of comparison cases in the analysis. Causal interpretation of these results should still be used with hesitation, however, as I am only able to account for observed confounding covariates when calculating the propensity of being assigned a suspension (Shadish, Cook, & Campbell, 2002). There are likely other student characteristics that are not measured and are related to suspension status and the included outcomes, such as motivation and engagement, that are not fully captured in my models.

Descriptive Results

Tables 2 and 3 provide frequencies of the number of suspensions by year and term and the number of suspended students by year and term. These frequencies differ, as some students are suspended multiple times within the same semester. Table 2 shows that each semester, the number of Superintendent's suspensions is much lower than the number of Principal's suspensions. Interestingly, I also note a nonlinear trend in the suspension rate: the number of suspensions increases after the first semester of high school, peaks during the second semester of sophomore year, and declines as students progress toward the year of graduation. This pattern may be due to changes in behavior and/or enrollment, as more troubled students may begin to drop out or enroll in alternative settings.

Table 2. Number of Suspensions by Year and Term

	Principal's Suspension	Superintendent's Suspension	Total Suspensions
2005, Semester 1	1505	384	1889
2005, Semester 2	2418	754	3172
2006, Semester 1	2501	645	3146
2006, Semester 2	2846	872	3718
2007, Semester 1	1761	643	2404
2007, Semester 2	2076	546	2622
2008, Semester 1	1780	474	2254
2008, Semester 2	1419	377	1796

The patterns evident in Table 2 are reflected in Table 3 as well. The number of students assigned a suspension also peaks in the second semester of sophomore year, which indicates that the peak in suspension is not due to the same students being suspended more frequently during this semester. As with the number of suspensions, the smallest numbers of students are suspended during their first semester of high school and the last semester of their fourth year: these frequencies are almost half those of the second semester of sophomore year. These suspensions are not distributed evenly across schools. For example, in the first semester of ninth grade, 45 percent of schools had no suspensions. In the second semester of tenth grade, when suspensions are at their highest, 32 percent of schools had no suspensions. In fact, in that term roughly 30 percent of schools account for approximately 80 percent of the suspensions.

Table 3. Number of Suspended Students by Year and Term

	Students Assigned Principal's Suspension	Students Assigned Superintendent's Suspension	Students Assigned any Suspension
2005, Semester 1	1321	376	1636
2005, Semester 2	2103	647	2626
2006, Semester 1	2113	590	2585
2006, Semester 2	2331	669	2873
2007, Semester 1	1520	604	2041
2007, Semester 2	1733	526	2177
2008, Semester 1	1495	459	1883
2008, Semester 2	1207	366	1507

Table 4 displays academic and socio-demographic differences between students who were and were not suspended at least once during their high school careers. Roughly 15% of the sample—over 11,000 students—had been suspended. These descriptive results generally confirm prior research regarding the types of high school students more likely to experience exclusionary disciplinary practices. Students who were suspended were almost twice as likely to have IEPs, and male suspended students outnumbered females almost two to one. In terms of racial/ethnic background, the clear finding is the overrepresentation of black students among those suspended. Approximately half of all suspended students were black, while they constitute less than one-third of the sample. Conversely, white and Asian students were far less likely to experience a suspension. I also find that those suspended were less likely to be Limited English Proficiency students, but were typically slightly older than students who were not suspended. School suspensions are also associated with much weaker academic backgrounds, with suspended students typically entering high schools with fewer academic skills. On average, suspended students in this cohort began high school with lower eighth-grade standardized test scores in both ELA (-0.406 *SD*s) and mathematics (-0.469 *SD*s).

During high school, suspended students continued to display fewer positive academic behaviors. On average, they missed over seven more days of school per year, and were late over four more days. Suspended students also typically passed a much smaller proportion of their classes. Whereas students who were never suspended passed over 80 percent of their classes on average per semester, suspended students passed fewer than 60 percent. As a result, they earned roughly 1.6 fewer credits per semester. Broken down by subjects, suspended students gained approximately 30 percent fewer credits in math and 25 percent fewer credits in English per semester compared to their peers who were never suspended. The graduation discrepancies are also grim: only 40 percent of students who received a suspension graduated within four years, compared to 70 percent of their non-suspended peers.

Table 4. Associations Between School Suspension and Students' Socio-Demographic and Academic Characteristics (n=70,130)

	No Suspensions (n=58,967)	At Least One Suspension (n=11,163)
Background Characteristics		
% Special Education	9.58***	19.1
% Free/Reduced-Price Lunch	45.8	44.8
% Female	52.9***	37.8
Race/Ethnicity***		
% White, non-Hispanic	14.2	7.3
% Black, non-Hispanic	31.1	50.8
% Hispanic	37.3	36.6
% Asian	16.7	4.7
% Other	.710	.590
% Limited English Proficiency	16.1***	9.59
Age at 9th Grade Entry	14.4***	14.5
	(0.660)	(0.748)
Middle School Academic Characteristics		
Eighth Grade ELA Scale Score	0.023***	-0.383***
	(1.13)	(0.940)
Eighth Grade Math Scale Score	0.064***	-.405
	(1.01)	(1.00)
Eighth Grade Absences	14.1***	19.4
	(15.1)	(16.3)
Eighth Grade Lateness	13.9***	24.6
	(21.9)	(24.8)
% Received Suspension in Middle School	4.03***	15.2
High School Academic Characteristics		
Average Days Absent Per Semester	10.1***	17.5
	(11.9)	(13.7)
Average Days Late Per Semester	5.90***	10.8
	(8.12)	(10.1)
Average Percent of Credits Passed Per	81.4***	59.7
Semester	(24.3)	(28.4)
Average Number of Credits Earned Per	5.96***	4.36
Semester	(2.24)	(2.34)
Average Percent of Math Credits Passed	77.2***	54.8
Per Semester	(28.6)	(32.3)
Average Number of Math Credits Earned	.929***	.666
Per Semester	(.423)	(.422)
Average Percent of ELA Credits Passed	82.6***	60.7
Per Semester	(26.3)	(32.0)
Average Number of ELA Credits Passed	1.17***	.885
Per Semester	(.514)	(.514)
% Graduated in 4 Years	69.4***	37.5
% Graduated in 5 Years	79.5***	51.8
% Graduated in 6 Years	79.8***	52.2
% Pass ELA Regents	78.3***	55.3
% Pass Math Regents	78.5***	54.1
% Pass Global History Regents	68.5***	41.1
% Pass US History Regents	76.9***	53.6
% Pass Science Regents	71.9***	46.7

Standard Deviations in Parentheses

*p<.05; **p<.01; ***p<.001

Significance for differences in MI data determined using F-Tests

This dichotomous grouping may mask differences between suspended and non-suspended that are even starker if I further disaggregate students by the type of suspension. Table 5 presents descriptive relationships between suspension and student characteristics using a four-level measure of suspension status: students who never received a suspension in high school, students who received at least one Principal's suspension but never received a Superintendent's suspension, students who received at least one Superintendent's suspension but never received a Principal's suspension, and students who received at least one Principal's and one Superintendent's suspension. On average, these groups are probably ordered from the least behavioral infractions to the greatest, with students who have received both kinds of suspension likely having the highest recidivism rate.

Broken down in this manner, my results suggest that the relationships established in Table 4 hold. As is expected given the literature on exclusionary discipline, as I move from students with no suspension to students with both levels of suspension, the percent of males, special education, and black students increases. In contrast, I note decreasing proportions of Limited English Proficient and white and Asian students. Similar trends also exist for academic characteristics in high school: students with at least one of both types of suspension have the worst records, with far lower mean eighth-grade English (-.558 *SD*) and Math (-.648 *SD*) scores, twice as many latenesses, and approximately two weeks of additional absences, on average, than their non-suspended peers during their eighth grade year. One in four of these suspended students was assigned at least one suspension in middle school, compared to one in 25 of the students who never received a suspension in high school.

In high school, these patterns persist. On average, students in the most suspended group were, on average, absent for two weeks more, were late twice as often, and earned roughly half as many credits, both overall and in math, per semester than their non-suspended peers. The graduation statistics also reveal considerable differences: whereas 70 percent of non-suspended students graduated in four years, only 41 percent of students who received at least one Principal's suspension, 39 percent of students who received at least on Superintendent's suspension, and 21 percent of students who received both levels of suspension do so.

Table 5. Associations Between School Suspension and Students' Socio-Demographic and Academic Characteristics (n=70,130)

	No Suspensions[1] (n=58,967)	At Least One Principal's Suspension (n=7,665)	At Least One Superintendent's Suspension (n=1,734)	At Least One Principal's and One Superintendent's Suspension (n=1,764)
Background Characteristics				
% Special Education***	9.58	17.52	18.6	26.7
% Free/Reduced-Price Lunch	45.8	45.2	45.4	42.7
% Female***	52.9	42.1	31.4	25.5
Race/Ethnicity***				
% White, non-Hispanic	14.2	8.47	5.77	3.91
% Black, non-Hispanic	31.1	48.0	52.0	61.9
% Hispanic	37.3	37.7	37.0	31.2
% Asian	16.7	5.23	4.61	2.27
% Other	.710	.570	.580	.680
% Limited English Proficiency***	16.1	10.2	9.63	7.03
Age at 9th Grade Entry	14.4	14.5***	14.6***	14.5***
	(0.660)	(0.753)	(0.919)	(0.822)
Middle School Academic Characteristics				
Eighth Grade ELA Scale Score	0.023	-0.347***	-0.390***	-0.535***
	(1.13)	(0.907)	(0.961)	(0.948)
Eighth Grade Math Scale Score	0.064	-0.372***	-0.373***	-0.584***
	(1.01)	(0.967)	(1.10)	(0.960)
Eighth Grade Absences	14.1	18.4***	19.6***	23.4***
	(15.1)	(15.3)	(17.3)	(18.3)
Eighth Grade Lateness	13.9	23.7***	24.8***	28.6***
	(21.9)	(24.4)	(26.6)	(26.0)
% Received Suspension in Middle School	4.06	13.0***	14.5***	25.9***
High School Academic Characteristics				
Average Days Absent Per Semester	10.1	16.0***	18.8***	22.6***
	(11.9)	(13.3)	(14.2)	(13.6)
Average Days Late Per Semester	5.90	10.6***	10.5***	12.3***
	(8.12)	(10.2)	(9.85)	(10.2)
Average Percent of Credits Passed Per	81.4	62.5***	60.7***	46.8***
Semester	(24.3)	(28.1)	(28.5)	(26.7)
Average Number of Credits Earned Per	5.96	4.58***	4.36***	3.40***
Semester	(2.24)	(2.33)	(2.29)	(2.14)
Average Percent of Math Credits Passed Per	77.2	57.2***	56.4***	42.4***
Semester	(28.6)	(31.9)	(32.8)	(30.8)
Average Number of Math Credits Earned Per	.929	.696***	0.679***	0.525***
Semester	(.423)	(.420)	(.424)	(.399)
Average Percent of ELA Credits Passed Per	82.6	63.8***	61.1***	47.0***
Semester	(26.3)	(31.5)	(32.1)	(30.6)
Average Number of ELA Credits Passed Per	1.17	.928***	0.883***	0.705***
Semester	(.514)	(.513)	(.507)	(.481)
% Graduated in 4 Years***	69.4	41.1	38.8	20.5
% Graduated in 5 Years***	79.5	56.0	51.2	34.6
% Graduated in 6 Years***	79.8	56.3	51.6	35.1

Standard Deviations in Parentheses

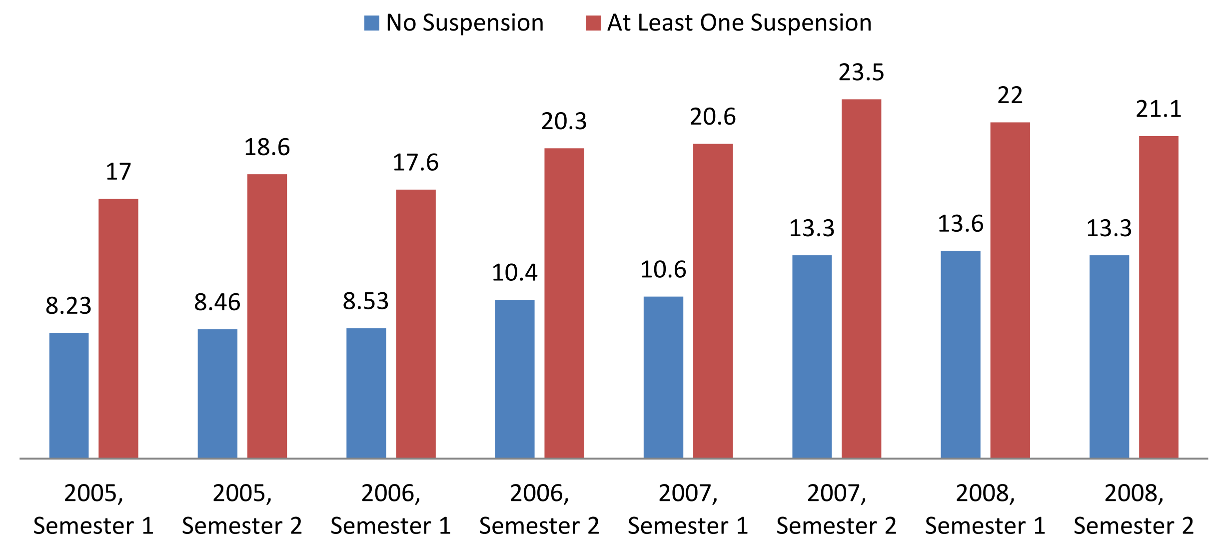
[1]For One-Way ANOVAs, No Suspensions serve as the comparison group.

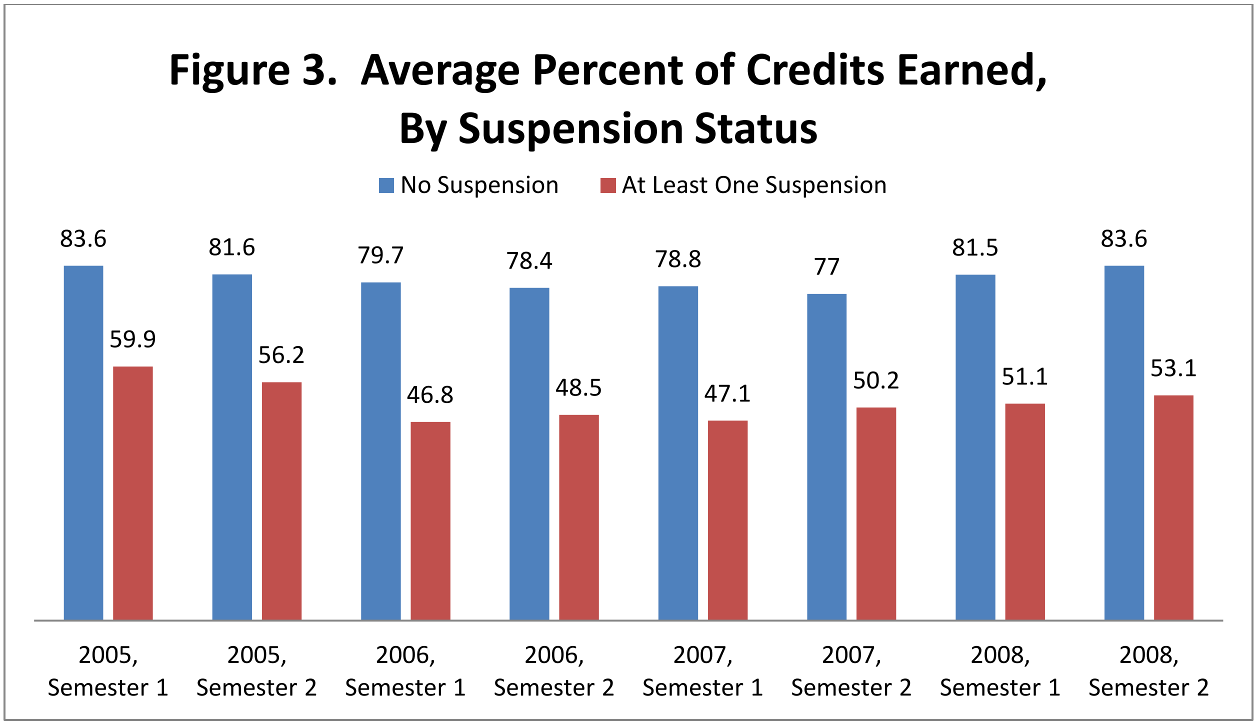
***p<.001

Significance for differences in MI data determined using F-Tests

Finally, I examined whether these trends differ across semesters. Since the suspension rate follows a nonlinear trend, looking at by-semester differences between suspended and non-suspended students may reveal nonlinear trends in differences. These semester-level means are presented in Table 6, and two semester-level relationships are depicted in Figures 1 and 2. In general, it appears possible that there may be nonlinear trends in these attendance and academic means; however, it is impossible to determine whether these differences would be statistically significant. Furthermore, it appears that both groups follow similar trends, which suggests that gaps between groups at one time period may not be significantly larger than gaps between groups at other time periods. For example, it appears that students have the fewest absences at the beginning of their high school careers, and that these absence rates gradually increase until they peak during their junior year and stabilize or slightly decrease their senior year. This trend basically holds for both groups. The same is true with the percent of credits earned. All students have the highest percentages of credits earned during their first and fourth years of high school, and it is not clear whether the differences between groups differ across years.

Figure 2. Average Days Absent Per Semester, By Suspension Status





Indeed, although illuminating, these simple descriptive results cannot disentangle the complicated associations among these overlapping student characteristics. The multivariate and quasi-experimental analyses, which will represent the analytic core of my dissertation, will build upon these preliminary findings by identifying the unique relationships between student and school characteristics and suspension and between suspension and short- and long-term outcomes.

Table 6. Associations Between Suspensions and High School Academic Characteristics

	2005, Semester 1		2005, Semester 2		2006, Semester 1		2006, Semester 2		2007, Semester 1		2007, Semester 2		2008, Semester 1		2008, Semester 2	
	No Suspension (n=67,979)	At Least One Suspension (n=1,636)	No Suspension (n=67,088)	At Least One Suspension (n=2,626)	No Suspension (n=67,149)	At Least One Suspension (n=2,585)	No Suspension (n= 66,693)	At Least One Suspension (n=2,873)	No Suspension (n=67,738)	At Least One Suspension (n=2,041)	No Suspension (n=67,367)	At Least One Suspension (n=2,177)	No Suspension (n= 67,902)	At Least One Suspension (n=1,883)	No Suspension (n=67,719)	At Least One Suspension (n=1,507)
Average Days Absent Per Semester	8.23***	17.0	8.46***	18.6	8.53***	17.6	10.4***	20.3	10.6***	20.6	13.3***	23.5	13.6***	22.0	13.3***	21.1
	(10.3)	(13.8)	(12.3)	(16.2)	(13.3)	(16.0)	(14.9)	(17.0)	(16.0)	(17.4)	(18.9)	(19.3)	(18.5)	(18.0)	(17.2)	(17.5)
Average Days Late Per Semester[1]	4.29***	7.43	5.78***	10.6	6.47***	13.1	6.82***	12.4	7.14***	14.2	8.02***	13.2	7.19***	13.6	--	--
	(8.11)	(9.80)	(10.5)	(13.1)	(11.7)	(15.5)	(11.7)	(14.8)	(12.3)	(15.9)	(13.3)	(15.6)	(11.9)	(15.5)	--	--
Average Percent of Credits Passed Per	83.6***	53.1	81.6***	51.1	79.7***	50.2	78.4***	47.1	78.8***	48.5	77.0***	46.8	81.5***	56.2	83.6***	59.9
Semester	(26.7)	(35.6)	(28.7)	(36.8)	(29.4)	(36.2)	(30.5)	(36.0)	(30.3)	(35.8)	(31.9)	(36.6)	(29.2)	(36.6)	(28.7)	(37.8)
Average Number of Credits Earned Per	5.88***	3.53	5.58***	3.42	5.93***	3.72	5.66***	3.40	6.12***	3.69	5.73***	3.50	5.84***	4.27	5.20***	4.03
Semester	(2.67)	(2.72)	(2.34)	(2.68)	(2.92)	(3.06)	(2.51)	(2.77)	(3.11)	(3.04)	(2.74)	(3.37)	(3.08)	(3.24)	(2.49)	(2.84)
Average Percent of Math Credits Passed Per	79.2***	47.4	76.6***	43.8	76.6***	47.9	75.7***	44.2	74.0***	44.5	72.9***	42.9	75.4***	49.2	77.4***	48.9
Semester	(40.1)	(49.4)	(42.1)	(49.3)	(41.7)	(49.2)	(42.5)	(49.0)	(43.1)	(48.5)	(43.9)	(48.7)	(42.1)	(48.8)	(41.0)	(48.5)
Average Number of Math Credits Earned Per	0.941***	0.555	0.863***	0.513	0.906***	0.569	0.846***	0.510	0.896***	0.545	0.823***	0.487	0.946***	0.603	0.909***	0.604
Semester	(.614)	(.628)	(.571)	(.632)	(.602)	(.626)	(.559)	(.610)	(.649)	(.643)	(.590)	(.586)	(.692)	(.662)	(.606)	(.667)
Average Percent of ELA Credits Passed Per	86.0***	57.7	83.3***	53.1	80.8***	50.5	79.5***	47.5	80.2***	48.7	78.6***	47.5	82.6***	56.6	85.4***	60.4
Semester	(33.5)	(48.0)	(36.4)	(48.7)	(38.1)	(48.3)	(39.0)	(48.1)	(38.0)	(47.6)	(39.2)	(47.5)	(35.7)	(46.5)	(33.3)	(46.2)
Average Number of ELA Credits Passed Per	1.20***	0.826	1.24***	0.756	1.10***	0.709	1.02***	0.650	1.15***	0.704	1.06***	0.655	1.19***	0.835	1.15***	0.895
Semester	(.726)	(.816)	(.686)	(.799)	(.709)	(.782)	(.654)	(.737)	(.805)	(.778)	(.723)	(.721)	(.814)	(.824)	(.718)	(.850)

Standard Deviations in Parentheses

[1]Data not available for 2008, second semester

***p<.001

Chapter 4. Suspension and Student- and School-Level Characteristics.

The simple descriptive statistics highlighted in the previous chapter indicate stark differences between students who are and are not suspended. Because some of the characteristics that are related to suspension, such as race/ethnicity, academic achievement and gender, are likely to both be associated with each other and with suspension status, it is methodologically more appropriate to estimate the relationship between suspensions and socio-demographic and academic characteristics in a regression framework, thereby allowing me to account for this complicated web of associations. The analyses in this chapter allow me to answer my first three research questions in a more sophisticated manner, as I am not only able to explore the relationship between student attributes and suspension, but also between school characteristics and suspension status. Because I am interested in school-level characteristics, these analyses exclude students in District 75 and 79 schools, as those schools purposefully have different school characteristics than those that serve the vast majority of New York City students. Schools in both of these districts are designed to educate students with demanding special needs, or who are in need of an alternative setting due to age, health, or academic or behavioral needs.

Logistic Regression Results

Much extant research estimates the relationship between suspension and student and school characteristics using one-level logistic regression. To replicate the results from this common analytic approach, I first fit logistic regression models that quantify the association between the odds of ninth grade suspension and student and school characteristics in ninth grade, in tenth grade and in all of high school. I estimate the results separately for ninth and tenth grade since this is when the bulk of suspensions occur and students may transfer schools between years. By running a separate regression for the tenth grade year, I am able to accurately match students to school characteristics.

Ninth Grade Suspension. Table 7 displays the ninth grade results. In Model 1, I limit the included covariates to student socio-demographic characteristics and find relationships consistent with those presented in the body of work on exclusionary discipline. I estimate that the odds of suspension are almost twice as high for students with IEPs as they are for general education students, but they are approximately 40 percent lower for English Language Learners compared to their general education peers. I also find that Hispanic students have 40 percent greater odds of suspension, whereas Asian students have approximately ten percent lower odds of suspension than their white peers, after adjusting for select socio-demographic characteristics. Unlike with the other races, the relationship between identifying as African-American and suspension status depends upon gender. Although black students and boys are far more likely to be suspended, these relationships depend upon one another. The odds of suspension for black males is over twice as that of white males, and the odds of suspension for black females compared to white males is thirty percent higher than it is for white females compared to white males.

Some of these relationships are partially explained by students' entering high school academic characteristics. In Model 2, controlling for academic and discipline history decreases the gap in odds of suspension for special education versus general education students to approximately 30 percent, as well as the gap between black and white males to 75 percent. However, it widens the gap for black females versus white females: black females now have 80 percent higher odds of suspension.

This model also reveals a very strong relationship between middle school suspension and high school suspension, which is responsible for the changes in the estimates for the relationship among suspension status, race, and gender. For male students, much of what appeared to be a high school gap is explained by whether students were suspended in the past. In contrast, suspension history acts as a suppressor measure, as including it increases the black/white female discipline gap. This suggests the suspension disparity for black/white females was not as large as it was in middle school as it is in ninth grade.

Regardless, the middle school discipline gaps are large: students who are suspended in middle school, either for short- or long-term, have over two times the odds of receiving a high school suspension as do students who were never suspended in middle school. This finding provides evidence that serving a suspension does not effectively prevent students from being suspended again in their futures, whether it be because students' behavior patterns remain unchanged or because personnel view students as likely offenders and treat them accordingly. The relationship between eighth grade test scores is also as one may suspect, with students with higher test scores having somewhat lower odds of suspension. Interestingly, this relationship is nonlinear. The drop in the odds of suspension associated with a one standard deviation increase in test scores is even larger for students with lower test scores. This drop becomes smaller as students approach average test scores. The further above average scores the students have, the larger the decrease in the odds of suspension yet again.

In Models 3, 4, and 5, I include school-level characteristics. In all three of these models, we see that students in large, non-select high schools have lower odds of suspension than students in non-select medium-sized high schools; students in large, select high schools have greater odds of suspension than students at medium-sized, select high schools; and students in medium-sized select schools have lower odds of suspension than do those in medium-sized non-select schools. As we will see in the multilevel framework, these findings are only a result of the fact that students are not nested in schools in this simplistic model.

Even after controlling for individuals' scores, these findings also indicate that higher levels of aggregate eighth grade ELA test scores are related to a trivial reduction in the odds of suspension, and that a ten percentage point increase in the percentage of students who received a suspension in middle schools is associated with a 30 percent increase in the odds of high school suspension. The racial/ethnic makeup of the students also weakly relates to the odds of suspension, with a 10 percentage point increase in the percent of black/Hispanic students being associated with a 6 percent decrease in the odds of suspension. These results suggest at the importance of peer effects and of selection into high school. Students are less likely to be suspended when they are surrounded by peers who are also less likely to be suspended.

Finally, these results highlight important relationships between school characteristics measured by the New York City School Survey. Higher scores on both the Safety and the Engagement factors are associated with approximately 15 percent lower odds of suspension. This suggests that at schools that are more orderly, with clearer expectations for student behavior and higher levels of respect among community members, as well as schools where students feel they have stronger relationships with staff and students and more interesting classes, students are less likely to be suspended. Conversely, a one standard deviation increase in the Communication score is related to a 20 percent increase in the odds of suspension. As Model 7 shows, the relationship between the Safety score and suspension is dependent upon gender, with females having decreased odds of suspension associated with higher Safety scores as compared to males. This suggests that females are more heavily influenced by their school environments than are males.

Table 7. The Relationship between the Odds of 9th Grade Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.63***	0.46***	0.46***	0.46***	0.46***	0.45***
Special Education	1.83***	1.28***	1.28***	1.29***	1.29***	1.29***
English Language Learner	0.60***	0.60***	0.60***	0.58***	0.60***	0.60***
Free/Reduced Price Lunch	0.88**	0.98	0.91*	0.91*	0.93	0.93
Age in 9th Grade (in years)	1.09*	0.98	0.96	0.96	0.96	0.96
Race[1]						
Black	2.37***	1.74***	1.66***	1.67***	1.62***	1.70***
Hispanic	1.43***	1.49***	1.34***	1.36**	1.34**	1.38**
Asian	0.85*	0.84	0.83	0.84	0.80	0.81
Other	1.19	1.13	1.08	1.07	1.04	1.07
Black*Female	1.38*	1.79***	1.73***	1.73***	1.70**	1.51*
Hispanic*Female	1.11	1.36	1.31	1.31	1.30	1.20
Asian*Female	1.27	0.92	0.92	0.91	0.91	0.88
Other*Female	0.99	0.99	0.98	0.98	0.94	0.86
8th Grade Absences (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.50***	2.50***	2.45***	2.46***	2.47***
Middle School Principal's Suspension		2.28***	2.31***	2.25***	2.27***	2.27***
8th Grade ELA Test Score[2]		0.79***	0.80***	0.81***	0.82***	0.82***
8th Grade Math Test Score[2]		0.78***	0.80***	0.81***	0.81***	0.81***
8th Grade ELA Test Squared		0.95***	0.95***	0.96***	0.96***	0.96**
8th Grade Math Test Squared		0.96***	0.96***	0.96***	0.97***	0.97***
High School Characteristics						
Percent Black/Hispanic[3]			0.97**	0.96**	0.94***	0.94***
Percent Free/Reduced Price Lunch[3]			1.05***	1.05***	1.05***	1.05***
Percent Female[3]			1.08***	1.10***	1.11***	1.12***
Percent Overage[3]			1.14***	1.10**	1.06	1.06
Percent Special Education[3]			1.16***	1.06	1.03	1.03
School Size[4]						
Small School			0.86	0.79	0.84	0.84
Large School			0.76*	0.67**	0.61***	0.61***
Select School			0.83	0.79	0.75*	0.75*
Small*Select			1.05	1.13	1.17	1.17
Large*Select			1.42*	1.69**	1.69**	1.69**
Mean 8th Grade ELA Score				1.00*	0.99***	0.99**
Mean 8th Grade Math Score				1.00	1.01	1.01
Percent Middle School Suspensions[3]				1.34***	1.34***	1.33***
Safety Survey Score[2]					0.86***	0.92
Engagement Survey Score[2]					0.82**	0.82**
Academic Survey Score[2]					0.95	0.95
Communication Survey Score[2]					1.19**	1.20**
Female*Percent Black/Hispanic						1.01
Female*Safety Survey Score						0.83***

*p<.05; **p<.01; ***p<.001; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points
[4]Compared to medium

Tenth Grade Suspension. Since the majority of suspensions occur in 10th grade, I create a separate model predicting tenth grade suspensions. These results are displayed in Table 8. Note that the number of students in this model decreases slightly, as some students who were enrolled in ninth grade are not enrolled in tenth grade. I examine these relationships in the same manner as with ninth grade suspensions, including six models, each of which builds upon the previous model. The relationships between tenth grade suspension and student and school characteristics are largely the same as those in ninth grade, with a few exceptions. At the outset, we see even smaller odds of suspension for Asian students compared to their white peers. Unlike in the ninth grade year, this relationship persists across all six models, indicating that Asian students have about 30 percent lower odds of suspension than do white students. A second difference is that we only note a nonlinear negative relationship between eighth grade math tests and the odds of suspension; the relationship between the odds of suspension and eighth grade ELA scores appear to be negative and linear. This means that the reduction in the odds of suspension associated with higher ELA scores is consistent across all scores, whereas reductions in suspension associated with math scores are more extreme at scores further from the average.

We notice the same school size and selectivity differences, along with a gap between small, selective schools and medium selective schools, with small schools having 40 percent higher odds of suspension. The tenth grade estimates suggest a relationship between the percentage of students who are overage and suspension, with a 10 percentage point increase in overage students related to an 11 percent increase in the odds of suspension. Finally, we see some differences in the relationship between the school survey results and the odds of suspension. Here, the only relationship that exists is between the Engagement factor and the odds of suspension, with a one standard deviation increase engagement associated with a 22 percent decrease in the odds of suspension.

Table 8. The Relationship between the Odds of 10th Grade Suspension and Student and School Characteristics. (n=68,424)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.54***	0.51***	0.51***	0.51***	0.51***	0.55***
Special Education	1.69***	1.20***	1.20***	1.20***	1.20***	1.20***
English Language Learner	0.72***	0.64***	0.64***	0.64***	0.65***	0.65***
Free/Reduced Price Lunch	0.83***	0.93***	0.85***	0.85***	0.87***	0.87***
Age in 9th Grade (in years)	1.04	0.95	0.93	0.93*	0.93*	0.93*
Race[1]						
Black	2.82***	1.86***	1.78***	1.79***	1.75***	1.74***
Hispanic	1.54***	1.50***	1.39***	1.41***	1.38***	1.37***
Asian	0.60***	0.68**	0.70***	0.71***	0.67**	0.67**
Other	1.44	0.97***	0.96	0.97	0.93	0.93
Black*Female	1.37*	1.46**	1.42**	1.42**	1.40*	1.44*
Hispanic*Female	1.10	1.00	0.97	0.97	0.96	0.99
Asian*Female	1.13	0.88	0.87	0.87	0.87	0.87
Other*Female	1.36	1.36	1.31	1.30	0.13	1.28
8th Grade Absences (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.29***	2.31***	2.30***	2.30***	2.30***
Middle School Principal's Suspension		2.57***	2.51***	2.48***	2.49***	2.49***
8th Grade ELA Test Score[2]		0.85***	0.86***	0.86***	0.86***	0.86***
8th Grade Math Test Score[2]		0.76***	0.78***	0.79***	0.79***	0.79***
8th Grade ELA Test Squared		0.98	0.99	0.99	0.99	0.99
8th Grade Math Test Squared		0.94***	0.95***	0.96***	0.96***	0.96***
High School Characteristics						
Percent Black/Hispanic[3]			0.96***	0.95***	0.93***	0.94***
Percent Free/Reduced Price Lunch[3]			1.00	1.00	1.00	1.00
Percent Female[3]			1.06***	1.07***	1.08***	1.08***
Percent Overage[3]			1.19***	1.16***	1.11***	1.11***
Percent Special Education[3]			1.17***	1.07	1.08	1.08
School Size[4]						
Small School			0.73*	0.69**	0.78*	0.78*
Large School			0.46***	0.42***	0.40***	0.40***
Select School			0.60***	0.59***	0.57***	0.57***
Small*Select			1.33*	1.40*	1.41*	1.42*
Large*Select			2.18***	2.47***	2.45***	2.45***
Mean 8th Grade ELA Score				1.00	1.00	1.00
Mean 8th Grade Math Score				0.99	1.00	1.00
Percent Middle School Suspensions[3]				1.17**	1.18**	1.18**
Safety Survey Score[2]					0.99	1.01
Engagement Survey Score[2]					0.78***	0.78***
Academic Survey Score[2]					1.05	1.05
Communication Survey Score[2]					1.03	1.05
Female*Percent Black/Hispanic						0.99
Female*Safety Survey Score						0.97

*p<.05; **p<.01; ***p<.001; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points
[4]Compared to medium

Suspension During High School. To determine how student and school characteristics are related to the odds of being suspended during the first four years of high school, I estimate a final logit model, presented in Table 9. As with moving between the ninth and tenth grade models, we observe many similarities when we move from those models to the model predicting suspension at any point, both in terms of statistical and substantive significance. The biggest suspension gaps continue to be related to race, gender, and discipline history. Females remain less likely to be suspended than males, and black students are still more likely to be suspended than white students. These relationships continue to be dependent upon one another, with black males having higher odds of suspension than white males, and black females having higher odds of suspension than white females. Students who were suspended in middle school have odds of suspension that are two and a half times greater than those who were never suspended.

The relationships between suspension and school size and selectivity are similar to those uncovered in the 9th grade models, but the relationship between school-average academic ability and suspension is different. Here, there is a trivial negative relationship between mean eighth grade math scores and suspension. The school survey results are very close to the estimates in the ninth grade model, with higher Engagement scores related to decreased odds of suspension, and higher Safety scores related to lower odds of suspension for females as compared to males. However, there is no effect for males for higher levels of perceived safety of the school environment.

Table 9. The Relationship between the Odds of High School Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.53***	0.45***	0.45***	0.45***	0.45***	0.44***
Special Education	1.64***	1.19***	1.18***	1.19***	1.19***	1.19***
English Language Learner	0.76***	0.66***	0.66***	0.66***	0.68***	0.68***
Free/Reduced Price Lunch	0.93**	1.01	0.95*	0.95	0.97	0.97
Age in 9th Grade (in years)	1.01	0.92***	0.91***	0.91***	0.91***	0.91***
Race[1]						
Black	2.46***	1.82***	1.75***	1.77***	1.73***	1.78***
Hispanic	1.36***	1.38***	1.27***	1.30***	1.28***	1.30***
Asian	0.75***	0.70***	0.70***	0.71***	0.680***	0.68***
Other	1.16	1.01	0.98	0.99	0.96	0.98
Black*Female	1.24*	1.48***	1.44***	1.45***	1.43***	1.33**
Hispanic*Female	0.99	1.12	1.09	1.09	1.08	1.03
Asian*Female	1.10	0.86	0.86	0.86	0.85	0.84
Other*Female	1.01	1.05	1.03	1.03	0.99	0.94
8th Grade Absences (Days)		1.00	1.00***	1.00***	1.00***	1.00***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.54***	2.55***	2.52***	2.53***	2.53***
Middle School Principal's Suspension		2.39***	2.42***	2.36***	2.38***	2.38***
8th Grade ELA Test Score[2]		0.81***	0.82***	0.83***	0.83***	0.83***
8th Grade Math Test Score[2]		0.77***	0.79***	0.80***	0.80***	0.80***
8th Grade ELA Test Squared		0.96	0.97***	0.97***	0.97***	0.97***
8th Grade Math Test Squared		0.96	0.96***	0.96***	0.96***	0.96***
High School Characteristics						
Percent Black/Hispanic[3]			0.97***	0.95***	0.94***	0.94***
Percent Free/Reduced Price Lunch[3]			1.02**	1.02**	1.02***	1.02***
Percent Female[3]			1.06***	1.07***	1.08***	1.09***
Percent Overage[3]			1.17***	1.13***	1.09***	1.09***
Percent Special Education[3]			1.16***	1.01	1.00	1.00
School Size[4]						
Small School			0.92	0.83	0.89	0.89
Large School			0.74**	0.64***	0.59***	0.59***
Select School			0.87	0.83	0.79*	0.79***
Small*Select			0.96	1.04	1.08	1.08
Large*Select			1.46***	1.78	1.81***	1.80***
Mean 8th Grade ELA Score				1.00	1.00	1.00
Mean 8th Grade Math Score				0.99**	0.99*	0.99*
Percent Middle School Suspensions[3]				1.30***	1.30***	1.30***
Safety Survey Score[2]					0.92**	0.96
Engagement Survey Score[2]					0.78***	0.78***
Academic Survey Score[2]					1.01	1.01
Communication Survey Score[2]					1.13***	1.14**
Female*Percent Black/Hispanic						1.01
Female*Safety Survey Score						0.91***

*p<.05; **p<.01; ***p<.001; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points
[4]Compared to medium

Multilevel Modeling with Adaptive Centering

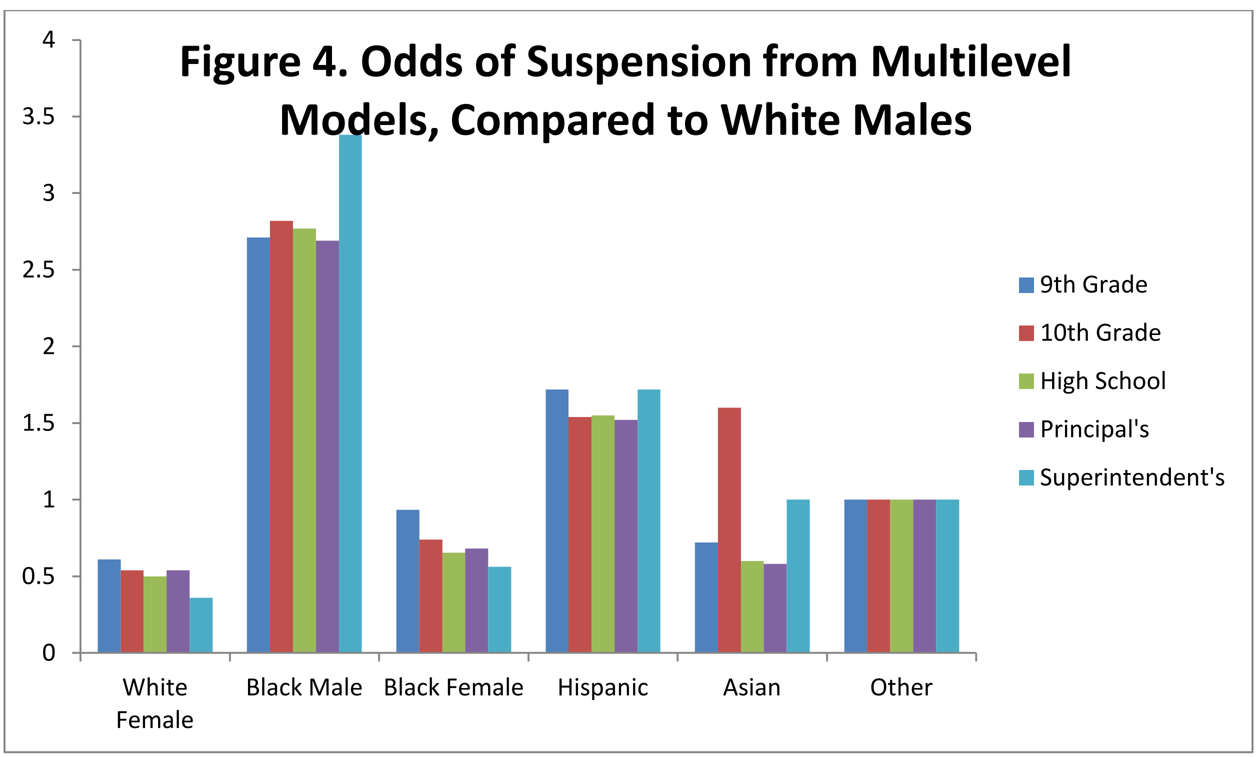
These initial estimates, like those presented in Chapter 3, provide us with a sense of what student and school characteristics are related to suspension. However, some of the uncovered relationships may be spurious since simple logit models do not take into account the nested nature of the data. This represents a central limitation of much extant research. To account for the nesting of students in schools, I calculate my final estimates of the relationships between suspension and student and school characteristics using multilevel logistic regression with adaptive centering. In these models, all student-level variables are centered within schools, allowing me to estimate the relationship between suspension and student-level characteristics within schools, while simultaneously explaining some of the variation in the odds of suspension between schools using school-level characteristics. An additional benefit of this approach is that it correctly accounts for the dependence of observations within schools, providing me with more accurate estimates of the statistical significance of the results. As I did previously, I investigate these relationships in ninth grade, in tenth grade, and in all of high school. As displayed in Table 10, between five and nine percent of the variation in likelihood of suspension occurs between schools, and including student-and school level controls cuts this variation to four to five percent of the variation between schools. I also conduct two additional sets of analyses to determine whether these relationships differ according to whether students earned a Principal's or Superintendent's suspension.

Table 10. Intraclass Correlations for Suspension Status

	Unconditional	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
9th Grade Suspension	0.051	0.052	0.048	0.042	0.040	0.037	0.037
10th Grade Suspension	0.070	0.071	0.066	0.057	0.055	0.052	0.052
Suspension in High School	0.088	0.091	0.083	0.067	0.060	0.054	0.054

Note: Model numbers correspond to models presented in Tables 11-13.

Ninth Grade Suspension. Table 11 displays the relationship between the odds of ninth grade suspension and student and school characteristics using multilevel modeling with adaptive centering. Model 1 presents the relationship between suspension and student socio-demographic measures. These relationships present within-school estimates, and are substantively the same as the relationships presented in the simple logistic regression. As before, the racial/ethnic and gender gaps are extreme. Since these results are difficult to interpret and are critical for understanding characteristics predicting suspension, racial/ethnic and gender gaps are graphically represented in Figure 4. Within schools, white females have forty percent lower odds of being suspended than their male peers; black females have about fifty percent higher odds of being suspended compared to white males than do white females compared to white males; and black males have about 270 percent higher odds of suspension compared to white males. Hispanic students have 72 percent higher odds of suspension, whereas Asian students have 28 percent lower odds of suspension, than their white peers. Based on the non-significant interactions, there are no gender disparities in the odds of suspension for Hispanic and Asian students.



Special education students are also far more likely to be suspended, with the odds of suspension 76 percent higher than the odds for general education students. English Language Learners, on the other hand, have considerably lower odds of suspension, when compared with the English-speaking students at their school. Older students have slightly higher odds of suspension, with a one year increase in age linked to a nine percent increase in the odds of suspension. Finally, students who receive Free or Reduced Price Lunch have somewhat lower odds of suspension than do their peers above the poverty line. However, it is likely that this relationship is not representative of the poverty gap, as the students who are labeled impoverished are those who were willing, able, and responsible enough to turn in their lunch forms.

In Model 2, I control for entering high school academic characteristics. Doing so substantially reduces the difference in odds of suspension between special education and general education students and black and white males, and completely explains the relationship between age and the odds of being suspended. After holding students' eighth grade characteristics constant, the odds that black males are suspended are two times those of white males, and students with IEPs have 30 percent higher odds of being suspended than do their general education peers. Better eighth grade attendance and fewer latenesses are associated with slightly lower odds of suspension, as are higher eighth grade English and math skills. The relationship between eighth grade test scores and suspension continues to appear nonlinear, with smaller decreases in the odds of suspension occurring for test scores higher than one standard deviation above the mean. The strongest relationships continue to be between suspension and race/ethnicity, gender, and discipline history. Students who were assigned a long-term suspension in middle school have odds of suspension that are almost three times greater than those of students who were never assigned a suspension. Students who were assigned a Principal's suspension have 140 percent higher odds of suspension than do students who were never suspended.

Because some of the variation in odds of suspension lies between schools, I include school-level measures in Model 3, which permits me to explain differences in suspension rate across schools. Controlling for these school-level effects leaves the within-school student estimates substantively unchanged. These Model 3 estimates statistically and substantively differ from those produced in the third model of the one-level logistic regression. Unlike the results presented in Table 7, I find no relationship between suspension status and school-level poverty, gender, age, size, or selectivity. These relationships previously appeared significant due to the fact that single-level models do not account for the dependent error structure of the observations and inflate the statistical power associated with school-level variables.

I do find a relationship between the racial/ethnic composition of the school and the odds of suspension, but it is now in the direction anticipated by the literature: increasing percentages of black/Hispanic students are associated with higher odds of suspension. I also find a stronger relationship between special education status and suspension than discovered in logistic regression. In this model, a ten percentage point increase in special education students is associated with 78 percent higher odds of suspension.

In Model 4, I add school-level measures of eighth grade ELA and Math skills, as well as the percent of ninth graders who earned middle school suspensions. Importantly, including these controls completely explains the relationship between the racial/ethnic composition of the student body and suspensions. It also decreases the odds gap due to special education status to 34 percent. This reduction is a result of the strong relationship between percent of middle school suspendees and high school suspension: a ten percentage point increase in middle school offenders is linked to a 61 percent increase in the odds of suspension. The estimated relationship between mean math scores and suspension, on the other hand, is weak: a one standard deviation increase in mean math scores is only associated with two percent lower odds of suspension.

To better understand how school environments relate to the odds of suspension, I include the School Survey factors in Model 5. Unlike in the simple logistic regression models, the only survey factor that is related to suspension is the Safety factor, with a one standard deviation increase associated with 25 percent lower odds of suspension. Not only does this finding make intuitive sense, but it also signals that the NYCDOE is collecting valuable information about the learning environment from the surveys. This relationship likely represents bidirectional causality, as suspended students may make the environment more disorderly, and a disorderly environment may lead to more students being suspended. These relationships explain the association between mean math scores and suspension from Model 4.

Finally, I examine which school-level characteristics influence the relationships between student-level characteristics and suspension. I find two mediating factors: the racial ethnic composition of the school and how safe the learning environment is. In particular, increasing percentages of black/Hispanic students are associated with increasing likelihood that females are assigned suspensions. A ten percentage point increase in the percent of black/Hispanic students is related to four percent higher odds of suspension for females. Furthermore, the safer the school environment, the less likely it is that females will be suspended: a one standard deviation increase on the Safety factor is related to 16 percent lower odds of suspension for females versus males. Notably, the relationship between the percent of students assigned middle school suspensions and the odds of high school suspension remains relatively unchanged across models.

Table 11. The Relationship between the Odds of 9th Grade Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.61***	0.64***	0.64***	0.64***	0.64***	0.43***
Special Education	1.76***	1.29***	1.28***	1.28***	1.28***	1.29***
English Language Learner	0.64***	0.62***	0.62***	0.62***	0.62***	0.62***
Free/Reduced Price Lunch	0.87**	0.90*	0.90*	0.90*	0.90*	0.90*
Age in 9th Grade (in years)	1.09*	0.96	0.96	0.96	0.96	0.96
Race[1]						
Black	2.71***	2.08***	2.11***	2.10***	2.11***	2.18***
Hispanic	1.72***	1.45***	1.47***	1.46***	1.46***	1.52***
Asian	0.72**	0.80*	0.78*	0.78*	0.79*	0.80
Other	1.39	1.04	1.06	1.05	1.06	1.09
Black*Female	1.53*	1.60**	1.63**	1.63**	1.63**	1.78**
Hispanic*Female	1.29	1.27	1.29	1.29	1.29	1.43*
Asian*Female	1.48	1.45	1.47	1.47	1.47	1.42
Other*Female	0.99	0.91	0.94	0.94	0.94	1.01
8th Grade Absences (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.78***	2.75***	2.74***	2.74***	2.75***
Middle School Principal's Suspension		2.40***	2.39***	2.38***	2.38***	2.37***
8th Grade ELA Test Score[2]		0.83***	0.83***	0.83***	0.83***	0.83***
8th Grade Math Test Score[2]		0.84***	0.84***	0.84***	0.84***	0.84***
8th Grade ELA Test Squared		0.96**	0.96**	0.96***	0.96**	0.96**
8th Grade Math Test Squared		0.98*	0.97*	0.97*	0.97*	0.97**
High School Characteristics						
Percent Black/Hispanic[3]			1.13***	1.07	1.01	1.02
Percent Free/Reduced Price Lunch[3]			1.00	1.00	1.02	1.02
Percent Female[3]			1.01	1.03	1.06	1.06
Percent Overage[3]			1.09	0.99	0.94	0.94
Percent Special Education[3]			1.78***	1.34*	1.23	1.23
School Size[4]						
Small School			1.09	0.92	1.17	1.18
Large School			1.17	0.84	0.70	0.71
Select School			0.84	0.84	0.90	0.90
Small*Select			0.85	0.99	0.89	0.88
Large*Select			1.09	1.78	1.66	1.66
Mean 8th Grade ELA Score				1.00	1.00	1.00
Mean 8th Grade Math Score				0.98*	0.99	0.99
Percent Middle School Suspensions[3]				1.61*	1.56*	1.56*
Safety Survey Score[2]					0.75*	0.73*
Engagement Survey Score[2]					0.83	0.84
Academic Survey Score[2]					1.06	1.06
Communication Survey Score[2]					1.05	1.05
Female*Percent Black/Hispanic						1.04*
Female*Safety Survey Score						0.84***
Constant	0.04	0.04***	0.01***	15785*	14592	13945

*p<.05; **p<.01; ***p<.001; All student-level variables group-mean centered; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points; [4]Compared to medium

Tenth Grade Suspension. Because the relationships between suspension and student and school characteristics may differ across grades, I also estimate these relationships in tenth grade, when the suspension rate is highest. Table 12 displays these relationships. I begin with Model 1, in which I include students' socio-demographic background characteristics. The tenth grade associations are analogous to those of Model 1 of the ninth grade multilevel model. White females have about 50 percent lower odds than white males of being suspended, and black males have odds of suspension 2.82 times that of white males. The odds ratio for black female suspension to white male suspension is about 40 percent than that of white females to black males, and Hispanic students have 50 percent greater odds of suspension than their white peers. Special education students are also considerably more likely to be suspended, while English Language Learners, Asians, and students receiving Free- or Reduced-Price Lunch have lower odds of suspension.

In Model 2, including students' middle school academic histories explains part of the relationship between suspension and special education status and race, but leaves the racial gaps substantively unchanged. Most glaringly, black males still have over two times the odds of suspension as white males. Adjusting for all student-level characteristics, a weak relationship between age and suspension appears, with older students having slightly lower odds of being suspended. As with Model 2 in ninth grade, there is a faint relationship between suspension and eighth grade attendance and lateness, a slight, linear relationship between suspension and eighth grade ELA abilities, and a moderate, nonlinear relationship between suspension and entering math abilities. Similar to ninth grade suspension, the strongest predictor of tenth grade suspension is middle school suspension status, with students who receive either a Superintendent's or a Principal's suspension having over 150 percent higher odds of suspension.

In Models 3, 4, and 5, I add school-level covariates to explain the between-school differences in the odds of suspension. Although there is a negative relationship between the racial/ethnic composition of the school and the odds of suspension in Models 3 and 4, this relationship is explained by the inclusion of the aggregate measure of middle school suspension. The middle school discipline rate also explains the weak relationship between school suspension and entering math abilities and special education. The relationship between high school suspension and middle school discipline rates is strong: a ten percent increase in the percentage of students who were assigned middle school suspension is associated with a 50 percent increase in the odds of high school suspension. This relationship persists in the final model, in which we see that the relationship between gender and 10th grade suspension rate does not depend on school-level race/ethnicity or the school safety factor.

Table 12. The Relationship between the Odds of 10th Grade Suspension and Student and School Characteristics. (n=68,424)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.54***	0.57***	0.56***	0.56***	0.56***	0.50***
Special Education	1.69***	1.24***	1.23***	1.23***	1.23***	1.23***
English Language Learner	0.72***	0.68***	0.68***	0.68***	0.68***	0.68***
Free/Reduced Price Lunch	0.83***	0.85***	0.85***	0.85***	0.85***	0.85***
Age in 9th Grade (in years)	1.04	0.93*	0.93*	0.93*	0.93*	0.93*
Race[1]						
Black	2.82***	2.16***	2.18***	2.17***	2.17***	2.19***
Hispanic	1.54***	1.30**	1.32***	1.32**	1.32**	1.33***
Asian	0.60***	0.66***	0.64***	0.64***	0.64***	0.64***
Other	1.44	1.15	1.17	1.16	1.16	1.17
Black*Female	1.37*	1.42*	1.44*	1.44*	1.44*	1.47*
Hispanic*Female	1.10	1.07	1.09	1.08	1.08	1.11
Asian*Female	1.13	1.11	1.11	1.11	1.11	1.07
Other*Female	1.36	1.36	1.39	1.38	1.38	1.41
8th Grade Absences (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.56***	2.54***	2.53***	2.53***	2.53***
Middle School Principal's Suspension		2.62***	2.61***	2.60***	2.60***	2.60***
8th Grade ELA Test Score[2]		0.87***	0.87***	0.87***	0.87***	0.87***
8th Grade Math Test Score[2]		0.79***	0.79***	0.79***	0.79***	0.79***
8th Grade ELA Test Squared		0.99	0.99	0.99	0.99	0.99
8th Grade Math Test Squared		0.96***	0.96***	0.96***	0.96***	0.96***
High School Characteristics						
Percent Black/Hispanic[3]			1.13***	1.08*	1.03	1.03
Percent Free/Reduced Price Lunch[3]			0.98	0.98	1.00	1.00
Percent Female[3]			0.96	0.97	0.99	0.99
Percent Overage[3]			1.15	1.10	1.04	1.04
Percent Special Education[3]			1.68***	1.28	1.20	1.20
School Size[4]						
Small School			1.03	0.89	1.22	1.22
Large School			0.81	0.60	0.53	0.53
Select School			0.66	0.65	0.73	0.73
Small*Select			0.98	1.11	0.93	0.93
Large*Select			1.52	2.30	2.06	2.06
Mean 8th Grade ELA Score				1.01	1.01	1.01
Mean 8th Grade Math Score				0.98*	0.98	0.98
Percent Middle School Suspensions[3]				1.56*	1.50*	1.50*
Safety Survey Score[2]					0.82	0.81
Engagement Survey Score[2]					0.89	0.89
Academic Survey Score[2]					1.09	1.09
Communication Survey Score[2]					0.91	0.91
Female*Percent Black/Hispanic						1.01
Female*Safety Survey Score						0.94
Constant	0.05***	0.05***	0.01***	68.2	239.8	236.9

*p<.05; **p<.01; ***p<.001; All student-level variables group-mean centered; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points; [4]Compared to medium

Suspension During High School. To improve upon the one-level logistic regression estimates predicting suspension at any point during high school, I present multilevel estimates in Table 13. These estimates are statistically and substantively similar to the estimates in the models predicting suspension in ninth grade. In Model 1, race, gender, language status, special education status, and poverty, all appear to be related to the odds of suspension. These estimates are slightly altered with the inclusion of middle school academic characteristics in Model 2, but then remain relatively unchanged through Model 6. As was the case with ninth and tenth grade suspension, within-school student-level measures are far more predictive of the odds of suspension than are school-level characteristics. In the final model, there is a nonlinear, negative relationship between suspension and middle school English and math abilities, and a trivial positive relationship between suspension and absence and lateness.

The strongest associations continue to be between suspension and race, gender and discipline history, with black males having twice the odds of suspension as white males, and black females having 40 percent higher odds ratio of suspension than white females. White females also have lower odds of suspension than white males. However, the relationship between gender and suspension also depends upon school-level characteristics. Safer school environments are related to decreased odds of suspension for all students, but this relationship is even stronger for females. On the other hand, females have slightly higher odds of suspension in schools with higher percentages of black/Hispanic students.

Prior suspensions, both at the individual level and in the aggregate, heavily relate to the odds of suspension. Students who were assigned a Principal's suspension in middle school have 151 percent higher odds of suspension and students who were assigned a Superintendent's suspension have 180 percent higher odds of suspension than do students who were never assigned a middle school suspension. At the school-level, a ten percentage increase in the percent of students who earned suspension is associated with a 77 percentage increase in the odds of suspension.

Table 13. The Relationship between the Odds of High School Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.50***	0.51***	0.51***	0.51***	0.51***	0.40***
Special Education	1.62***	1.19***	1.19***	1.19***	1.19***	1.19***
English Language Learner	0.77***	0.71***	0.71***	0.71***	0.71***	0.71***
Free/Reduced Price Lunch	0.92**	0.94*	0.94*	0.94*	0.94*	0.94*
Age in 9th Grade (in years)	1.01	0.90***	0.90***	0.90***	0.90***	0.91***
Race[1]						
Black	2.77***	2.14***	2.15***	2.14***	2.15***	2.17***
Hispanic	1.55***	1.31***	1.32***	1.32***	1.32***	1.34***
Asian	0.60***	0.65***	0.64***	0.64***	0.65***	0.64***
Other	1.31	1.01	1.01	1.01	1.01	1.03
Black*Female	1.31**	1.34**	1.35**	1.35**	1.35**	1.40**
Hispanic*Female	1.10	1.08	1.08	1.08	1.08	1.13
Asian*Female	1.13	1.11	1.11	1.11	1.11	1.06
Other*Female	1.02	1.00	1.01	1.01	1.02	1.05
8th Grade Absences (Days)		1.00***	1.00***	1.00***	1.00***	1.00***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.81***	2.80***	2.79***	2.79***	2.80***
Middle School Principal's Suspension		2.53***	2.52***	2.52***	2.52***	2.51***
8th Grade ELA Test Score[2]		0.84***	0.84***	0.84***	0.84***	0.84***
8th Grade Math Test Score[2]		0.81***	0.81***	0.81***	0.81***	0.81***
8th Grade ELA Test Squared		0.97***	0.97***	0.97***	0.97***	0.97***
8th Grade Math Test Squared		0.97***	0.97***	0.97***	0.97***	0.97***
High School Characteristics						
Percent Black/Hispanic[3]			1.13***	1.08*	1.02	1.03
Percent Free/Reduced Price Lunch[3]			0.99	0.99	1.00	1.00
Percent Female[3]			0.96	0.98	1.00	1.00
Percent Overage[3]			1.14*	1.07	1.03	1.03
Percent Special Education[3]			1.72***	1.25*	1.14	1.14
School Size[4]						
Small School			1.07	0.89	1.06	1.06
Large School			1.01	0.70	0.59	0.59
Select School			0.83	0.79	0.82	0.82
Small*Select			0.83	0.97	0.93	0.93
Large*Select			1.19	1.95	1.86	1.86
Mean 8th Grade ELA Score				1.01	1.00	1.00
Mean 8th Grade Math Score				0.98**	0.99*	0.99*
Percent Middle School Suspensions[3]				1.84***	1.76***	1.77***
Safety Survey Score[2]					0.76**	0.75**
Engagement Survey Score[2]					0.86	0.87
Academic Survey Score[2]					1.01	1.01
Communication Survey Score[2]					1.09	1.09
Female*Percent Black/Hispanic						1.03*
Female*Safety Survey Score						0.91***
Constant	0.15***	0.15***	0.03***	253.8	173.9	173.4

*p<.05; **p<.01; ***p<.001; All student-level variables group-mean centered; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points; [4]Compared to medium

Principal's and Superintendent's Suspensions. Not only could the relationship between suspension and student and school characteristics vary between years, but it also could vary according to whether students were assigned a short- or long-term suspension. Thus, I estimate two final sets of models, one that examines the odds of receiving a Principal's suspension and one that calculates the odds of being assigned a Superintendent's suspension.

Table 14 displays the Principal's suspension predictions. These estimates are virtually the same as those for the prediction of any suspension in high school, presented in Table 13. In Model 1, white males have drastically lower odds of suspension than black males, and white females have considerably lower odds of suspension than black females. Special education and Hispanic students are also more likely to be suspended, whereas Asian students, English Language Learners, and students receiving free or reduced price lunch as less likely to be suspended.

Controlling for entering high school characteristics in Model 2 explains the relationship between free or reduced price lunch and suspension and decreases the magnitude of the relationship among suspension and race and gender, as well as between suspension and special education status. In addition, it surfaces a weak relationship between age and suspension, with older students having slightly lower odds of suspension, and nonlinear, negative relationships between eighth grade English and math test scores and the odds of suspension. Finally, it reveals the strong relationship between discipline history and odds of suspension, with prior suspension being associated with a two and a half increase in the odds of suspension.

These relationships endure through Model 6, in which I include all student- and school-level measures and interactions. In this final model, there is an even stronger relationship between middle school suspension rates and high school suspension, with a 10 percentage point increase in middle school suspension rate associated with a doubling of the odds of high school suspension. How safe the school environment is continues to be negatively related to the odds of suspension, with an even bigger effect for females over males. In addition, we continue to see a trivial interaction between the racial composition of the student body and gender, with females having slightly higher odds of suspension at higher levels of black/Hispanic population.

Table 14. The Relationship between the Odds of High School Principal's Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.54***	0.56***	0.56***	0.56***	0.56***	0.44***
Special Education	1.63***	1.20***	1.19***	1.19***	1.19***	1.20***
English Language Learner	0.79***	0.72***	0.72***	0.72***	0.72***	0.72***
Free/Reduced Price Lunch	0.93*	0.96	0.96	0.96	0.96	0.95
Age in 9th Grade (in years)	0.98	0.87***	0.87***	0.87***	0.87***	0.88***
Race[1]						
Black	2.69***	2.07***	2.07***	2.07***	2.07***	2.10***
Hispanic	1.52***	1.28***	1.29***	1.29***	1.29***	1.30***
Asian	0.58***	0.63***	0.63***	0.63***	0.63***	0.63***
Other	1.33	1.03	1.03	1.04	1.03	1.04
Black*Female	1.26*	1.29*	1.30*	1.30*	1.30*	1.34**
Hispanic*Female	1.06	1.03	1.04	1.04	1.04	1.08
Asian*Female	1.16	1.14	1.14	1.14	1.14	1.08
Other*Female	1.09	1.07	1.08	1.08	1.08	1.12
8th Grade Absences (Days)		1.00	1.00***	1.00***	1.00***	1.00***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		2.49***	2.48***	2.48***	2.48***	2.48***
Middle School Principal's Suspension		2.44***	2.44***	2.43***	2.43***	2.42***
8th Grade ELA Test Score[2]		0.85***	0.85***	0.85***	0.85***	0.85***
8th Grade Math Test Score[2]		0.79***	0.79***	0.79***	0.79***	0.79***
8th Grade ELA Test Squared		0.97***	0.97***	0.97***	0.97***	0.97***
8th Grade Math Test Squared		0.97***	0.96***	0.96***	0.96***	0.96***
High School Characteristics						
Percent Black/Hispanic[3]			1.11**	1.05	1.00	1.00
Percent Free/Reduced Price Lunch[3]			0.99	0.98	1.00	1.01
Percent Female[3]			0.97	0.99	1.02	2.02
Percent Overage[3]			1.13	1.06	1.05	1.01
Percent Special Education[3]			1.76***	1.24	1.11	1.11
School Size[4]						
Small School			0.97	0.79	0.89	0.89
Large School			1.02	0.68	0.55	0.55
Select School			0.76	0.71	0.71	0.71
Small*Select			0.89	1.08	1.09	1.09
Large*Select			1.27	2.21	2.18	2.17
Mean 8th Grade ELA Score				1.01	1.00	1.00
Mean 8th Grade Math Score				0.98*	0.99	0.99
Percent Middle School Suspensions[3]				2.06***	1.99***	1.99***
Safety Survey Score[2]					0.72**	0.71**
Engagement Survey Score[2]					0.79	0.79
Academic Survey Score[2]					1.02	1.02
Communication Survey Score[2]					1.22	1.22
Female*Percent Black/Hispanic						1.03*
Female*Safety Survey Score						0.89***
Constant	0.11***	0.11***	0.03***	576.0	252.3	253.2

*p<.05; **p<.01; ***p<.001; All student-level variables group-mean centered; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points; [4]Compared to medium

Table 15 presents the Superintendent's suspension predictions. These findings reveal important information about the relationship between suspension and race/ethnicity. Indeed, across all models, the racial suspension gaps are even greater in Superintendent's suspension than they are in Principal's suspension. In the final model, black males have odds three times those of white males of receiving a Superintendent's suspension, and black females have odds that are twice those of white females. Unlike the previous models, suspension gaps for Hispanic students also depend upon gender, with Hispanic males having 77 percent higher odds of suspension than white males and Hispanic females having 69 percent higher odds of suspension than white females. The racial/ethnic composition of the student body is also related to the odds of suspension, with higher odds of suspension associated with higher percentages of black and Hispanic students. This relationship is even stronger for female students. The safety of the school environment is negatively related to the odds of Superintendent's suspension, but does not interact with gender. Lastly, discipline history remains one of the strongest predictors of suspension and the school and student level. In particular, having at least one middle school Superintendent's suspension is linked to a tripling of the odds of being assigned a Superintendent's suspension in high school.

Table 15. The Relationship between the Odds of High School Superintendent's Suspension and Student and School Characteristics. (n=69,050)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student Characteristics						
Female	0.36***	0.37***	0.37***	0.37***	0.37***	0.15***
Special Education	1.64***	1.24***	1.23***	1.23***	1.23***	1.23***
English Language Learner	0.69***	0.67***	0.67***	0.67***	0.67***	0.67***
Free/Reduced Price Lunch	0.85***	0.88**	0.88**	0.88**	0.88**	0.88**
Age in 9th Grade (in years)	1.05	0.94	0.94*	0.94*	0.94*	0.94
Race[1]						
Black	3.38***	2.69***	2.82***	2.81***	2.83***	3.15***
Hispanic	1.72***	1.48***	1.57***	1.57***	1.58***	1.77***
Asian	0.77	0.84	0.83	0.82	0.83	0.84
Other	1.19	0.93	0.97	0.97	0.98	1.06
Black*Female	1.56*	1.63*	1.73*	1.72*	1.73*	2.26**
Hispanic*Female	1.18	1.19	1.26	1.26	1.27	1.69*
Asian*Female	1.41	1.41	1.46	1.46	1.47	1.39
Other*Female	0.46	0.45	0.47	0.47	0.48	0.59
8th Grade Absences (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
8th Grade Latenesses (Days)		1.01***	1.01***	1.01***	1.01***	1.01***
Middle School Sup's Suspension		3.38***	3.32***	3.30***	3.30***	3.31***
Middle School Principal's Suspension		2.25***	2.24***	2.23***	2.23***	2.22***
8th Grade ELA Test Score[2]		0.84***	0.84***	0.84***	0.84***	0.84***
8th Grade Math Test Score[2]		0.87**	0.87**	0.86**	0.86**	0.86**
8th Grade ELA Test Squared		0.99	0.99	0.99	0.99	0.99
8th Grade Math Test Squared		0.97**	0.97**	0.97**	0.96**	0.96**
High School Characteristics						
Percent Black/Hispanic[3]			1.20***	1.17***	1.12***	1.16***
Percent Free/Reduced Price Lunch[3]			1.00	1.00	1.01	1.01
Percent Female[3]			0.92***	0.92**	0.94*	0.94*
Percent Overage[3]			1.17**	1.14*	1.09	1.09
Percent Special Education[3]			1.50***	1.25*	1.16	1.16
School Size[4]						
Small School			1.14	1.02	1.28	1.28
Large School			0.93	0.76	0.68	0.68
Select School			0.84	0.81	0.88	0.88
Small*Select			0.89	0.97	0.85	0.85
Large*Select			1.00	1.43	1.31	1.31
Mean 8th Grade ELA Score				1.01*	1.01	1.01
Mean 8th Grade Math Score				0.98*	0.99	0.99
Percent Middle School Suspensions[3]				1.36*	1.29*	1.30*
Safety Survey Score[2]					0.82*	0.81*
Engagement Survey Score[2]					0.99	0.99
Academic Survey Score[2]					0.94	0.94
Communication Survey Score[2]					0.96	0.97
Female*Percent Black/Hispanic						1.11***
Female*Safety Survey Score						0.93
Constant	0.04***	0.04***	0.01***	0.40	0.35	0.29

*p<.05; **p<.01; ***p<.001; All student-level variables group-mean centered; [1]Compared to white; [2]Measure is z-scored; [3]Unit is 10 percentage points; [4]Compared to medium

Discussion

When examined together, these models reveal many important findings about the relationship between suspension and student and school characteristics. First, utilizing single-level logistic regression to model the relationship between school level characteristics and student-level suspension may lead to incorrect statistically significant findings, as these models ignore the dependent nature of observations within schools. The fact that some of the significant relationships estimated in extant literature are not significant in my multi-level models, such as enrollment and school-level academic abilities, could be due to the fact that I have nested students within schools. These differences also may be due to the fact that I have access to a measure indicating the safety of the school environment, and that I include aggregate levels of middle school suspension. Previous studies may be picking up information about the school learning environment and school-level discipline history when they find significant relationships between academic skills and likelihood of suspension.

Second, within-school student characteristics are far more predictive of the odds of suspension than are between-school measures. On both levels, however, discipline history is strongly related to the odds of suspension, so much so that it often explains other relationships purported to be related to suspension status, such as racial/ethnic school composition. This finding calls into question the efficacy of suspension if suspensions are supposed to either act as a deterrent for future negative behavior. If this were the case, we would expect there to be a negative relationship between middle school suspension and high school suspension. Of course, the positive relationship that is present in all models could also be a result of the way students are perceived: having been suspended in the past, these students may be labeled as troublesome and school personnel may be more likely to suspend them in high school. The aggregate findings also suggest that peer effects play a role in the likelihood of suspension, both in terms of the school-level middle school suspension rate and the safety of the school environment, as increased odds of suspension are associated with more unsafe environments. Critically, there is likely cyclic causality in this relationship, as having a more unsafe environment could lead to more offences, which could lead to a more unsafe environment. Likewise, the causal chain could begin with having more students who tend to exhibit unsafe behaviors.

Across all models, the racial/ethnic and special education gaps in suspension rates may be most troubling, as my research confirms the conclusions drawn in extant literature: black males are excluded at much higher rates than white males. Black females are also suspended at higher rates than white females. Hispanic students and special education students are also more likely to be excluded from school. These gaps mirror academic gaps that exist throughout students' k-12 academic trajectory, thereby raising question about the role suspension plays in these students' high school outcomes, to which I now turn my attention in the following chapter.

Finally, my results highlight the subtle differences that exist in these relationships across time and between different kinds of suspension. For example, the school environment is predictive of suspensions in ninth grade, but not in tenth. This may be because students are more heavily influenced by their context when they first enter high school. Disaggregating the type of suspension also reveals stark differences in prediction estimates, with black/white racial gaps being much larger for Superintendent's suspension than for Principal's suspension.

Although I improve upon past work by introducing metrics that estimate perceptions of the learning environment and utilize multi-level modeling techniques, these models also have limitations. These models would be vastly improved by more valid measures of students' socioeconomic status. Much literature suggests that socioeconomic status is a strong predictor of likelihood of suspension. Given that socioeconomic status likely relates to many of the predictors of interest, such as entering academic skills, and that our measure may be more of a proxy for whether students are willing, able, and responsible enough to turn in their lunch forms than for socioeconomic status, these results are likely biased.

Ultimately, these results present a clear descriptive picture of which attributes are associated with suspension, but it does not allow me to make empirically-based inferences on why these attributes are related to suspension. This work is essential to disrupting the uneven distribution of discipline, as it will both enable policymakers and educators to determine when these relationships represent the unfair application of suspension, and it will also help educators identify interventions that may help reduce suspension incidents.

Chapter 5. Suspension and Short-term Outcomes

The unequal distribution of suspension by student social and academic backgrounds is even more problematic if suspensions are negatively related to student outcomes, as this would suggest that suspensions exacerbate socio-demographic inequality. To determine whether suspensions are linked to academic achievement, I first examine the relationship between suspensions and credit accumulation and attendance in high school. These are two crucial outcomes, as credit accumulation, especially in ninth grade, is predictive of high school graduation, and attendance is positively related to student achievement (Allensworth & Easton, 2005; Neild, 2009; Roderick & Camburn, 1999). Utilizing multilevel modeling is critical given that a substantial portion of the variation in these outcomes lies between schools, as shown in Tables 16 and 17.

Table 16. Intraclass Correlations for 9th Grade Outcomes

	Unconditional	Model 1	Model 2	Model 3
9th Grade Credits	0.112	0.107	0.145	0.120
9th Grade Math Credits	0.093	0.088	0.101	0.085
9th Grade English Credits	0.080	0.077	0.090	0.076
9th Grade Lateness	0.369	0.368	0.395	0.383
9th Grade Absence	0.145	0.140	0.068	0.053

Note: Model numbers correspond to models presented in Tables 18 and 19.

Table 17. Intraclass Correlations for 10th Grade Outcomes

	Unconditional	Model 1	Model 2	Model 3
10th Grade Credits	0.124	0.120	0.173	0.156
10th Grade Math Credits	0.096	0.092	0.112	0.098
10th Grade English Credits	0.093	0.089	0.106	0.097
10th Grade Lateness	0.435	0.435	0.459	0.449
10th Grade Absence	0.157	0.155	0.086	0.070

Note: Model numbers correspond to models presented in Tables 20 and 21.

Credit Accumulation and Attendance in Ninth Grade

Ninth Grade Credit Accumulation. Because success in ninth grade is closely related to graduation, I begin with an analysis limited to the ninth grade year. Table 18 displays the relationships between suspension in the first semester of ninth grade and credit accumulation in ninth grade. In order for students to graduate, they must earn 44 credits, including 6 math credits and 8 English credits. Consequently, in these models, which nest students within their ninth grade schools, I utilize three outcomes: percent of all classes passed, percent of math classes passed, and percent of English classes passed. In Model 1, I establish the unadjusted relationship between the percent of math classes passed and suspension. This relationship is strong: students receiving at least one Principal's suspension in the fall of ninth grade have a 23.6 percentage point lower pass rate than students who are never suspended. Additionally, students receiving at least one Superintendent's suspension have a math credits pass rate that is 31.4 percentage point lower than their non-suspended peers.

Since suspended students differ from students who never earn suspension, I add student-level covariates in Model 2. Adding these measures explains much of the relationship between suspension and the percentage of math credits earned. The reduction in passage rate associated with Principal's suspension decreases to 8.4 percentage points, and the reduction linked to Superintendent's suspension lowers to 8.3 percentage points. Although the magnitude of these estimates is not nearly as large as those in Model 1, suspension is still one of the strongest predictors of math passage rate, the others being entering math abilities and discipline history.

In Model 3, I include school-level covariates, as well as interactions between school and student-level variables. The only school-level variable that mediates the relationship between suspension and the percent of math credits earned is the percent of students who earned middle school suspensions. Therefore, other interactions were removed from the model in the interest of parsimony. In this model, we see that the association between Principal's suspension and math credit accumulation remains virtually unchanged, with a Principal's suspension now associated with an 8.3 percentage point decrease in math credits earned. The relationship between Superintendent's suspension and math credit passage rate differs in this model. This association now depends on school-level discipline history. Though a Superintendent's suspension is linked to a 9.4 percentage point drop in the percent of suspensions earned in schools with an average percent of students who earned middle school suspensions, students earning Superintendent's suspensions at schools that have below average percentages of middle school suspendees have even greater decreases in their math passage rate. For example, students receiving Superintendent's suspensions at schools with a middle school suspension rate of ten percentage points below average now have a 15.1 percentage point decrease in their math credit accumulation rate, compared to the 9.4 percentage point decrease suspended students experience at schools with average middle school suspension rates.

On the other hand, suspended students at schools with above average percentages of middle school suspension take less of a credit hit when they are suspended. For example, students assigned Superintendent's suspensions at schools with the percent of middle school suspendees 10 percentage points above the mean only have, on average, a 3.7 percentage point drop in the percent of math credits they pass. Although this finding may initially appear counterintuitive, it is worth noting that this interaction likely signals something about academic expectations at the school rather than about student's performance when they return from suspension. Since class passage is not an objective standard shared across schools, it is possible that schools with higher percentages of middle school suspendees have lower academic expectations, making it easier for students to pass classes upon return from suspension. Conversely, it is also possible that these schools have better systems for reintroducing suspended students back into their home school, as they may deal more frequently with Superintendent's suspensions, or that these schools are more likely to grant students credit based on their work at ALCs.

I conduct the same investigation for both percent of English credits passed and percent of all credits passed in ninth grade to determine whether the patterns that exist for math credit accumulation differ by subjects and if they are representative of what occurs across subjects. These results are also displayed in Table 18. Indeed, my results suggest that the patterns are similar across all three outcomes. With regard to English passage rates, in Model 1, the unadjusted gap between suspended and non-suspended students is large: suspensions are associated with an approximately 25 percentage point decrease in the percent of English credits passed. After controlling for student's socio-demographic and background characteristics in Model 2, this detrimental effect decreases to 10.6 percentage points for Principal's suspensions, and 4.2 percentage points for Superintendent's suspensions. Current and past disciplinary statuses remain two of the strongest predictors of English credit accumulation in this model, along with English Language learner status.

In Model 3, I add school-level covariates, as well as an interaction between school-level middle school suspension rates and high school suspension. As with mathematics, doing so reveals that the relationship between Superintendent's suspension and English credit accumulation more heavily depends upon the percent of students who earned middle school suspensions. Whereas students who attend schools with average levels of middle school suspendees have English passage rates 6 percentage points lower than their non-suspended peers, students assigned Superintendent, this relationship is more extreme for schools with below and above average middle school suspension rates. Dropping the middle school suspension rate by 10 percentage points below the mean is associated with an additional 11.1 percentage point decrease in the percent of English credits passed; however, increasing the middle school suspension rate by 10 percentage points above the mean is related to an English credit passage rate of 5.1 percentage points above that of non-suspendees, holding all other variables constant.

The relationship between suspension and the percentage of all credits that a student passes are analogous to the relationship between suspension and percentage of all math credits passed. Again, we see that the unadjusted relationship between suspension and passage rate is much stronger than the adjusted relationship. After holding school and student characteristics constant, being assigned a Principal's suspension in the first semester of ninth grade is linked to a 10.1 percentage point decrease in the percent of all credits earned in ninth grade. The relationship between Superintendent's suspension and credit accumulation continues to be dependent upon the aggregate level of middle school suspension. The magnitude of this dependency is similar to that in the math model, with a 10 percentage point change in the middle school attendance rate associated with a 6 percentage point change in percent of credits earned. In addition to high school suspension, students' middle school suspension records continue to be the strongest predictors of credit accumulation, followed by student- and school-level math abilities.

Table 18. The Relationship between Suspension and Credits Earned in 9th Grade

	Percent of Math Credits			Percent of English Credits			Percent of Credits Passed		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Principal's Suspension	-0.24***	-0.08***	-0.08***	-0.24***	-0.11***	-0.11***	-0.24***	-0.10***	-0.10***
Superintendent's Suspension	-0.31***	-0.08***	-0.09***	-0.27***	-0.04**	-0.06***	-0.30***	-0.08***	-0.09***
Absences		-0.01***	-0.01***		-0.01***	-0.01***		-0.11***	-0.01***
Lateness		-0.01***	-0.01***		0.00***	0.00***		0.00***	0.00***
Female		0.05***	0.05***		0.06***	0.06***		0.04***	0.04***
Free/Reduced Price Lunch		0.03***	0.03***		0.03***	0.03***		0.03***	0.03***
Age in 9th Grade		0.00	0.00		0.00	0.00		0.00	0.00
Race[1]									
Black		-0.04***	-0.04***		-0.03***	-0.03***		-0.03***	-0.04***
Hispanic		-0.05***	-0.05***		-0.03***	-0.04***		-0.03***	-0.04***
Asian		-0.01	-0.01		-0.02***	-0.02***		-0.02***	-0.02***
Other		-0.07***	-0.08***		-0.06***	-0.06***		-0.05***	-0.05***
Special Education		0.06***	0.06***		0.02***	0.02***		0.03***	0.03***
English Language Learner		0.01	0.01		-0.05***	-0.05***		-0.02***	-0.02***
8th Grade Absences		0.00***	0.00***		0.00***	0.00***		0.00***	0.00***
8th Grade Latenesses		0.00***	0.00***		0.00***	0.00***		0.00***	0.00***
Middle School Superintendent's Suspension		-0.05***	-0.05***		-0.07***	-0.08***		-0.06***	-0.06***
Middle School Principal's Suspension		-0.08***	-0.08***		-0.07***	-0.07***		-0.07***	-0.07***
8th Grade ELA Exam[2]		0.00	0.00		0.02***	0.02***		0.02***	0.02***
8th Grade Math Exam[2]		0.10***	-0.10***		0.03***	0.03***		0.05***	0.05***
School Characteristics									
Percent Female[3]			0.00			0.00			0.01
Percent Free/Reduced Price Lunch[3]			0.01			0.00			0.00
Age in 9th Grade			0.01			0.06			0.01
Percent Black/Hispanic[3]			0.01**			0.01			0.01
Percent Special Education[3]			0.01			0.00			0.00
Percent English Language Learner[3]			0.00			-0.00			-0.00
Percent Middle School Suspension[3]			0.00			0.00			0.01
Mean 8th Grade ELA Score			0.02			0.01			0.02
Mean 8th Grade Math Score			-0.04			-0.03			-0.05**
9th Grade Superintendent's Suspension * Mean Middle School Suspension			0.06*			0.11***			0.06***
9th Grade Principal's Suspension * Mean Middle School Suspension			-0.02			-0.02			-0.01
Constant	0.79***	0.98***	0.96***	0.85***	1.03***	1.01***	0.84***	1.01***	0.99***
Variance Components									
School	0.12	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Residual	0.13	0.10	0.10	0.09	0.07	0.07	0.06	0.04	0.04

*p<.05; **p<.01; ***p<.001
All school-level variables grand-mean centered
[1]Compared to white
[2]Measure is z-scored
[3]Unit is 10 percentage points

Ninth Grade Attendance. I also examine whether suspensions are related to attendance, including absences and latenesses. These results are presented in Table 19. For both outcomes, Model 1 presents the unadjusted relationship. I find that receiving a Principal's suspension in the first semester of ninth grade is related to six additional latenesses, and over two weeks of additional absences. Moreover, being assigned a Superintendent's suspension is associated with four additional latenesses, and over four weeks of additional absences over the course of the ninth grade year.

In the second models, I include student characteristics. These characteristics explain all of the relationship between Superintendent's suspension and latenesses, and most of the relationship between Principal's suspension and latenesses. In other words, some of the lateness deficits we attributed to suspension in the first model were really a result of student characteristics, rather than the suspension itself. In addition, including these covariates cuts the absence effects of Principal's suspension in half, and the Superintendent's suspension by approximately eight days. However, being assigned a suspension is still related to over one week of additional absences for a Principal's suspension, and almost three weeks of additional absences for a Superintendent's suspension. Recall that students are only considered absent during their suspension if they do not show up for instruction at their school or at the ALC.

In the third models, I add school characteristics into the models, as well as interaction terms between student-level suspension and school-level middle school suspension rates. I find that the school-level suspension rate does not affect the relationship between student-level suspension and lateness, yet it does mediate the relationship between Superintendent's suspension and absences. The direction of this interaction is the opposite of what it was for credit accumulation. In other words, students receiving Superintendent's suspensions at schools with below average middle school suspension rates have *fewer* absences than do students with average or above average levels of middle school suspension rates. Each 10 percentage point change in middle school suspension rate is associated with a two and a half day change in expected absences. The gaps due to attendance remain large: students receiving a Principal's suspension miss over a week more of school during ninth grade, and students receiving a

Superintendent's suspension at schools with average levels of middle school suspension miss approximately three more weeks of school than non-suspended students.

Table 19. The Relationship between Suspension and Absence and Lateness in 9th Grade

	Days Late			Days Absent		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Principal's Suspension	6.05***	0.23***	2.10***	14.3***	6.50***	6.71***
Superintendent's Suspension	4.03***	0.00	0.11	23.0***	14.7***	15.1***
Absences		0.07***	0.07***		--	--
Lateness		--	--		0.17***	0.17***
Female		0.19	0.18		-0.01	-0.11
Free/Reduced Price Lunch		-0.08	-0.09		-0.91***	-0.94***
Age in 9th Grade		-0.03	-0.01		0.23	0.27*
Race[1]						
Black		2.20***	2.18***		0.40	0.26
Hispanic		1.56***	1.53***		1.73***	1.60***
Asian		0.06	0.05		0.91***	0.91***
Other		0.42	0.40		0.42	0.31
Special Education		0.31	0.29		1.62***	1.54***
English Language Learner		-1.24***	-1.22***		-1.06***	-0.97***
8th Grade Absences		0.00	0.00		0.70***	0.70***
8th Grade Latenesses		0.19	0.19***		0.07***	0.07***
Middle School Superintendent's Suspension		1.45**	1.43**		2.37***	2.35***
Middle School Principal's Suspension		1.66***	1.66***		3.07***	3.10***
8th Grade ELA Exam[2]		-0.47***	-0.47***		-0.38***	-0.36***
8th Grade Math Exam[2]		-0.48***	-0.48***		-1.04***	-1.01***
School Characteristics						
Percent Female[3]			0.39			0.01
Percent Free/Reduced Price Lunch[3]			0.33			-0.15
Age in 9th Grade			-3.54			6.24***
Percent Black/Hispanic[3]			0.76*			0.50**
Percent Special Education[3]			1.61**			0.59*
Percent English Language Learner[3]			0.51			-0.24
Percent Middle School Suspension[3]			-0.87			-0.83
Mean 8th Grade ELA Score			1.75			-0.37
Mean 8th Grade Math Score			-0.38			0.04
9th Grade Superintendent's Suspension * Mean Middle School Suspension			-0.79			-2.58*
9th Grade Principal's Suspension * Mean Middle School Suspension			1.58			-1.45
Constant	11.0***	6.87***	5.72**	15.8***	0.00	-1.02
Variance Components						
School	107.3	108.0	103.1	56.5	17.0	13.3
Residual	184.1	165.5	166.4	347.8	239.3	239.7

*p<.05; **p<.01; ***p<.001

All school-level variables grand-mean centered

[1]Compared to white

[2]Measure is z-scored

[3]Unit is 10 percentage points

Credit Accumulation and Attendance in Tenth Grade

Tenth Grade Credit Accumulation. The relationship between suspension and credit accumulation may vary between ninth and tenth grade; thus, I also estimate the relationship between suspension in the first semester of tenth grade and credit accumulation for math, English, and all classes during the tenth grade year. Table 20 displays these relationships. These results are statistically and substantively very similar to those from ninth grade. Looking at Model 1 for all tenth grade credit outcomes shows that without adjusting for student or school characteristics, Principal's and Superintendent's suspensions are associated with a 25 percentage point decrease in the percent of credits earned in math and English, and including all tenth grade courses.

Adjusting these estimates for student characteristics in the second models results in over a 50 percent decrease in the strength of the relationship between Principal's suspension and the percent of all credits, math credits and English credits earned, as well as a over an 80 percent decrease in the strength of the relationship between Superintendent's suspension and each credit outcome. Now, a Principal's suspension yields approximately a 20 percentage point decrease in the percent of credits earned, and a Superintendent's suspension is associated with a three to four percentage point decrease in the percent of credits earned.

In each of the third models, I control for school-level characteristics, and I estimate the effect of the percent of students earning middle school suspensions on the relationship between high school suspension and credit accumulation. With regard to math and English classes, the relationship between Principal's suspension and percent of credits earned remains unchanged from Model 2 to Model 3. In the case of all classes, however, this relationship depends on the percent of students who were assigned a middle school suspension, as does the relationship between Superintendent's suspension and percent of math, English, and all credits earned.

Students assigned a Superintendent's suspension at a school with the average percent of middle school suspendees have a credit passage rate that is 5.6 percentage points lower than their non-suspended peers, and students assigned a Principal's suspension rate have one 10 percentage points lower. If these students were at schools with middle school suspension rates 10 percentage points higher than average, they would be expected to increase their passage rate by 7.1 percentage points in the case of Superintendent's suspension, and 2.1 percentage points, in the case of Principal's suspension. As was the case with interpreting the ninth grade results, these results may seem unexpected. However, the higher passage rate at schools with higher percentages of students with discipline problems may be a result of lower academic expectations or improved support systems for students who are assigned suspensions.

Table 20. The Relationship between Suspension and Credits Earned in 10th Grade

	Percent of Math Classes Passed			Percent of English Classes Passed			Percent of Classes Passed		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Principal's Suspension	-0.25***	-0.10***	-0.10***	-0.26***	-0.11***	-0.11***	-0.25***	-0.10***	-0.10***
Superintendent's Suspension	-0.24***	-0.03*	-0.04**	-0.25***	-0.03*	-0.04**	-0.26***	-0.04***	-0.06***
Absences		-0.01***	-0.01***		-0.01***	-0.01***		-0.01***	-0.01***
Lateness		-0.01***	-0.01***		-0.01***	-0.01***		-0.01***	-0.01***
Female		0.04***	0.04***		0.07***	0.07***		0.05***	0.05***
Free/Reduced Price Lunch		0.03***	0.03***		0.03***	0.02***		0.03***	0.03***
Age in 9th Grade		0.00	0.00***		0.00	0.00		0.00	0.00
Race[1]									
Black		-0.04***	-0.04***		-0.03***	-0.03***		-0.04***	-0.04***
Hispanic		-0.06***	-0.06***		-0.05***	-0.05***		-0.04***	-0.04***
Asian		0.00	0.00		-0.02***	-0.02***		-0.02***	-0.02***
Other		-0.06***	-0.06***		-0.03*	-0.03*		-0.04***	-0.04***
Special Education		0.06***	0.06***		0.03***	0.03***		0.03***	0.03***
English Language Learner		0.01**	0.01**		-0.02***	-0.02***		0.00	0.00
8th Grade Absences		0.00**	-0.01***		0.00***	0.00***		0.00***	0.00***
8th Grade Latenesses		0.00***	-0.01***		0.00***	0.00***		0.00***	0.00***
Middle School Superintendent's Suspension		-0.02*	-0.02*		-0.04***	-0.04***		-0.03***	-0.03***
Middle School Principal's Suspension		-0.06***	-0.06***		-0.06***	-0.06***		-0.05***	-0.05***
8th Grade ELA Exam[2]		0.01*	0.01**		0.03***	0.03***		0.02***	0.02***
8th Grade Math Exam[2]		0.08***	0.09***		0.03***	0.03***		0.04***	0.04***
School Characteristics									
Percent Female[3]			0.01*			-0.00			0.00
Percent Free/Reduced Price Lunch[3]			0.00			0.00			0.00
Age in 9th Grade			0.02			0.01			0.03
Percent Black/Hispanic[3]			0.01*			0.01			0.01
Percent Special Education[3]			0.01			0.02***			0.02**
Percent English Language Learner[3]			0.01			0.01			0.00
Percent Middle School Suspension[3]			-0.02*			-0.02*			-0.02*
Mean 8th Grade ELA Score			0.00			0.02			0.02
Mean 8th Grade Math Score			-0.01			-0.02			-0.02
10th Grade Superintendent's Suspension * Mean Middle School Suspension			0.07**			0.07***			0.07***
10th Grade Principal's Suspension * Mean Middle School Suspension			0.00			0.03			0.02*
Constant	0.77***	0.95***	0.93***	0.80***	1.00***	0.98***	0.80***	0.96***	0.95***
Variance Components									
School	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Residual	0.13	0.09	0.09	0.11	0.08	0.08	0.08	0.04	0.04

*p<.05; **p<.01; ***p<.001

All school-level variables grand-mean centered

[1]Compared to white

[2]Measure is z-scored

[3]Unit is 10 percentage points

Tenth Grade Attendance. The results for tenth grade attendance outcomes and suspension in the first half of tenth grade are presented in Table 21. As was the case with credit accumulation, these results are very similar to those from the ninth grade models. In the adjusted relationships, Principal's suspension is related to eight additional latenesses and 15 additional absences in the tenth grade year. Furthermore, Superintendent's suspension is linked to four additional latenesses and 22 additional absences. However, adjusting these relationships for student characteristics halves the relationship between Principal's suspension and latenesses and absences, as well as the relationship between Superintendent's suspension and absences, and explains the relationship between Superintendent's suspension and latenesses.

My third models reveal that the link between Principal's suspension and latenesses depends upon the percentage of students who earned middle school suspensions, with above average levels of middle school suspension associated with more latenesses. Consistent with the ninth grade results, the relationship between Superintendent's suspension and absences also depends on aggregate levels of middle school suspension, with a ten percentage point increase in the percent of students receiving middle school suspension being related to an additional week of absences for students at schools with above-average levels of middle school suspendees.

Table 21. The Relationship between Suspension and Absence and Lateness in 10th Grade

	Days Late			Days Absent		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Principal's Suspension	8.13***	4.81***	5.05***	14.1***	7.20***	7.13***
Superintendent's Suspension	4.21***	0.78	0.47	21.7***	13.5***	14.3***
Absences		0.00	0.00		--	--
Lateness		--	--		0.10***	0.10***
Female		-0.08	-0.08		0.61***	0.59***
Free/Reduced Price Lunch		-0.21	-0.21		-2.73***	-2.75***
Age in 9th Grade		-0.32*	-0.31*		0.06	0.06
Race[1]						
Black		2.59***	2.57***		0.59	0.51
Hispanic		1.81***	1.80***		2.83***	2.75***
Asian		-0.52*	-0.54*		1.55***	1.60***
Other		-1.04	-1.05		0.80	0.74
Special Education		-0.50*	-0.52*		1.44***	1.35***
English Language Learner		-1.40***	-1.41***		-2.16***	-2.06***
8th Grade Absences		0.02**	0.02**		0.71***	0.72***
8th Grade Latenesses		0.18***	0.18***		0.11***	0.12***
Middle School Superintendent's Suspension		0.68	0.67		4.21***	4.11***
Middle School Principal's Suspension		1.66***	1.68***		4.42***	4.44***
8th Grade ELA Exam[2]		-0.65***	-0.66***		-0.79***	-0.76***
8th Grade Math Exam[2]		-0.80***	-0.81***		-1.31***	-1.30***
School Characteristics						
Percent Female[3]			0.45			0.50*
Percent Free/Reduced Price Lunch[3]			0.30			-0.22
Age in 9th Grade			-5.74			1.71
Percent Black/Hispanic[3]			1.27*			0.38
Percent Special Education[3]			1.81**			1.08**
Percent English Language Learner[3]			0.64			-0.25
Percent Middle School Suspension[3]			-0.30			-1.56**
Mean 8th Grade ELA Score			0.81			-2.07
Mean 8th Grade Math Score			1.60			-4.68***
10th Grade Superintendent's Suspension * Mean Middle School Suspension			1.87			-4.68***
10th Grade Principal's Suspension * Mean Middle School Suspension			-2.61**			1.11
Constant	15.8***	15.4***	14.0***	17.9***	3.33	2.69
Variance Components						
School	203.2	206.3	199.5	92.4	36.7	29.6
Residual	264.0	243.6	245.3	508.9	388.6	390.2

*p<.05; **p<.01; ***p<.001

All school-level variables grand-mean centered

[1]Compared to white

[2]Measure is z-scored

[3]Unit is 10 percentage points

Within-Student Estimates of Credit Accumulation and Attendance

Even though I include multiple controls in the two-level models estimating the relationship between suspension and credit accumulation and attendance, it is likely that my models still suffer from endogeneity problems, as suspended students likely differ from their non-suspended peers in ways that are difficult to measure. Thus, it is possible that my estimates are biased. Therefore, I estimate the relationship between suspension and these outcomes in one additional way: a three-level model, in which students serve as their own counterfactual. In these models, all measurement-level variables are centered within students, and all student-level variables are centered within schools. Therefore, the suspension effects are within-student estimates of how outcomes differ during the semester of suspension compared to semesters in which no suspension was assigned. Each student has up to eight time points in the model, one for each of the first eight semesters of high school. In the final models, I include a measure for time, centered around the second semester of tenth grade, when suspensions are at their highest.

Percent of Credits Passed. Table 22 displays the relationship between suspension and the percent of credits passed for math, English, and all classes. In contrast to the previous analyses, this table highlights the fact that suspension is only associated with a very small decrease in percent of math, English, or all credits earned compared with semesters without suspension. In the first models, the unadjusted relationships are displayed. Here we see that a Principal's suspension is linked to a five percentage point reduction in the percent of math, English, or all credits earned, and a Superintendent's suspension is associated with a 10 percentage point decline. After controlling for attendance and adjusting for student characteristics, the decrease in credit accumulation shrinks to between three and four percent for Principal's suspension, and two and four percent for Superintendent's suspension. Unlike in the two-level models, high school suspension is more weakly related to credit accumulation than race, eighth grade math scores, and discipline history across all models. These results differ because I am estimating the within-student suspension differences: the suspension coefficients indicate how much worse, on average, students are expected to do the semester of suspension compared to other semesters. In contrast, the student characteristic coefficients explain the variation in credit accumulation between students.

In the third models, I include school characteristics, as well as an indicator for time and an interaction between time and suspension status. The time measure is a year and semester indicator centered at the second term of sophomore year, taking values -3 to eight. The suspension estimates remain stable from Models 2 to 3. The relationship between Principal's suspension and percent of credits earned is weakly dependent on time with all three outcomes; however the relationship between Superintendent's suspension and credit accumulation only depends on time in terms of math credits earned. Where time mediates the relationship between suspension and credit accumulation, it means that the relationship between suspension and credit accumulation differs depending on the year and term of the suspension. In all cases of time dependency, being suspended before the second semester of tenth grade is associated with greater reductions in passage rate. For example, being assigned a Principal's suspension in the first semester of ninth grade is associated with an additional 3.6 percentage points decrease in math credit accumulation and a Superintendent's suspension is related to an additional 5.1 percentage point decrease when compared to being assigned a suspension in the second semester of tenth grade.[6] This dependency, though weak, highlights the importance of success in ninth grade, as missteps are often related to more detrimental outcomes when they occur earlier in

[6] Since time is centered at the second semester of the sophomore year, there is not any additional decrease associated with suspension occurring in that semester. However, in the first semester of ninth grade, time equals negative three. This means that there is an additional deficit associated with suspension of (-3*.012) for Principal's suspension and (-3*.017) for Superintendent's suspension.

students' high school careers.

Table 22. The Relationship between Suspension and Credits Earned

	Percent of Math Classes Passed (n=401,621)			Percent of English Classes Passed (n=421,973)			Percent of Classes Passed (n=429,255)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Measurement Level									
Principal's Suspension	-0.05***	-0.03***	-0.03***	-0.06***	-0.04***	-0.04***	-0.05***	-0.04***	-0.04***
Superintendent's Suspension	-0.10***	-0.04***	-0.04***	-0.10***	-0.03***	-0.03***	-0.11***	-0.04***	-0.04***
Absences		-0.01***	-0.01***		-0.01***	-0.01***		-0.01***	-0.01***
Lateness		-0.01***	0.00***		0.00***	0.00***		0.00***	0.00***
Time[4]			0.00***			0.00***			0.00***
Student Characteristics									
Female		0.05***	0.05***		0.07***	0.07***		0.05***	0.05***
Free/Reduced Price Lunch		0.05***	0.05***		0.05***	0.05***		0.05***	0.05***
Age in 9th Grade		0.00	0.00		0.00	0.00		0.00	0.00
Race[1]									
Black		-0.05***	-0.05***		-0.04***	-0.04***		-0.05***	-0.05***
Hispanic		-0.07***	-0.07***		-0.07***	-0.07***		-0.06***	-0.06***
Asian		-0.01*	-0.01*		-0.03***	-0.03***		-0.02***	-0.02***
Other		-0.05***	-0.05***		-0.05***	-0.05***		-0.05***	-0.05***
Special Education		0.05***	0.05***		0.01**	0.01**		0.01***	0.01***
English Language Learner		0.03***	0.03***		-0.01*	-0.01*		0.02***	0.02***
8th Grade Absences		-0.01***	-0.01***		-0.01***	-0.01***		-0.01***	-0.01***
8th Grade Latenesses		0.00***	0.00***		0.00***	0.00***		0.00***	0.00***
Middle School Superintendent's Suspension		-0.08***	0.08***		-0.09***	-0.09***		-0.09***	-0.09***
Middle School Principal's Suspension		-0.10***	-0.10***		-0.10***	-0.10***		-0.09***	-0.09***
8th Grade ELA Exam[2]		0.01***	0.01***		0.03***	0.03***		0.03***	0.03***
8th Grade Math Exam[2]		0.09***	0.09***		0.04***	0.04***		0.05***	0.05***
School Characteristics									
Percent Female[3]			0.00			0.00			0.00
Percent Free/Reduced Price Lunch[3]			0.01**			0.01**			0.01**
Age in 9th Grade			-0.14***			-0.13***			-0.13***
Percent Black/Hispanic[3]			0.00			-0.01*			-0.01*
Percent Special Education[3]			0.01**			0.01*			0.01*
Percent English Language Learner[3]			0.02***			0.01***			0.01***
Percent Middle School Suspension[3]			-0.05***			-0.04***			-0.03***
Mean 8th Grade ELA Score			0.00			0.02			0.021
Mean 8th Grade Math Score			0.14***			0.09***			0.10***
Superintendent's Suspension *Time			0.02***			0.00			0.00
Principal's Suspension*Time			0.01***			0.01**			0.01***
Constant	0.74***	0.73***	2.74***	0.78***	0.77***	2.69***	0.769***	0.76***	2.65***
Variance Components									
School	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Student	0.06	0.05	0.05	0.06	0.05	0.05	0.06	0.04	0.04
Residual	0.11	0.11	0.11	0.08	0.07	0.07	0.03	0.02	0.02

*p<.05; **p<.01; ***p<.001

All Measurement and Student Variables group-mean centered.

[1]Compared to white

[2]Measure is z-scored

[3]Unit is 10 percentage points

[4]Centered around the second semester of 10th grade

Attendance. Table 23 displays the relationship between suspension and attendance. There are stronger associations between suspension and absence – especially with regard to Superintendent's suspension – than there are between suspensions and credit accumulation. In the first models, we look at the within-student, unadjusted estimates of the relationship between suspension and absences and latenesses. In semesters where students received Principal's suspensions, they have approximately one additional lateness and one additional absence. In semesters where students were assigned Superintendent's suspensions, they have approximately one less lateness, but over a week of additional absences. In the second models, where I control for school-level covariates, the suspension estimates are substantively and statistically unchanged. Superintendent's suspension most strongly related absences, with middle school suspension record having the next strongest association.

In the third models, I add school-level metrics, as well as time and its interaction with suspension status. Substantively, time mediates the relationship between suspension status and attendance in the opposite way that it affects the relationship between suspension status and credit accumulation. In other words, suspension in semesters preceding the second semester of tenth grade is associated with fewer absences than suspension in eleventh and twelfth grade. For example, Principal's and Superintendent's suspension in the second semester of eleventh grade is associated with two roughly two additional absences compared to suspension in the second semester of tenth grade. Students suspended in the second semester of twelfth grade have roughly one day of additional lateness, compared with students suspended in the second semester of tenth grade. In this model, the mean entering math abilities of the students is strongly related to lateness and absences, as is the percentage of students who received a suspension in middle school.

Table 23. The Relationship between Suspension and Absence and Lateness (n=430,862)

	Days Late			Days Absent		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Measurement Level						
Principal's Suspension	1.29***	1.29***	1.24***	0.85***	0.99***	0.94***
Superintendent's Suspension	-1.11***	-1.02***	-0.94***	6.79***	6.80***	6.42***
Absences		-0.02***	-0.06***		--	--
Lateness		--	--		-0.02***	-0.06***
Student Characteristics						
Female		-0.26***	-0.26***		0.10	0.13
Free/Reduced Price Lunch		-0.08	-0.13*		-1.71***	-1.80***
Age in 9th Grade		-0.10*	-0.11*		0.16	0.17
Race[1]						
Black		1.44***	1.42***		0.49**	0.43**
Hispanic		0.97***	1.00***		1.69***	1.71***
Asian		-0.25*	-0.23*		0.69***	0.69***
Other		-0.03	-0.15		0.90*	0.69
Special Education		-0.15	-0.13		0.95***	0.98***
English Language Learner		-0.69***	-0.86***		-1.03***	-1.30***
8th Grade Absences		0.00	0.01		0.40***	0.40***
8th Grade Latenesses		0.09***	0.09***		0.07***	0.07***
Middle School Superintendent's Suspension		0.72***	0.81***		2.68***	2.87***
Middle School Principal's Suspension		1.03***	1.10***		2.86***	3.00***
8th Grade ELA Exam[2]		-0.29***	-0.300***		-0.43***	-0.46***
8th Grade Math Exam[2]		-0.39***	-0.408***		-0.87***	-0.91***
School Characteristics						
Percent Female[3]			-0.01***			0.05
Percent Free/Reduced Price Lunch[3]			0.18			-0.41***
Age in 9th Grade			-1.46			6.33***
Percent Black/Hispanic[3]			0.53**			0.45***
Percent Special Education[3]			0.28			0.19
Percent English Language Learner[3]			-0.03			-0.49**
Percent Middle School Suspension[3]			-0.28			2.95***
Mean 8th Grade ELA Score			1.27			0.96
Mean 8th Grade Math Score			-3.46*			-5.17***
Time[4]			0.71			0.94***
Superintendent's Suspension*Time			0.25**			0.80***
Principal's Suspension*Time			0.33***			1.09***
Constant	8.63***	8.87***	23.8	12.7	12.1	-82.9***
Variance Components						
School	44.3	45.2	38.9	36.1	41.7	12.6
Student	31.0	26.0	26.5	128.4	80.4	83.7
Residual	67.4	67.0	65.7	75.3	71.0	66.5

*p<.05; **p<.01; ***p<.001

All Measurement and Student Variables group-mean centered.

[1]Compared to white

[2]Measure is z-scored

Discussion

The relationship between suspension and short-term outcomes appears the same or more detrimental for Superintendent's suspension than for Principal's suspension, and gaps appear larger between students than within students. The negative associations also are stronger for absences than for credit accumulation.

The difference between the two- and three-level model findings likely suggests that there are differences between students that are unmeasured, and therefore unaccounted for, between students. For these reasons, the results presented in this chapter, though they may get closer to causal estimates than do simple regression estimates, should not be interpreted causally. This discrepancy also suggests that students' performance is relatively stable, and that receiving a suspension in any given semester does not provide that much of a shock to student outcomes. This stability is troubling given that students who earn these suspensions tend to consistently have worse outcomes than those students who are not suspended. These results, like those in the previous chapter, suggest that suspensions do not help put students back on track.

These differences also suggest that there may be some lingering effects of suspension beyond the semester in which students are suspended. These lasting results may be the reason why there are stronger relationships between suspension and attendance, for example, between students in the two-level models, than within students in the three-level models. It is possible that being assigned a suspension in the first semester negatively impacts your attachment to school and likelihood to attend for the remainder of the year.

The larger effect sizes for Superintendent's suspension is not surprising, given that students receiving these suspensions miss school for longer periods of time, nor is the fact that suspension is more strongly linked to attendance than to credit accumulation. Attendance is an objective metric. Students either attend school or not. Credit accumulation, on the other hand, is more subjective, as teachers and schools may have varying academic expectations depending on their student body. Furthermore, schools may have different policies for reintegrating students into the school environment after serving a suspension.

The fact that the percent of students who received suspensions in middle school mediates the relationship between suspension and credit accumulation and attendance reinforces this point. Students assigned suspensions at schools with higher levels of discipline problems are more likely to miss even more school, but have smaller decreases in credit accumulation associated with suspension. At schools with high levels of students with previous infractions, it may be easier for students to earn all of their credits. It is also possible that the pass rate across students is lower, making the differences between students who are and are not suspended smaller.

Chapter 6. Suspension and Long-term Outcomes

Whether students ultimately succeed in high school is of central concern for students' life trajectories. Therefore, for my final set of analyses, I examine the relationship between suspension and passing Regents exams, Regents exam scores, and four-, five- and six-year graduation. Regents exams are critical for students. Students in the 2005 cohort had to pass five tests with a score of 65 – one math, one English, one science, the Global History, and the U.S. History exam – in order to graduate with a Regents Diploma. Students who passed two exams with a 65, but the others with scores of 55, graduated with a Local Diploma. Graduating with a Regents versus a Local Diploma is essential for post-secondary and labor market success. The scores on Regents examinations are important, above and beyond whether a student reached a 65. These scores reflect students' academic achievement, and are a reflection of their mastery of k-12 academic content. Subsequently, students must earn at least a 75 on English and math Regents exams to pass out of remedial courses in the City University system. Additionally, students who earn above 90 on their exams graduate with a Regents Diploma with Honors.

Regression Estimates for Regents Performance

I begin by using ordinary least squares (OLS) regression to estimate the relationship between suspension and Regents performance, as regression is a basic way of obtaining estimates for the average treatment effect of suspension on Regents performance, and regression is commonly used in the suspension literature to estimate the relationships between suspension and student performance. Rather than using logistic regression with the dichotomous outcome of passing the Regents or not, I use ordinary least squares (OLS) regression to estimate linear probability models (LPMs). These are methodologically appropriate because I want my results to be comparable across techniques used in this chapter. My final method is propensity score matching. Since odds ratios are not collapsible, I cannot use logistic regression with matching and must instead estimate risk differences.

I define the suspension “treatment” to be any suspension in the first three semesters of high school. For the outcome, I include the highest test score on each respective exam that was taken after the first semester of tenth grade. Within each analysis, students who took the relevant exam in the second semester of tenth grade or later are included, yielding varying sample sizes.

Because I am ultimately estimating these relationships in a causal framework using the quasi-experimental method of propensity score matching, throughout this chapter I limit the included covariates to those that can be considered pre-treatment. This way, my results are comparable across techniques. Using the pretreatment variables, I try to make suspended students and non-suspended students as similar as possible. Subsequently, any observed differences in outcomes are theoretically attributable to the treatment, rather than to student characteristics. Though set up in a causal framework, these results should still be seen as associations because there are unmeasured student characteristics that are not held constant in these models.

Table 24 displays the relationship between suspension and whether students passed each Regents exam with a score of 55. In the first models, I estimate the unadjusted relationship between suspension and the probability of passing the exam. These results suggest that suspended students are 25 percentage points less likely to pass the English or U.S. History Regents, 24 percentage points less likely to pass the math Regents, 23 percentage points less likely to pass a science Regents, and 29 percentage points less likely to pass the Global History Regents.

After controlling for pre-treatment covariates in the second models, these estimates shrink by over 50 percent. Now, suspension is associated with a 13 percentage point decrease in the likelihood of passing the English Regents, a 12 percentage point decrease in the likelihood of passing a math or the U.S. History Regents, an 11 percentage point reduction in the likelihood of passing the Science Regents, and a 14 percentage point decline in the probability of passing the Global History Regents.

Table 24. Relationships between Suspension and Likelihood of Passing Regents Exams

	Pass ELA Regents		Pass Math Regents		Pass Science Regents		Pass Global History Regents		Pass US History Regents	
	(n=63,329)		(n=42,478)		(n=37,084)		(n=59,890)		(n=55,928)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.25***	-0.13***	-0.24***	-0.12***	-0.23***	-0.11***	-0.29***	-0.14***	-0.25***	-0.12***
Female		0.05***		0.13**		-0.02***		-0.01		0.01**
Black		-0.05***		-0.07***		-0.07***		-0.09***		-0.05***
Hispanic		-0.05***		-0.07***		-0.07***		-0.11***		-0.07***
Asian		-0.03***		-0.01		0.01		-0.02**		-0.03***
Other		-0.05**		-0.07**		-0.04		-0.05*		-0.03
Special Education		-0.25***		-0.23***		-0.19***		-0.18***		-0.20***
English Language Learner		-0.19***		0.00***		-0.03**		-0.03***		-0.04***
Free/Reduced Price Lunch		0.03***		0.05***		0.04***		0.02***		0.02***
Age in 9th Grade		-0.05***		-0.02***		-0.02***		-0.03***		-0.03***
8th Grade ELA Score		0.06***		0.03***		0.08***		0.09***		0.07***
8th Grade Math Score		0.08***		0.15***		0.11***		0.09***		0.09***
Middle School Principal's Suspension		-0.06***		-0.05***		-0.05***		-0.05***		-0.05***
Middle School Superintendent's Suspension		-0.06***		-0.06***		-0.07***		-0.08***		-0.08***
8th Grade Absences		0.00***		0.00***		0.00***		0.00***		0.00***
8th Grade Latenesses		-0.00***		0.00***		0.00***		0.00***		0.00***
Constant	0.77***	1.52***	0.68***	1.13***	0.61***	1.04***	0.66***	1.25***	0.75***	1.30***

*p<.05; **p<.01; ***p<.001

Table 25 presents the associations between suspension and Regents Exam scores. Again, in the first models, I explore the unadjusted relationship between suspension and test scores, and find that suspension is associated with a 15-19 point decline in exam performance. After holding constant pre-treatment student characteristics, these estimates are also halved, so that suspension is associated with approximately an eight point decrease on each type of exam.

Table 25. Relationship between Suspension and Regents Exam Scores[1]

	English (n=63,329)		Math (n=42,478)		Science (n=37,084)		Global History (n=59,890)		US History (n=55,926)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
9th/10th Grade Suspension[2]	-17.9***	-8.41***	-14.9***	-7.51***	-16.1***	-7.69***	-18.1***	-8.25***	-19.2***	-8.62***
Female		3.79***		1.60***		0.97***		0.35		1.13***
Black		-4.42***		-4.06***		-5.92***		-5.83***		-4.91***
Hispanic		-4.90***		-5.34***		-5.97***		-6.46***		-6.19***
Asian		-2.74***		1.07*		-1.07		-1.89***		-2.61***
Other		-7.13***		-7.27***		-8.02***		-5.86***		-3.54*
Special Education		-12.0***		-7.84***		-8.36***		-8.82***		-9.43***
English Language Learner		-7.21***		1.45***		0.73		-1.64***		-0.69
Free/Reduced Price Lunch		2.49***		4.14***		3.02***		2.10***		1.81***
Age in 9th Grade		-2.01***		-0.12		-0.37		-1.21***		-1.40***
8th Grade ELA Score		5.04***		2.25***		4.09***		5.55***		5.43***
8th Grade Math Score		5.38***		7.11***		6.06***		5.12***		6.20***
Middle School Principal's Suspension		-4.15***		-4.11***		-4.62***		-4.32***		-4.72***
Middle School Superintendent's Suspension		-5.37***		-4.14***		-3.41***		-5.67***		-6.79***
8th Grade Absences		-0.31***		-0.36***		-0.35***		-0.34***		-0.38***
8th Grade Latenesses		-0.08***		-0.09***		-0.10***		-0.09***		-0.13***
Constant	67.7***	103.5***	63.0***	74.6***	87.2***	75.9***	63.2***	90.5***	69.4***	98.6***

*p<.05; **p<.001; ***p<.001

[1]Performance on exams taken during or after the second semester of 10th grade.

[2]Earned at least one Principal's or Superintendent's Suspension during the 9th grade or the first half of 10th grade.

Fixed Effects Estimates for Regents Performance

The regression estimates presented in Tables 24 and 25 could be improved upon by taking account the nested nature of students within schools. Because I am interested in estimating the relationship between suspension and Regents outcomes and am not focused on explaining school-level factors also related to Regents exam performance, I employ school-fixed effects. This approach allows me to estimate the within-school average treatment effect adjusted for all unmeasured differences between schools. This means that I am comparing suspended students to non-suspended students in the same school. Table 26 presents the associations between suspension and likelihood of passing the Regents exam. The unadjusted school-fixed effects results are smaller than are those from the regular regression analyses, with suspension associated with a 20 percentage point decrease in the probability of passing the English or a math

Regents exams, an 18 percentage point decline in the likelihood of passing a science or the U.S. History Regents, and a 21 percentage point decline in probability of passing the Global History Regents. After adjusting for student-level characteristics in the second models, suspended students are ten percentage points less likely to pass a science or the U.S. History exam, 11 percentage points less likely to pass a math Regents, and twelve percentage points less likely to pass the English or Global History exams.

Table 26. Relationships between Suspension and Likelihood of Passing Regents Exams

	English (n=63,329)		Math (n=42,478)		Science (n=37,084)		Global History (n=59,890)		US History (n=55,928)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.20***	-0.12***	-0.20***	-0.11***	-0.18***	-0.10***	-0.21***	-0.12***	-0.18***	-0.10***
Female		0.05***		0.01		-0.03***		-0.01***		0.00
Black		-0.02***		-0.03**		-0.05***		-0.06***		-0.02**
Hispanic		-0.03***		-0.04***		-0.05***		-0.07***		-0.04***
Asian		-0.02**		0.02*		0.02		0.00		-0.01*
Other		-0.03		-0.04		-0.02		0.02		0.01
Special Education		-0.23***		-0.23***		-0.18***		-0.18***		-0.19***
English Language Learner		-0.18***		0.00		-0.05***		-0.04***		-0.05***
Free/Reduced Price Lunch		0.03***		0.06***		-0.05***		0.04***		0.03***
Age in 9th Grade		-0.05***		-0.02***		-0.03***		-0.03***		-0.03***
8th Grade ELA Score		0.06***		0.03***		0.08***		0.09***		0.07***
8th Grade Math Score		0.08***		0.14***		0.11***		0.09***		0.09***
Middle School Principal's Suspension		-0.06***		-0.06***		-0.06***		-0.06***		-0.06***
Middle School Superintendent's Suspension		-0.06***		-0.06**		-0.06**		-0.07***		-0.08***
8th Grade Absences		0.00***		0.00***		0.00***		0.00***		0.00***
8th Grade Latenesses		0.00***		0.00***		0.00***		0.00***		0.00***
Constant	0.76***	1.49***	0.67***	1.15***	0.60***	1.12***	0.65***	1.21***	0.75***	1.27***

*p<.05; **p<.01; ***p<.001

The associations between suspension and Regents exam scores calculated using school-fixed effects are displayed in Table 27. As with the likelihood of passing Regents exams, these unadjusted estimates are lower than they were when calculated using regression because I have already limited the comparison to students within the same school, thereby making a more plausible comparison between students who are and are not suspended. In these first models, suspension is associated with a 12-14 point decline in Regents scores. Holding constant student characteristics in the second models explains much of the relationship between suspension and

Regents exam scores, with the score decreases now ranging from 7.5-8.5 points.

Table 27. Relationship between Suspension and Regents Exam Scores[1]

	English (n=63,329)		Math (n=42,478)		Science (n=37,084)		Global History (n=59,890)		US History (n=55,928)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
9th/10th Grade Suspension[2]	-14.2***	-8.28***	-12.7***	-7.44***	-13.0***	-7.53***	-13.9***	-7.98***	-14.0***	-8.02***
Female		3.67***		1.31***		0.58*		-0.08		0.68**
Black		-2.19***		-1.67***		-3.24***		-3.83***		-2.36***
Hispanic		-3.70***		-3.77***		-3.93***		-4.74***		-4.41***
Asian		-2.11***		2.15***		-0.11		-1.17**		-1.78***
Other		-5.30***		-5.18***		-5.30***		-3.87***		-0.84
Special Education		-11.8***		-7.89***		-8.43***		-9.11***		-9.60***
English Language Learner		-7.40***		0.94*		-0.89*		-2.93***		-1.67***
Free/Reduced Price Lunch		3.06***		4.75***		3.93***		3.07***		3.00***
Age in 9th Grade		-2.00***		-0.38		-0.80**		-1.31***		-1.48***
8th Grade ELA Score		4.65***		1.69***		3.25***		4.74***		4.87***
8th Grade Math Score		5.11***		6.65***		5.38***		4.57***		5.67***
Middle School Principal's Suspension		-4.50***		-4.46***		-5.11***		-4.65***		-5.00***
Middle School Superintendent's Suspension		-4.95***		-3.71***		-2.93*		-5.25***		-6.50***
8th Grade Absences		-0.28***		-0.33***		-0.32***		-0.31***		-0.34***
8th Grade Latenesses		-0.08***		-0.08***		-0.09***		-0.08***		-0.12***
Constant	67.4***	101.4***	62.8***	75.6***	59.0***	78.9***	62.9***	89.7***	69.1***	97.0***

*p<.05; **p<.001; ***p<.001

[1]Performance on exams taken during or after the second semester of 10th grade.

[2]Earned at least one Principal's or Superintendent's Suspension during the 9th grade or the first half of 10th grade.

Propensity Score Estimates for Regents Performance

To improve upon these fixed-effects estimates, I utilize one final identification strategy. Since suspended students are vastly different than their non-suspended peers, I utilize multilevel propensity score matching. This technique restricts the analysis to a comparison of suspended and non-suspended students who have analogous propensities toward suspension. With this reduced sample, it is more plausible that we meet the regression linearity assumption and that we have enough overlap between treatment and control groups. Importantly, these estimates no longer are of the average treatment effect; rather, they represent the average effect of the treatment on the treated. In other words, they quantify how differently suspended students would be expected to perform had they not been assigned suspensions.

Table 28 displays the propensity score estimates for the links between suspension and the probability of passing Regents exams. The first models present the matched estimates without covariate adjustment. These estimates are similar to those after holding constant student characteristics with both regression and the school-fixed effects models. More specifically, suspension is associated with a nine percentage point decrease in the probability of passing a science Regents, and an 11 percentage point decrease in the likelihood of passing the English, a math, the Global History, or the U.S. History Regents exams. These findings remain unchanged when we include covariate adjustments in the second models.

Table 28. Relationships between Suspension and Likelihood of Passing Regents Exams

	English (n=8222)		Math (n=7089)		Science (n=5472)		Global History (n=8247)		US History (n=6580)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.11***	-0.11***	-0.11***	-0.11***	-0.09***	-0.09***	-0.11***	-0.11***	-0.11***	-0.11***
Female		0.07***		0.03		-0.01		-0.02*		0.00
Black		-0.05*		-0.06*		-0.08*		-0.09**		-0.06*
Hispanic		-0.04		-0.06*		-0.08*		-0.13***		-0.07*
Asian		-0.01		0.01		-0.05		-0.03		-0.07
Other		-0.11		-0.12		-0.18		-0.09		-0.06
Special Education		-0.17***		-0.18***		-0.11***		-0.11***		-0.13***
English Language Learner		-0.13***		0.00		-0.01		0.01		0.00
Free/Reduced Price Lunch		0.03*		0.05***		0.06**		0.03**		0.03
Age in 9th Grade		-0.05***		-0.04*		-0.04**		-0.05**		-0.05**
8th Grade ELA Score		0.12***		0.04***		0.11***		0.13***		0.12***
8th Grade Math Score		0.10***		0.15***		0.11***		0.09***		0.10***
Middle School Principal's Suspension		-0.05**		-0.04		-0.03		-0.04*		-0.05
Middle School Superintendent's Suspension		-0.04		-0.01		-0.04		-0.08**		-0.05
8th Grade Absences		0.00***		0.00***		0.00***		0.00***		0.00***
8th Grade Latenesses		0.00***		0.00***		0.00***		0.00***		0.00*
Constant	0.63***	1.54***	0.54***	1.27***	0.47***	1.31***	0.476***	1.49***	0.609***	1.58***

*p<.05; **p<.01; ***p<.001

In Table 29, I share the estimates for the reduction in Regents scores. These results suggest that suspended students would be expected, on average, to score between seven and eight points higher on the Regents exam had they never been suspended. These results are substantively consistent with those presented with the previous methods. Although these differences may sound small, it is quite possible that seven or eight points could mean the difference between reaching a 65 or not, thereby changing the type of diploma the student earned or earning a diploma at all.

Table 29. Relationship between Suspension and Regents Exam Scores[1]

	English (n=8222)		Math (n=7089)		Science (n=5472)		Global History (n=8247)		US History (n=6580)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
9th/10th Grade Suspension[2]	-7.69***	-7.78***	-7.38***	-7.29***	-7.23***	-7.19***	-7.50***	-7.43***	-7.94***	-7.87***
Female		5.13***		2.17**		2.42*		0.70		0.89
Black		-5.24***		-2.74		-7.87***		-5.95***		-5.89**
Hispanic		-5.13***		-4.65**		-8.43**		-7.15***		-7.29***
Asian		-3.48		-0.34		-4.45		-3.76		-5.70*
Other		-15.7**		-11.9		-23.0*		-7.77		-9.68
Special Education		-11.1***		-6.78***		-6.73***		-8.17***		-8.26***
English Language Learner		-4.45**		-0.08		1.58		-0.44		2.26
Free/Reduced Price Lunch		4.22***		5.90***		5.75***		4.33***		3.25**
Age in 9th Grade		-2.35**		-1.19		-1.52		-1.99*		-3.07*
8th Grade ELA Score		7.38***		3.17***		4.98***		7.03***		7.42***
8th Grade Math Score		5.81***		6.32***		5.72**		4.80***		6.57***
Middle School Principal's Suspension		-3.90**		-3.73**		-4.90**		-4.30***		-2.95
Middle School Superintendent's Suspension		-4.79*		-2.42		-2.62		-6.07**		-5.52
8th Grade Absences		-0.35***		-0.40***		-0.35***		-0.34***		-0.41***
8th Grade Latenesses		-0.07***		-0.08***		-0.10**		-0.08***		-0.11***
Constant	57.5***	107.7***	55.5***	87.8***	50.5***	91.3***	52.6***	99.8***	58.2***	122.3***

*p<.05; **p<.001; ***p<.001

[1]Performance on exams taken during or after the second semester of 10th grade.

[2]Earned at least one Principal's or Superintendent's Suspension during the 9th grade or the first half of 10th grade.

Regression Estimates for Likelihood of Graduation

Perhaps the most important high school outcome is whether a student graduates. Consequently, I conduct analyses investigating the links between suspension and graduation using the same three methods used for Regents scores: regression, school-fixed effects, and propensity score matching. I utilize two treatments in these examinations: one that is any suspension in the first three semesters of high school, and one that is suspension in ninth grade. I include this second treatment since success in ninth grade is crucial for persistence through high school.

Regression estimates for the relationship between suspension in the first three semesters and four-, five- and six-year graduation status is presented in Table 30. Although the reduction in probability is approximately 30 percentage points when student characteristics are not held constant, these estimates are reduced by half after controlling for student-level measures. Suspended students are 18 percentage points less likely to graduate in four years and 17 percentage points less likely to graduate in five or six years. As shown in Table 31, the results for suspension in ninth grade are virtually the same.

Table 30. Relationships between Suspension and Likelihood of Graduation (n=70,130)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.34***	-0.18***	-0.31***	-0.17***	-0.31***	-0.17***
Female		0.07***		0.05***		0.05***
Black		-0.06***		-0.04***		-0.04***
Hispanic		-0.09***		-0.06***		-0.06***
Asian		-0.06***		-0.05***		-0.05***
Other		-0.06**		-0.05**		-0.05**
Special Education		0.03***		0.02**		0.01**
English Language Learner		-0.06***		-0.03***		-0.03***
Free/Reduced Price Lunch		0.05***		0.07***		0.07***
Age in 9th Grade		-0.02***		-0.02***		-0.03***
8th Grade ELA Score		0.05***		0.04***		0.03***
8th Grade Math Score		0.09***		0.07***		0.07***
Middle School Principal's Suspension		-0.10***		-0.10***		-0.10***
Middle School Superintendent's Suspension		-0.08***		-0.10***		-0.11***
8th Grade Absences		-0.01***		-0.01***		-0.01***
8th Grade Latenesses		0.00***		0.00***		0.00***
Constant	0.67***	1.09***	0.78***	1.23***	0.78***	1.25***

*p<.05; **p<.01; ***p<.001

Table 31. Relationships between Suspension and Likelihood of Graduation (n=70,130)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During 9th Grade	-0.34***	-0.18***	-0.32***	-0.16***	-0.32***	-0.16***
Female		0.08***		0.05***		0.05***
Black		-0.06***		-0.04***		-0.04***
Hispanic		-0.09***		-0.07***		-0.07***
Asian		-0.06***		-0.05***		-0.05***
Other		-0.06**		-0.05**		-0.05**
Special Education		0.03***		0.01**		0.01*
English Language Learner		-0.06***		-0.03***		-0.03***
Free/Reduced Price Lunch		0.05***		0.07***		0.07***
Age in 9th Grade		-0.02***		-0.02***		-0.03***
8th Grade ELA Score		0.05***		0.04***		0.04***
8th Grade Math Score		0.09***		0.07***		0.07***
Middle School Principal's Suspension		-0.10***		-0.10***		-0.10***
Middle School Superintendent's Suspension		-0.09***		-0.11***		-0.11***
8th Grade Absences		-0.01***		-0.01***		-0.03***
8th Grade Latenesses		0.00***		0.00***		-0.01***
Constant	0.66***	1.08***	0.77***	1.23***	0.77***	1.24***

*p<.05; **p<.01; ***p<.001

Fixed Effects Estimates for Likelihood of Graduation

Tables 32 and 33 display the school-fixed effects results for the relationship between suspension and graduation status. These results are substantively equivalent to those presented in the regression models once covariates are included in the model. We see that suspension in the first three semesters of high school is linked to an 18 percentage point decrease in the likelihood of four-year graduation and a 17 percentage point reduction in the probability of five- or six-year graduation. We also see that ninth grade suspension is related to a 17 percentage point reduction in the likelihood of four-year graduation, and a 16 percentage point decrease in the probability of five- or six-year graduation.

Table 32. Relationships between Suspension and Likelihood of Graduation (n=70,130)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.29***	-0.18***	-0.28***	-0.17***	-0.27***	-0.17***
Female		0.07***		0.05***		0.05***
Black		-0.04***		-0.02***		-0.02***
Hispanic		-0.09***		-0.06***		-0.06***
Asian		-0.04***		-0.03***		-0.03***
Other		-0.04		-0.04***		-0.03
Special Education		0.04***		0.02***		0.02
English Language Learner		-0.06***		-0.04***		-0.03***
Free/Reduced Price Lunch		0.06***		0.08***		0.08***
Age in 9th Grade		-0.02***		-0.02***		-0.03***
8th Grade ELA Score		0.04***		0.03***		0.03***
8th Grade Math Score		0.09***		0.07***		0.07***
Middle School Principal's Suspension		-0.10***		-0.10***		-0.10***
Middle School Superintendent's Suspension		-0.07***		-0.09***		-0.09***
8th Grade Absences		-0.01***		-0.01***		-0.01***
8th Grade Latenesses		0.00***		0.00***		0.00***
Constant	0.67***	1.06***	0.77***	1.21***	0.78***	1.22***

*p<.05; **p<.01; ***p<.001

Table 33. Relationships between 9th Grade Suspension and Likelihood of Graduation (n=70,130)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During 9th Grade	-0.29***	-0.17**	-0.27***	-0.16***	-0.27***	-0.16***
Female		0.07***		0.05***		0.05***
Black		-0.05***		-0.02***		-0.02***
Hispanic		-0.09***		-0.06***		-0.06***
Asian		-0.04***		-0.03***		-0.03***
Other		-0.03		-0.03*		-0.03
Special Education		0.03***		0.02***		0.02***
English Language Learner		-0.06***		-0.04***		-0.03***
Free/Reduced Price Lunch		0.06***		0.08***		0.08***
Age in 9th Grade		-0.02***		-0.02***		-0.03***
8th Grade ELA Score		0.04***		0.04***		0.03***
8th Grade Math Score		0.09***		0.07***		0.07***
Middle School Principal's Suspension		-0.11***		-0.11***		-0.11***
Middle School Superintendent's Suspension		0.08***		-0.10***		-0.10***
8th Grade Absences		-0.01***		-0.01***		-0.01***
8th Grade Latenesses		0.00***		-0.00***		0.00***
Constant	0.66***	1.05***	0.77***	1.20***	0.77***	1.22***

*p<.05; **p<.01; ***p<.001

Propensity Score Estimates for Likelihood of Graduation

Finally, I estimate these associations using propensity score matching, which provides the average effect of the treatment on the treated. These results, displayed in Tables 34 and 35, suggest that the average effects of the treatment on the treated are just one percentage point smaller than the average treatment effects estimated with school-fixed effects. In other words, students who were assigned suspension in the first three semesters of high school would be expected to have a probability of four-year graduation that is 17 percentage points lower than if they had not been suspended, and likelihoods of five- and six-year graduation that are 16 percentage points lower. Similarly, ninth grade suspension is related to a 16 percentage point decrease in the probability of four- and five-year graduation, and a 15 percentage point reduction in the likelihood of six-year graduation.

Table 34. Relationships between Suspension and Likelihood of Graduation (n=7,251)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension During First Three Semesters	-0.18***	-0.17***	-0.17***	-0.16***	-0.17***	-0.16***
Female		0.08***		0.07***		0.07***
Black		-0.04		-0.03		-0.03
Hispanic		-0.08**		-0.07**		-0.06**
Asian		-0.05		-0.03		-0.03
Other		-0.06		-0.08		-0.08
Special Education		0.05**		0.04*		0.04*
English Language Learner		0.00		0.01		0.02
Free/Reduced Price Lunch		0.07***		0.10***		0.10***
Age in 9th Grade		-0.03**		-0.04***		-0.05***
8th Grade ELA Score		0.07***		0.06***		0.06***
8th Grade Math Score		0.08***		0.07***		0.08***
Middle School Principal's Suspension		-0.09***		-0.10***		-0.10***
Middle School Superintendent's Suspension		-0.06*		-0.07**		-0.07**
8th Grade Absences		-0.01***		-0.01***		-0.01***
8th Grade Latenesses		0.00***		0.00***		0.00***
Constant	0.51***	1.19***	0.63***	1.45***	0.63***	1.47***

*p<.05; **p<.01; ***p<.001

Table 35. Relationships between 9th Grade Suspension and Likelihood of Graduation (n=4,180)

	Graduate in 4 Years		Graduate in 5 Years		Graduate in 6 Years	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Suspension during 9th Grade	-0.17***	-0.16***	-0.16***	-0.16***	-0.16***	-0.15***
Female		0.09***		0.07**		0.07***
Black		-0.05		-0.03		-0.03
Hispanic		-0.09**		-0.06*		-0.06*
Asian		-0.06		-0.02		-0.02
Other		-0.09		-0.11		-0.11
Special Education		0.06*		0.04		0.04
English Language Learner		0.00		0.00		0.01
Free/Reduced Price Lunch		0.06***		0.09***		0.09***
Age in 9th Grade		-0.04**		-0.05**		-0.05**
8th Grade ELA Score		0.07***		0.06***		0.06***
8th Grade Math Score		0.08***		0.07***		0.08***
Middle School Principal's Suspension		-0.08**		-0.07*		-0.07*
Middle School Superintendent's Suspension		-0.05*		-0.06		-0.06*
8th Grade Absences		-0.01***		-0.01***		-0.01***
8th Grade Latenesses		0.00***		0.00***		0.00***
Constant	0.49***	1.23***	0.62***	1.47***	0.62***	1.48***

*p<.05; **p<.01; ***p<.001

Discussion

Suspension is strongly related to end of high school outcomes. Students suspended in the first three semesters of high school are much less likely to pass their Regents exams, and have somewhat lower exam scores. Furthermore, students suspended in ninth grade or in the first three semesters of high school are considerably less likely to graduate high school in four-, five-, or six-years. These estimates are consistent across three estimation methods. In addition, the propensity score matching estimates are similar whether matching with replacement, matching without replacement, or matching without replacement and a common support restriction are utilized.

Though I utilize two quasi-experimental techniques, fixed effects and propensity score matching, causal interpretation of these results should be used with caution. School-fixed effects take into account school-level characteristics that differ across schools and may be related to suspension and end of high school outcomes, but they do not adjust for within-school unmeasured varying characteristics. These characteristics also could be related to both suspension and the included outcomes, such as student engagement, or non-cognitive traits such as grit and resilience. Therefore, my estimates still may suffer from endogeneity. Propensity score matching does not remove endogeneity issues, but it does further restrict my analysis to students who are extremely similar on measured characteristics, thereby providing me with better overlap between suspended and non-suspended students, and making it so that the linearity assumption of regression need only apply in a more limited range.

Another limitation of these models is the specification of the treatment and outcome variables. The suspension treatment may be vastly different depending on when students received the suspension: perhaps the treatment of suspension in ninth grade is very different than suspension in tenth grade. Similarly, I am unable to account for the different times that students took the exams using these modeling techniques. The relationship between suspension and Regents outcomes may vary based on when students took the exams. Not only am I unable to model this variation, but I also do not estimate how the relationship varies depending on how many times students took the test. I include students' highest test scores in my analyses. It is quite possible that suspended students may have to take the test multiple times before passing the exam. Therefore the relationship between suspension and exam performance would look different if I included an outcome of their first attempt versus their best attempt at passing the exam.

Failing to account for the timing of the tests also means that I am comparing exam results from different iterations of the same test, in the case of math[7], English, Global History, or U.S. History. Furthermore, I do not differentiate between the science subject in which students earned their highest score, as students are able to pass any science exam to meet the graduation requirements. As with the other measurement limitations, it is possible that the estimated relationships would differ if I redefined these sets of outcomes. Even with these methodological limitations, the estimates presented in this chapter should be considered strong evidence that suspension is negatively related to student outcomes. Although I am unable to assert that suspension causes these outcomes, the consistency of the negative relationships leads me to conclude that suspensions are unable to, on average, improve students' academic trajectories.

[7] There may be a few students who took the Integrated Algebra rather than the Math A exam. This would mean that they took and passed their math exam in or after June 2008.

Chapter 7. Conclusion

School discipline continues to be a topic at the forefront of education policy debates, as discipline gaps have persisted since being brought to the nation's attention in 1975. Indeed, in 2011, Secretary of Education Arne Duncan highlighted school discipline as a matter of federal education policy through his Supportive School Discipline Initiative. This policy creates a partnership between the Departments of Education and Justice in an effort to cultivate the use of practices that reduce the need for exclusionary discipline within schools. One of the main thrusts of this initiative is to increase research on school discipline to help practitioners and policy makers make informed decisions about how to eliminate discipline gaps and better serve all students.

This study extends the literature on school discipline by employing rigorous methods to better quantify relationships between student and school characteristics and suspension and estimate the negative associations between school suspensions and student outcomes. I began by exploring average differences between New York City high school students who did and did not earn suspensions at any point. My results are consistent with extant research: I find that a greater percentage of suspended students were African-American, had IEPs, and were male. These suspended students also had lower initial academic performance, on average, than their non-suspended peers, and once in high school, had weaker school attendance, gained fewer credits, and earned lower grades each semester.

I then estimated these relationships in multivariate frameworks, starting with the relationship between student and school characteristics and the odds of suspension. I found that student socio-demographic and academic characteristics continue to be strong predictors of suspension. These student-level attributes are much stronger predictors of suspension than are school characteristics. In particular, race, gender, special education status, and previous disciplinary infractions are strongly related to the odds of being assigned suspension in high school. Altogether, these disquieting results suggest even greater incidences of exclusionary disciplinary practices among already disadvantaged student populations.

Both sets of analyses exploring the characteristics of suspended students draw attention to the differences that exist between the population of high school students who are and are not suspended and provide evidence that sophisticated methods are required to examine the links between suspensions and student outcomes. Researchers must account for the fact that by the time students enter high school, suspended students are vastly different than their peers who never earn suspensions. Failing to account for these differences would produce estimates of suspension effects that in fact spuriously reflect other differences between students who are and are not suspended. For this reason, I first utilized two-level models to estimate the relationship between suspension and student outcome while controlling for all available student and school covariates. I then employed a methodology that allowed me to estimate these associations within students. Using large-scale data and student fixed-effects permitted me to examine average differences between an individual student's performance the semester(s) he/she was and was not suspended. Doing so allowed me to eliminate all unmeasured differences between students and to make more valid inferences about the negative associations between suspension and student outcomes.

The relationships between suspension and credit accumulation are negative, but not as large as I assumed they would be. I hypothesize that the magnitude of these relationships are due to the subjectivity of granting students' passing grades, and of varying systems of reintegrating students back into home school environments. Given that I control for attendance, it is also possible that students who attend their alternative instruction complete enough work to maintain a rate of credit accumulation that is only slightly lower than the rate in non-suspended semesters. The attendance gaps are more troubling than the credit accumulation gaps. Attendance matters not only for student learning, but also for students' sense of belonging to the academic community.

I concluded my analyses with an investigation of the relationship between suspension and long-term outcomes. These estimates underscore many of the concerns addressed in extant research. It appears that students who are suspended are less likely to pass their Regents exams, are more likely to earn lower exam scores, and are less likely to graduate in four, five, or six years than their non-suspended peers. These negative relationships, in conjunction with the overrepresentation of marginalized socio-demographic groups, make the use of suspensions problematic. Although some of the examined relationships were not as severe as I anticipated, the evidence still suggests that suspended students do not achieve the same level of outcomes as those students who are not suspended.

Changes to New York City Discipline

These findings raise questions about NYCDOE's suspension policy, in particular, given that I focus on the 2005 cohort of NYCDOE high school students. Importantly, since 2010, the number of suspensions in New York City has fallen. In the 2012-2013 school year, 53,465 suspensions were issued, compared to 73,441 in the 2010-2011 school year (Decker, 2013). Approximately 11,600 of these were Superintendent's suspensions, compared to 15,055 in 2010-2011 (Decker, 2013). Although the number of suspensions has declined, discipline gaps persist between black and Hispanic students, who were assigned almost 90 percent of suspensions in 2011-2012, and white students, who represented approximately seven percent of suspensions (Decker, 2013). Students with IEPs also continue to be suspended at higher rates than their general education peers (Decker, 2013). Furthermore, these suspension numbers remain higher than they were before the mayor's office gained control of the schools: in the 2000-2001 school year, students served 48,471 suspensions, and in the 2001-2002 school year, 28,449[8] suspensions were assigned (NYCLU, 2011).

The recent decline in suspensions is partially attributed to the 2010 passage of the Student Safety Act. This law mandates that the NYCDOE and the NYPD report arrest, suspension, and expulsion data four times per year, in the aggregate and disaggregated by race, gender, age, and special education and language status (Walz, 2010). That same year, the NYCDOE decreased the number of zero tolerance infractions in the discipline code from 29 to 21, the first decrease since 2001, when the code included 14 zero tolerance infractions (NYCLU, 2011). The first discipline data release brought the number of exclusionary discipline incidents to the public attention, subsequently increasing pressure on the NYCDOE to revise its discipline code (Cromidas, 2012b). In response, the 2012-2013 Discipline Code emphasized that teachers should immediately address students' behavioral needs and advise counseling before resorting to assigning consequences for misbehavior (Baker, 2012; Cromidas, 2012a). In addition, the NYCDOE eliminated suspension for low-level offences, such as being late or bringing cell phones to school, and restricted the number of days to five that kindergarten through third graders could be suspended for midlevel transgressions, such as engaging in minor physical conflicts or drawing graffiti on school property (Baker, 2012). The discipline code was revised again in 2013-2014. It includes 62 infractions, 16 of which are zero tolerance infractions and 42 of which could result in suspension (NYCLU, 2013).

The NYCDOE has also increased its efforts to reduce discriminatory incidents, including harassment, intimidation, and bullying, through the Respect for All (RFA) program, which launched in the 2007-2008 school year. In that year, teams of staff members from schools

[8] This low number may be a result of a record-keeping error (NYCLU, 2011).

serving grades 6-12 attended a two-day RFA training, and the RFA website provided staff with instructional and professional development materials to help them better teach students about promoting respect and combating discrimination (NYCDOE, 2014b). The following year, all schools were required to identify a RFA liaison and to articulate an RFA plan aimed at reducing bullying and promoting respect for all community members. This same year, RFA instructional materials were linked to the NYCDOE Student Bill of Rights and Discipline Code. In 2009-2010, the NYCDOE included RFA measures in its School Quality Review, thereby formalizing RFA as part of the school evaluation system, and it extended training to k-5 schools. Since then, the NYCDOE has made annual improvements to the RFA program, including releasing new professional development opportunities, creating the *Respect for All Award* to honor schools that excel in their RFA implementation, expanding instructional material to address cyberbullying, and expanding the RFA library.

Promising Alternatives to Suspension: Guiding Principles

My findings underscore the need to better serve suspended students. Extant literature provides some evidence about how to improve discipline policy to meet the needs of both general and special populations. In response to the growing body of evidence that suggests racial discrimination in the use of exclusionary discipline policy, the federal Department of Education released a 2014 guidance letter on how schools and districts can meet their federal non-discrimination obligation in school discipline policy, as well as *Guiding Principles: A Resource Guide for Improving School Climate and Discipline* (USDOE, 2014). In this report, the USDOE argues that establishing a safe school environment is a prerequisite for creating a great school that facilitates students' academic success, and it condemns the current use of exclusionary discipline in attempting to create safe learning environments. Additionally, the USDOE outlines three key principles that are necessary for fairly fostering a healthy school climate, as well as action steps associated with realizing these tenets. The first principle states that "schools that foster positive school climates can help to engage all students in learning by preventing problem behaviors and intervening effectively to support struggling and at-risk students" (USDOE, 2014). To meet this principle, the USDOE suggests that schools write school climate goals and then use evidence-based approaches to meet these goals. In particular, the USDOE indicates that tiered support strategies are a promising means of meeting students' socio-emotional and behavioral needs and of creating safe school environments.

One such initiative that is largely seen as successful is Positive Behavioral Interventions and Supports (PBIS). This school-wide approach, which is rooted in frameworks for public health interventions and is named in the USDOE *Guidlines*, utilizes a three-tiered approach to behavioral intervention (Horner & Sugai, 2006; Sugai & Horner, 2005; Walker et al., 1996). Implementation of PBIS has been associated with decreased office referrals, and reductions in suspension and expulsion (Bradshaw, Mitchell, & Leaf, 2010; Muscott, Mann, & LeBrun, 2008; Simonson, et al., 2012). This system places the onus for managing student behavior on adults. Staff members are required to change the way they develop and communicate expectations, respond to student behavior, track and measure student behavior, and modify expectations, responses, and supports (Bradshaw, Mitchell, & Leaf, 2010; Netzel & Eber, 2003).

Schools and districts begin implementing PBIS on the universal level: they clearly develop positive behavioral expectations and have teachers conduct explicit lessons on how to use appropriate behavior. In this system, both good and poor behaviors are acknowledged: behaviors in line with expectations are rewarded, and non-compliant actions are handled using a continuum of consequences. Groups of students who need further behavioral attention are included in universal interventions and secondary interventions, which are actions targeted to help specific sub-groups meet school-wide behavioral expectations. Finally, students' needs may be addressed on the individual or intensive level, if behavioral needs are great enough (Bradshaw, Mitchell, & Leaf, 2010; Muscott, et al., 2008; Simonson, et al., 2012).

Using a framework consistent with PBIS is commonly mandated for students with special needs. The amended Individuals with Disabilities Education Act of 1997 (IDEA) requires that students with Individualized Education Programs (IEPs) whose behavior may interfere with learning are supposed to be individually assessed using a Functional Behavior Assessment (FBA). Furthermore, IDEA endorses PBIS as a means of effectively addressing the behavioral needs of students with IEPs. This process is also mandated by New York State Education Law (Advocates for Children, 2013). If the results of the FBA suggest that the student's behavior hinders learning, schools are required to create a Behavior Intervention Plan to address identified behavioral challenges. However, the degree to which schools are following these legal requirements is subject to question, and opponents of exclusionary discipline policy cite disproportional rates of suspension for special education students in arguments for increased use of individual behavior plans and decreased use of suspension (Advocates for Children, 2013).

In addition to encouraging schools to use tiered supports, this principle necessitates that SROs roles are clearly defined and are focused on students' welfare, and that continuing professional development be provided to all staff members with regard to encouraging positive behavior and responding appropriately to behavioral infractions that occur. The USDOE also advocates instruction in non-cognitive skills to help students learn socio-emotional competencies such as problem-solving and resiliency, and promotes community-based partnerships with organizations that can enhance schools' mental health support systems and social and emotional learning programs.

Social and emotional learning (SEL) can help students foster self-discipline and positive mental health, thereby preventing disciplinary infractions that may lead to suspension. Through SEL programs, students learn self-awareness, self-management, social awareness, relationship skills, and responsible decision making (CASEL, 2014). Fostering these competencies in children more inclined toward behavioral infractions may be particularly helpful in preventing disciplinary incidents, as children who struggle with aggression have trouble controlling their emotions (Eisenberg, et al., 2006; Hoffman, 2000). Moreover, these programs may increase students' ability to build and maintain relationships with peers and school staff members, thereby making students feel safer and more comfortable and emotionally attached to school (for review, see Osher, et al., 2010). Although SEL programs vary across schools, commonly shared features include curriculum lessons on social skills and social and emotional development, with opportunities to apply the competencies learned in the curriculum in class, at home, or in service learning trips (Osher, et al., 2010).

A meta-analysis of 213 universal, school-based SEL programs concludes that these school-wide interventions are positively related to academic performance and social and emotional skills (Durlak, et al., 2011). Furthermore, participation in universal SEL programming is related to decreases in problematic behavior, including disruptive class conduct, aggression, noncompliance, and bullying. Programs that are especially effective utilize SAFE implementation, which means they "use sequenced step-by-step training approach, use active forms of learning, focus sufficient time on skill development, and have explicit learning goals" (Durlak, et al., 2011, p. 408). These findings are consistent with an earlier meta-analysis of 165 studies of school-based prevention programs for problematic student behavior, which found that cognitive-behavioral and behavioral instructional programs are related to decreases in disciplinary infractions (Wilson, Gottfredson, & Najaka, 2001). Programs that explicitly taught behaviors, and then asked students to apply what they learned and that provided students with feedback, were particularly successful.

The second *Guiding Principle* states that "schools that have discipline policies or codes of conduct with clear, appropriate, and consistently applied expectations and consequences will help students improve behavior, increase engagement, and boost achievement" (USDOE, 2014). The USDOE encourages schools to set rules and expectations that are developmentally appropriate, suitable for students with special needs, and aligned with the school-wide systems put in place in conjunction with the first principle. School administrators should engage all community stakeholders in developing the discipline code, including families, students and school staff (USDOE, 2014). Exclusionary discipline should be reserved for exceptional cases, and, when used, schools must provide alternate education for removed students.

In lieu of purely punitive systems, schools should create consequences that teach students to learn from their past behavior. One approach that does so is the restorative justice approach to school discipline, which is a framework that relies upon social engagement as motivation for adhering to school expectations for conduct (Morrison & Vaandering, 2012). Restorative justice models focus on supporting students who transgress upon the community's code of conduct as they transfer their negative emotions, such as anger and aggression, to positive emotions, such as empathy and interest in other perspectives. This transformation occurs by involving those most affected in a process that allows the community members to repair the damage done in the incident, to decrease the risk of similar incidents happening in the future, to work toward healing those who were hurt by the infraction, and to empower the community to participate in the conflict resolution (Morrison & Vaandering, 2012; Pavelka, 2013). Common restorative justice programs include peacemaking circles, peer mediation, peer/accountability boards, and community conferences (Morrison & Vaandering, 2012; Pavelka, 2013).

Finally, schools must work to ensure discipline is equitably and fairly applied by building staff capacity and by utilizing data to implement and alter discipline policy. Staff professional development should include instruction on increasing teachers' abilities to effectively manage their classrooms, and data collection should include individual teachers' application of discipline, as well as school-wide trends. Collecting teacher-level data through observation, in addition to administrative records, may better inform teachers' professional development needs. Utilizing observation tools that focus on teacher-student interactions, such as the Classroom Assessment Scoring System (CLASS) can help reveal areas of strength and growth with regard to teachers' ability to create positive learning environments and appropriately interact with all of their students (Allen, et al., 2013; La Paro, Pianta, & Stuhlman, 2004; Pianta & Hamre, 2009). School-level analysis is critical, as consistently applied expectations are positively related to students' sense of safety (Gregory, et al., 2010), and monitoring these data can help indicate whether student groups are being disproportionately targeted.

Directions for Future Research

My study raises numerous directions for future research, as many of the questions central to suspension remain unanswered. Even though I attempt to use methods that provide more plausible counterfactuals for suspended students, there are likely issues of endogeneity with all presented models. To make these comparisons even more convincing, it would be helpful to have information on the types of infractions that were related to suspension. These data would allow for adjustment based on the level of past infractions, as well as enable explorations of uneven assignment based on infraction type across socio-demographic groups. It would also be interesting to obtain behavioral data on students who were not suspended, as this would provide helpful context when making judgments about the unequal distribution of exclusionary discipline. Qualitative research may be most appropriate for exploring undocumented student behavior, as this removes the role of school staff members in interpreting student behavior as worthy of documentation or not.

These results also suggest that more precise estimates between student and school characteristics and between suspension and student outcomes are needed for students in elementary and middle school. It is possible that the negative relationships would be stronger in the earlier grades, especially in middle school. Understanding when these effects are most negative would help educators and policy makers make more informed decisions about high-leverage points for adjusting school discipline policy.

In addition, this study raises questions about the mechanisms behind the observed negative relationships. For example, why do students have a lower credit passing rate in the semester of suspension than in other semesters? Why does this depend on the discipline history of their peers? Why are the effects starker for attendance than for credit accumulation? I suggest that these differences are a result of the subjectivity of grades and academic expectations, but more work needs to be done to examine the education provided to students when they are excluded from class. It is quite possible that students who attend their alternative education setting, whether that be in their school or at an ALC, receive approximately the same level of education as they do in their traditional classrooms. More research needs to be conducted on the access to and quality of education students receive once they are sentenced with a long-term suspension and sent to an ALC. Furthermore, future research should attend to the reintegration of students into their home schools following a long-term suspension, and the way that schools attempt to make discipline policy that benefits not just the school as a whole, but also the offending student. Another line of research that is warranted is examining the extent to which the relationships between suspension and outcomes vary as a function of the type of school students attend. More specifically, I wonder if the relationships between suspension and student outcomes are different at Second Opportunity schools than at traditional NYCDOE schools. It is possible that students who attend these schools following suspension have differing trajectories as a result of their alternative placement.

Indeed, these relationships highlight the fact that suspensions may be designed to serve non-suspended students. In designing and implementing discipline policy, districts and schools may make trade-offs between the academic well-being of the majority of the student body, and the well-being of each individual student, including those who may be troubled. These trade-offs are problematized by limited resources. Schools and districts may find it easier to exclude students who may detract from their learning environment, even if doing so does not help those students improve their own outcomes.

Moreover, the overall low performance of suspended students when compared with non-suspended students suggests that neither suspensions, nor the schools overall, appear to be meeting these students' needs. It is possible that suspended students in more recent high school cohorts are being better served – or that the relationships between suspension and outcomes have weakened – given the numerous changes in NYCDOE discipline policy and due to the increasing efforts to implement the RFA curriculum. It seems plausible that the relationships uncovered in this work would be consistent with those pertaining to the current cohort of high school students, but that once students who experienced the shifting discipline policies in elementary school attend high school, the estimates would change. Investigating the efficacy of these changes is out of the scope of this work, but demands future attention.

Ultimately, additional research on the reasons behind disproportionate suspension rates, on discipline systems particular to New York - such as the use of Alternate Learning Centers - and on systems utilized nationwide, such as PBIS, is critical if we are to better understand our use of exclusionary discipline and to ensure that all students are guaranteed the right to a quality education. Utilizing a discipline system that benefits the students most likely to be excluded will not only help those students in the short term, but will also likely play a role in disrupting the cycle of poverty and recidivism to which these students too often belong.

References

The Advancement Project and The Civil Rights Project at Harvard University. (2000). Opportunities Suspended: The Devastating Consequences of Zero Tolerance and School Discipline, Report from a National Summit on Zero Tolerance.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the Classroom Assessment Scoring System-Secondary. *School Psychology Review*, *42*(1), 76-98.

Allensworth, E.M., & Easton, J.Q. (2005). The on-track indicator as a predictor of high school graduation. Chicago, IL: Consortium on Chicago School Research. Retrieved from https://ccsr.uchicago.edu/sites/default/files/publications/p78.pdf

Allison, P.D. (2002). *Missing Data.* Thousand Oaks, CA: Sage University Paper No. 136.

American Civil Liberties Union (ACLU) and New York Civil Liberties Union (NYCLU), (2007). Criminalizing the Classroom: The Over-Policing of New York City Schools.

American Psychological Association (APA) Zero Tolerance Task Force. (2008). Are zero tolerance polices effective in the schools? An evidentiary review and recommendations. *American Psychologist*, *63*(9), 852-862.

Anderson, J.A. (1988). Cognitive styles and multicultural populations. *Journal of Teacher Education, 24*(1), 2-9.

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. doi: 10.1177/0013124506286947

Aud, S., Fox, M.A., & KewelRamani, A. (2010). *Status and trends in the education of racial and ethnic minorities (NCES 2010-015).* Washington, DC: National Center for Educational Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved October 13, 2012, from http://www.air.org/files/AIR-NCESracial_stats__trends1.pdf

Aud, S., KewalRamani, A., and Frohlich, L. (2011). America's Youth: Transitions to Adulthood (NCES 2012-026). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Baker, A. (2012, August 28). New code aims to ease suspension of students. *The New York Times.* Retrieved from http://www.nytimes.com/2012/08/29/education/new-york-education-department-revises-student-disciplinary-code.html?_r=0

Banks, C. (2013). Criminal Justice Ethics: Theory and Practice. Thousand Oaks, CA: Sage Publications.

Bear, G. G. (2000). School suspension and expulsion. (pp. 183-185). US: American Psychological Association. doi: 10.1037/10522-074

Beger, R.R. (2002). Expansion of police power in public schools and the vanishing rights of students. *Social Justice, 29*(1/2), 119-130.

Beger, R.R. (2003). The worst of both worlds: School security and the disappearing Fourth Amendment rights of students. *Criminal Justice Review, 28,* 336-354.

Bennett, C.O., & Harris, J.J. (1982). Suspensions and expulsions of male and black students: A study of the causes of disproportionality. *Urban Education, 16*(4), 399-423.

Bernstein, N. (2005). *All Alone in the World: Children of the Incarcerated.* NY: New Press.

Biskupic, J. (2000, March 1). Police search powers argued at high court. *Washington Post.*

Bowditch, C. (1993). Getting rid of troublemakers: High school disciplinary procedures and the production of dropouts. *Social Problems, 40,* 293-507.

Bowles, S., & Gintis, H. (1976). School in Capitalist America. New York, NY: Basic Books.

Bracy, N.L. (2010). Circumventing the law: Students' rights in schools with police. *Journal of Contemporary Criminal Justice, 26*(3), 294-315.

Bradley, C. L., & Renzulli, L. A. (2011). The complexity of non-completion: Being pushed or pulled to drop out of high school. *Social Forces, 90*(2), 521-545. doi: 10.1093/sf/sor003

Brown, B. (2006). Understanding and assessing school police officers: A conceptual and methodological comment. *Journal of Criminal Justice, 34*(6), 591-604.

Brown, T.M. (2007). Lost and turned out: Academic, social, and emotional experiences of students excluded from school. *Urban Education, 42*(5), 432-455.

Bronfenbrenner, U. (1979). *The Ecology of Human Development.* Cambridge, MA: Harvard University Press.

Brooks-Gunn, J., & Duncan, G. (1997). The effects of poverty on children. *The Future of Children: Children and Poverty, 7*(2), 55-71.

Bruns, E. J., Moore, E., Stephan, S. H., Pruitt, D., & Weist, M. D. (2005). The impact of school mental health services on out-of-school suspension rates. *Journal of Youth and Adolescence,34*(1), 23-30. doi: 10.1007/s10964-005-1333-z

Burkam, D.T., LoGerfo, L.F., Ready, D.D., & Lee, V.E. (2007). The differential effects of repeating kindergarten. *Journal of Education for Students Placed at Risk, 12*(2), 103-136.

Carter, P. (2006). Straddling boundaries: Identity, culture, and school. *Sociology of Education, 79*(4), 304-328.

Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Cambridge, MA: Washington Research Project.

Christle, C. A., Jolivette, K., & Nelson. C.M. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education, 28*(6), 325-339. doi: 10.1177/07419325070280060201

Christle, C., Nelson, C. M., & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education & Treatment of Children, 27*(4), 509-526.

Collaborative for Academic, Social, and Emotional Learning (CASEL). (2014). What is social and emotional learning? Retrieved from http://www.casel.org/social-and-emotional-learning

Costenbader, V.K., & Markson, S. (1994). School suspension: A survey of current polices and practices. *NASSP Bulletin, 78*(103), 103-107. doi: 10.1177/019263659407856420

Costenbader, V., & Markson, S. (1998). School suspension. *Journal of School Psychology, 36*(1), 59-82. doi: 10.1016/S0022-4405(97)00050-2

Cromidas, R. (2012a, August 28). Disciplinary code revisions could reduce student suspensions. *Chalkbeat New York.* Retrieved from http://ny.chalkbeat.org/2012/08/28/disciplinary-code-revisions-could-reduce-student-suspensions/

Cromidas, R. (2012b, November 16). Suspension rates continue to raise concerns, even as they drop. *Chalkbeat New York.* Retrieved from http://ny.chalkbeat.org/2012/11/16/suspension-rates-continue-to-raise-concerns-even-as-they-drop/

Decker, G. (2013, October 31). Suspensions fall, but disparities for some students persist. *Chalkbeat New York.* Retrieved from http://ny.chalkbeat.org/2013/10/31/suspensions-fall-but-disparities-for-some-students-remain/

Delpit, L. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review, 58*(3), 280-298.

Delpit, L. (2006). Other People's Children: Cultural Conflict in the Classroom. New York, NY: The New Press.

Devine, P.G., & Elliott, A.J. (2000). Are racial stereotypes really fading? The Princeton trilogy revisited. In C. Stangor (Eds.), *Stereotypes and Prejudice: Essential Readings.* New York: Psychology Press.

Devoe, J. F., & Darling-Churchill, K. E. (2008). Status and trends in the education of American Indians and Alaska Natives: 2008 (NCES 2008-084). Washington, DC: National Center for Education Statistics, Institute of Educational Sciences, U.S. Department of Education.

Devoe, J.F., Peter, K., Kaufman, P., Miller, A., Noonan, M., Snyder, T.D., & Baum, K. (2004). *Indicators of crime and safety: 2004*(U.S. Departments of Education and Justice, NCES 2005-002/NCJ 205290). Washington, DC: U.S. Government Printing Office.

Diamond, J.B., Randolph, A., & Spillane, J.P. (2004). Teachers' expectations and sense of responsibility for student learning: The importance of race, class, and organizational habitus. *Anthropology and Education Quarterly, 35*(1), 75-98.

Dinkes, R., Cataldi, E. F., Lin-Kelly, W. (2007). *Indicators of school crime and safety: 2007. (NCES 2008-021/NCJ 219553).* Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice.

Dohrn, B. (2001). Look out, kid, it's something you did: The criminalization of children. In Valerie Polakow (ed.), *The Public Assault on America's Children.* New York: Teachers College Press.

Eckholm, E. (2010a, March 18). School suspensions lead to legal challenge. *The New York Times.* Retrieved from http://www.nytimes.com/2010/03/19/education/19suspend.html

Eckholm, E. (2010b, October 8). Ruling limits state's power in school suspensions. *The New York Times.* Retrieved from http://www.nytimes.com/2010/10/09/us/09suspend.html

Entman, R.M., & Rojeck, 2000. *The black image in the white mind: media and race in America.* Chicago: University of Chicago Press.

Farkas, G. (2003). Racial disparities and discrimination in education: What do we know, how do we know it, and what do we need to know? *Teachers College Record, 105*(6), 1119-46.

Farkas, G., Grobe, R., Sheehan, D., & Shuan, Y. (1990). Cultural resources and school success: Gender, ethnicity, and poverty groups within an urban district. *American Sociological Review, 55*, 127-142.

Fenning, P., & Bohanon, H. (2006). Schoolwide discipline policies: An analysis of the discipline code of conduct. In C.M. Evertson & C.S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 1021-1040). Mahwah, NJ: Lawrence Erlbaum.

Fenning, P., & Rose, J. (2007). Overrepresentation of African-American students in exclusionary discipline the role of school policy. *Urban Education, 42*(6), 536-559. doi: 10.1177/0042085907305039

Fine, M. (1991). *Framing dropouts: Notes on the politics of an urban public high school.* Albany, NY: State University of New York Press.

Franklin, M. (1992). Culturally sensitive instructional practices for African-American learners with disabilities. *Exceptional Children, 59*(2), 115-122.

Gable, R.A. (2004). Hard times and uncertain future: Issues that confront the field of emotional/behavioral disorders. *Education and Treatment of Children, 27,* 341-352.

Gable, R.A., Bullock, L.M., & Evans, W.H. (2006) Changing perspectives on alternative schooling for children and adolescents with challenging behavior. *Preventing School Failure, 51*(1), 5-9.

Gallagos, A.H., & White, C.R. (2013). Preventing the school-justice connection for youth in foster care. *Family Court Review, 51*(3), 460-468.

Garland, D. (2001). The culture of control: Crime and social order in contemporary society. Chicago, IL: The University of Chicago Press.

Gastic, B. (2013). *Disproportionality in school discipline in Massachusetts.* Manuscript submitted for publication.

Gay, G. (2006). Connections between classroom management and culturally responsive teaching. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management* (pp. 735-786). Mahwah, NJ: Lawrence Erlbaum.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Gilliam, F.D., & Iyengar, S. (2000). Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science, 44*(3), 560-573.

Gootman, E. (2004, July 30). Students' suit on special ed becomes class action. *The New York Times.* Retrieved from http://www.nytimes.com/2004/07/30/nyregion/students-suit-on-special-ed-becomes-class-action.html

Gordon, R., Piana, L. D., & Keleher, T. (2000). *Facing the consequences: An examination o racial discrimination in U.S. public schools*. Oakland, CA: ERASE Initiative, Applied Research Center.

Gregory, A., Cornell, D., Fan, X., Sheras, P., Shih, T., & Huang, F. (2010). Authoritative school discipline: High school practices associated with lower bullying and victimization. *Journal of Educational Psychology, 102*(2), 483-496.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher, 39*(1), 59-68. doi: 10.3102/0013189X09357621

Hancock, L. (2001). Framing children in the news: The face and color of youth crime in America. In V. Polakow (Ed.), *The Public Assault on America's Children: Poverty, Violence, and Juvenile Justice* (pp. 78-100). New York: Teachers College Press.

Harball, E. (2012, May 1). New and improved suspension centers: The city's best kept secret? *School Stories.* Retrieved from http://school-stories.org/2012/05/new-and-improved-suspension-centers-the-citys-best-kept-secret/

Havalchek, A.H., White, C.R., O'Brien, K., Pecora, P., & Sepulveda, M. (2009). Foster care experiences and educational outcomes of young adults formerly placed in foster care. *School Social Work Journal, 34*(1), 1-27.

Hawkins, J.D., Smith, B.H., & Catalano, R.F. (2004). Social development and social and emotional learning. In J.E. Zins, R.P. Weissberg, M.C. Wang, & H.J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 135-150). New York: Teachers College Press.

Herda-Rapp, A. (2003). The social construction of local school violence threats by the news media and professional organizations. *Sociological Inquiry, 73,* 545-574.

Hirano, K., Imbens, G.W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica, 71,* 1161-1189.

Hemphill, S. A., Toumbourou, J. W., Herrenkohl, T. I., McMorris, B. J., & Catalano, R. F. (2006). The effect of school suspensions and arrests on subsequent adolescent antisocial behavior in Australia and the United States. *Journal of Adolescent Health, 39*(5), 736-744. doi: 10.1016/j.jadohealth.2006.05.010

Hyman, I.A. (1997). *School discipline and school violence: A teacher variance approach.* Boston: Allyn & Bacon.

Hyman, I.A., & Perone, D.C. (1998). The other side of school violence: Educator policies and practices that may contribute to student misbehavior. *Journal of School Psychology, 36*(1), 7-27.

Irvine, J. J. (2002). *In search of wholeness: African American teachers and their culturally competent classroom management practices.* New York: Palgrave.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* New York, NY: Harper.

Jenkins, A. H. (1982). The psychology of the Afro-American. Elmsford, NY: Permagon Press.

Jenkins, P. (1997). School delinquency and the school social bond. *Journal of Research in Crime and Delinquency, 34*(3), 337-367.

Jennings, W.G., Khey, D.N., Maskaly, J., & Donner, C.M. (2011). Evaluating the relationship between law enforcement and school security measures and violent crime in schools. *Journal of Police Crisis Negotiations, 11,* 109-124.

Johnson, D. (1999, December 1). Schools' new watchword: Zero tolerance. *The New York Times.* Retrieved from http://www.nytimes.com/1999/12/01/us/schools-new-watchword-zero-tolerance.html?pagewanted=all&src=pm

Johnson, I.M. (1999). School violence: The effectiveness of a school resource officer program in a southern city. *Journal of Criminal Justice, 27*(2), 173-192.

Kenney, D., & Watson, S. (1998). Crime in the schools: Reducing fear and disorder with student problem solving. Washington, DC: Police Executive Research Forum.

Kinsler, J. (2011). Understanding the black-white school discipline gap. *Economics of Education Review, 30*(6), 1370-1383. doi: 10.1016/j.econedurev.2011.07.004

Krezmien, M. P., Leone, P. E., & Achilles, G. (2006). Suspension, race, and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders, 14*(4), 217-226.

Kupchik, A. (2010). Homeroom security: School discipline in an age of fear. New York, NY: New York University Press.

Kupchik, A., & Bracy, N.L. (2009). The news media on school crime and violence: Constructing dangerousness and fueling fear. *Youth Violence and Juvenile Justice,7*(2), 136-155.

Lait, M. (1999, June 3). Officers easily blend into school drug scene. *Los Angeles Times.*

Langberg, J.B., & Fedders, B.A. (2013). How juvenile defenders can help dismantle the school-to-prison pipeline: A primer on educational advocacy and incorporating clients' education histories and records into delinquency representation. *Journal of Law and Education*, *42*(4), 653-689.

La Paro, K.M., Pianta, R.C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal*, *104*(5), 409-426.

Lareau, A., & Horvat, E. M. (1999). Moments of social inclusion and exclusion: Race, class, and cultural capital in family-school relationships. *Sociology of Education, 72*(1), 37-53.

Lawrence, R., & Mueller, D. (2003). School shootings and the man-bites-dog criterion of newsworthiness. *Youth Violence and Juvenile Justice, 1*(4), 330-345.

Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children,34*(2), 167-192. doi: 10.1353/etc.2011.0014

The Legal Aid Society. (2013). Know your rights: School suspensions. Retrieved from http://www.legal-aid.org/media/152368/kyr_school_suspensions_en.pdf

Leone, P. E., Mayer, M. J., Malmgren, K., & Meisel, S. M. (2000). School violence and disruption: Rhetoric, reality, and reasonable balance. *Focus on Exceptional Children, 33*(1), 1-20.

Leone, P.E., & Weinberg, L. (2010). *Addressing the unmet needs of children and youth in the juvenile justice and child welfare systems.* Washington, DC: The Center for Juvenile Justice Reform, Georgetown University.

Lindle, J.C. (2008). School safety: Real or imagined fear? *Educational Policy, 22*(1), 28-44.

Lipsky, M. (1980). *Street-level Bureaucracy: Dilemmas of the Individual in Public Services.* New York, NY: Russell Sage Foundation.

Martens, B.K., & Meller, P.J. (1990). The application of behavioral principles to educational settings. In T.B. Gukin & C.R. Reynolds (Eds.), *Handbook of School Psychology* (pp. 612-634). New York: Wiley.

May, D.C., Fessel, S.D., & Means, S. (2004). Predictors of principals' perceptions of school resource officer effectiveness in Kentucky. *American Journal of Criminal Justice*, *29*(1), 75-93.

Mayer, S.E. (1997). *What money can't buy: Family income and children's life chances.* Cambridge, MA: Harvard University Press.

Mayer, M.J., & Leone, P.E. (1999). A structural analysis of school violence and disruption: Implications for creating safer schools. *Education and Treatment of Children, 22*(3), 333-356.

Maynard-Moody, S., & Musheno, M. (2003). *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service.* Ann Arbor, MI: University of Michigan Press.

McCarthy, J.D., & Hoge, D.R. (1987). The social construction of school discipline: Racial disadvantage out of universalistic processes. *Social Forces, 65*(4), 1101-1120.

McCorkle, R.C., & Miethe, T.D. (1998). The political and organizational response to gangs: An examination of a moral panic in Nevada. *Justice Quarterly, 15,* 41-64.

McIntyre, T. (1993). Reflections on the impact of the proposed definition for emotional and behavioral disorders: Who will still fall through the cracks and why. *Behavioral Disorders, 18*(2), 148-160.

Miller, J.M., Gibson, C., Ventura, H.E., & Schreck, C.J. (2005). Reaffirming the significance of context: The Charlotte School Safety Program. *Journal of Criminal Justice, 33*(5), 477-485.

Morgan, S.L., & Harding, D.J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological Methods and Research, 35,* 3-60.

Morrison, G. M., Anthony, S., Storino, M., & Dillon, C. (2001). An examination of the disciplinary histories and the individual and educational characteristics of students who participate in an in-school suspension program. *Education and Treatment of Children, 24*(3), 276-93.

Morrison, B.D., & Vaandering, D. (2012). Restorative justice: Pedagogy, praxis, and discipline. *Journal of School Violence, 11*(2), 138-155.

Mukherjee, E. (2007). *Criminalizing the classroom: The over-policing of New York City schools.* New York: New York Civil Liberties Union.

NAACP Legal Defense & Educational Fund, Inc., "Dismantling the School-to-Prison Pipeline" (2005).

Nathanson, L., McCormick, M., & Kemple, J.J. (2013). Strengthening assessments of school climate: Lessons from New York City. New York, NY: The Research Alliance for New York City Schools. Retrieved from http://media.ranycs.org/2013/011

National Center for Education Statistics. (2006). *Indicators of school crime and safety, 2006.* Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.

The New York Civil Liberties Union (NYCLU). (2011). Education interrupted: The growing use of suspension in New York City's public schools. Retrieved from http://www.nyclu.org/files/publications/Suspension_Report_FINAL_noSpreads.pdf

The New York Civil Liberties Union (NYCLU). (2013). A, b, c, d, STPP: How school discipline feeds the school-to-prison pipeline. Retrieved from http://www.nyclu.org/files/publications/nyclu_STPP_1021_FINAL.pdf

New York City Department of Education (NYCDOE). (2014a). Alternate Learning Center. http://schools.nyc.gov/Offices/ALC/default.htm

Neild, R.C. (2009). Falling off track during the transition to high school: What we know and what can be done. *The Future of Children, 19*(1), 53-76.

New York City Department of Education (NYCDOE). (2014b). Respect for All. http://schools.nyc.gov/RulesPolicies/RespectforAll/Commitment/default.htm

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly, 90*(4), 1003-1018.

Noguera, P.A., & Akom, A. (2000). The opportunity gap. *Wilson Quarterly, 24,* 86.

Osher, D., Bear, G.G., Sprague, J.R., & Doyle, W. (2010). How can we improve school discipline? *Educational Researcher, 39*(1), 48-58.

Pavelka, S. (2013). Practices and policies for implementing restorative justice within schools. *The Prevention Researcher, 20*(1), 15-17.

Pianta, R.C., & Hamre, B.K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109-119.

Raffaele Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development, 2003*(99), 17-33. doi: 10.1002/yd.52

Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children, 26*(1), 30-51.

Raffaele Mendez, L. M., Knoff, H. M., & Ferron, J. M. (2002). School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district. *Psychology in the Schools, 39*(3), 259-277. doi: 10.1002/pits.10020

Ransom, C.J. (1999). Does the end justify the means? Use of juveniles as government informants, helpful to society while harmful to the child. *Journal of Juvenile Law, 20,* 108-129.

Raudenbush, S.W. (2009). Adaptive centering with random effects: An alternate to the fixed effects model for studying time-varying treatments in school settings. *Education Finance and Policy, 4*(4), 468-491.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models* (2nd ed.). Thousand Oaks, CA: Sage.

Reschly, D. (1997). *Disproportionate minority representation in general and special education: Patterns, issues, and alternatives.* Des Moines, IA: Drake University, Mountain Plans Regional Resource Center.

Reyes, A.H. (2006). *Discipline, achievement, and race: Is zero tolerance the answer?* Lanham, MD: Rowman & Littlefield Publishing Group.

Rist, R.C. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review, 40*(3), 411-51.

Roberts, D. (2004). The social and moral cost of mass incarceration in African American communities. *Stanford Law Review, 56*(127), 1271-1305.

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal, 36,*303-343.

Roque, M.(2010). Office discipline and student behavior: Does race matter? *American Journal of Education, 116*(4), 557-581.

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

Rubin, D.B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91,* 473-489.

Rubin, D.B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127,* 757-763.

Rubin, D.B., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics, 52,* 249-264.

Rueda, R. S., & Forness, S. R. (1994). Childhood depression: Ethnic and cultural issues in special education. In R. Peterson & S. Ishii-Jordan (Eds.). *Multicultrual issues in the education of students with behavioral disorders.* Cambridge, MA: Brookline Press. (pp. 40-62).

Santos, F. (2011, January 27). Sharp rise in suspension at city's schools is cited. *The New York Times.* Retrieved from http://www.nytimes.com/2011/01/28/nyregion/28suspend.html

Scherr, T. (2007). Education experiences of children in foster care: Meta-analyses of Special Education, retention, and discipline rates. *School Psychology International, 28*(4), 419-436.

Scott, T.M., & Barrett, S.B. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of school-wide PBS. *Journal of Positive Behavior Interventions, 6,* 21-27.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton Mifflin.

Shores, R.E., Gunter, P., & Jacks, S. (1993). Classroom management strategies: Are they setting events for coercion? *Behavioral Disorders, 18*, 92-102.

Skiba, R. (2000). Zero tolerance, zero evidence: An analysis of school disciplinary practice. Policy Research Rep. No. SRS2. Bloomington, IN: Indiana Education Policy Center.

Skiba, R. (2002). Special education and school discipline: A precarious balance. *Behavioral Disorders, 27*(2), 81-97.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review,34*(4), 317-342. doi: 10.1023/A:1021320817372

Skiba, R., & Peterson, R. (1999). The dark side of zero tolerance. *Phi Delta Kappan, 80*(5), 372-382.

Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children, 20*(3), 295-315.

Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10*(3), 297-306.

Suh, S., Suh, J., & Houston, I. (2007). Predictors of categorical at-risk high school dropouts. *Journal of Counseling & Development, 85*(2), 196-203.

Theriot, M.T. (2009). School resource officers and the criminalization of student behavior. *Journal of Criminal Justice, 37*(3), 280-287.

Theriot, M. T., Craun, S. W., & Dupper, D. R. (2010). Multilevel evaluation of factors predicting school exclusion among middle and high school students. *Children and Youth Services Review, 32*(1), 13-19. doi: 10.1016/j.childyouth.2009.06.009

Tobin, T., Sugai, G., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders, 4*(2), 82-94. doi: 10.1177/106342669600400203

Townsend, B. (2000). The disproportionate discipline of African-American learners: Reducing school suspensions and expulsions. *Exceptional Children, 66*(3), 381-391.

United States Department of Education (USDOE). (2014). *Guiding Principles: A Resource Guide for Improving School Climate and Discipline.* Washington, DC.

Vavrus, F., & Cole, K. (2002). "I didn't do nothin'": The discursive construction of school suspension. *The Urban Review, 34*(2), 87-111. doi: 10.1023/A:1015375215801

Von Hippel, P.T. (2007). Regression with missing ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology, 37*, 83-117.

Wald, J., & Losen, D.J. (2003). Defining and redirecting a school-to-prison pipeline. In J. Wald & D.J. Losen (Eds.), *Deconstructing the School-to-Prison Pipeline: New Directions for Youth Development* (pp.9-16). San Francisco, CA: Jossey-Bass.

Walker, H.M., Horner, R.H., Sugai, G., Bullis, M., Sprague, J.R., Bricker, D., & Kaufman, M.J. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders, 4*, 194-209.

Walker, H., & Sprague, J. (1999). The path to school failure, delinquency, and violence: Causal factors and some potential solutions. *Intervention in School and Clinic, 35*(2), 67-73. doi: 10.1177/105345129903500201

Wallace, J. M., Jr, Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *Negro Educational Review, 59*(1-2), 47-62.

Walz, M. (2010, December 20). After two years, Council moves to change school safety reports. *Chalkbeat New York.* Retrieved from http://ny.chalkbeat.org/2010/12/20/after-two-years-council-moves-to-change-school-safety-reports/

Wehlage, G. G., & Rutter, R. A. (1986). Dropping out: How much do schools contribute to the problem? *Teachers College Record, 87*(3), 374-92.

Westat, Inc. (2004). Summary of task force meeting on racial/ethnic disproportionality in special education. Washington, DC: Author.

Wildeman, C. (2009). Parental imprisonment, the prison boom, and the concentration of childhood disadvantage. *Demography, 46*, 265-280.

Willis, P. (1977). Learning to Labour: How Working Class Kids Get Working Class Jobs. New York, NY: Columbia University Press.

Wilson, D.B., Gottfredson, D.C., & Najaka, S.S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology, 17*(3), 247-272.

Zhang, D., Katsiyannis, A., & Herbst, M. (2004). Disciplinary exclusions in special education: A 4-year analysis. *Behavioral Disorders*, *29*(4), 337-347.