THREE ESSAYS IN APPLIED MICROECONOMICS

Chunxiao Li

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2017

Approved by:

Donna Gilleskie

Jane Fruehwirth

Valentin Verdier

David Guilkey

Helen Tauchen

Ju Hyun Kim

ABSTRACT

CHUNXIAO LI: Three Essays in Applied Microeconomics.
(Under the direction of Donna Gilleskie)

This dissertation is comprised of three independent chapters. The first chapter studies the effectiveness and consequences of exclusionary school discipline. Exclusionary school discipline techniques, such as out-of-school suspension, are often criticized for their inability to improve students' behavior, their adverse effects on students' achievement outcomes and their disproportionate use on minority students. Using large-scale administrative data on North Carolina public school students, I find that harsher disciplinary rules (measured by higher out-of-school suspension likelihood) significantly deter students from committing first offenses, but that they are less effective (or ineffective) for repeat offenses. I also find that their adverse effects on offending students' achievement outcomes, such as end-of-grade test scores and high school dropout probability, are much smaller than the effects documented in the existing literature. In addition, I find that harsher disciplinary rules could significantly improve the academic achievement of middle school students with no offense record. To carefully address endogeneity and selection issues in a large-scale data context, my preferred identification strategy combines the instrumental variable method and a machine learning cluster method (k-means). These findings suggest that current policy reform of exclusionary school discipline should carefully balance its benefits and costs for different student populations.

The second chapter explores the equity in exclusionary school discipline between black and white students and among students from families with different economic backgrounds. The existing literature and popular press report that black students face out-of-school suspension with much greater frequency than white students. Using administrative data on North Carolina public school students over eight academic years, I find that the racial disparity depends importantly on the type

of offenses when black and white students are compared within the same school. While black students are more likely to be suspended, for example, for fighting, theft and sexual harassment, white students are more likely to be suspended for insubordination, disrespect toward faculty, or leaving class without permission. I also find that Economically Disadvantaged students are consistently more likely to be suspended out-of-school for different types of offenses, even if the comparison is within schools.

The third chapter studies the impacts of social contacts, such as spouses, friends, siblings, parents or children on individual smoking behavior. To identify endogenous social interaction effects, we model an individual and her social contacts' smoking behaviors as a simultaneous move game with complete information. We also allow an individual's smoking behavior to depend on her previous behavior and unobserved heterogeneity. Using unique data from the Framingham Heart Study, which includes complementary social network data, we find statistically significant endogenous social interaction effects of spouses and friends on individual smoking behavior. We also find that endogenous social interaction effects from siblings or parents are not statistically significant after disentangling them from homophily. In addition, we find that the effects of social contacts' cardiovascular disease shocks on individual smoking behavior are not statistically significant.

ACKNOWLEDGMENTS

I would like to thank my advisor, Donna Gilleskie, for her guidance over the years. Without her persistent support and encouragement, this dissertation would not have been possible. I also thank my committee members, Jane Fruehwirth, Valentin Verdier, David Guilkey, Helen Tauchen and Ju Hyun Kim for their advice, guidance and support through this project. In addition, I thank Jonathan Williams, Brian Mcmanus, Tiago Pires, Peter Arcidiacono, Patrick Bayer, Stephane Bonhomme, Thibaut Lamadon, Jeffrey Wooldridge and participants at the UNC applied microeconomics workshops for their valuable feedback. I acknowledge data access provided by the North Carolina Education Research Data Center and computing resources provided by UNC ITS Research Computing. Finally, I thank my family members for their love and support. I thank my father, Xingmin Li, my grandmother, Yumei Chen, and especially my wife, Yue Wu; this journey would not have been possible without their unconditional support, love and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE EFFECTIVENESS AND CONSEQUENCES OF EXCLUSIONARY SCHOOL DISCIPLINE

1.1 Introduction

In U.S. public schools, exclusionary school discipline techniques, such as out-of-school suspension or expulsion, are commonly used methods to address student misbehaviors – ranging from severe misconduct (e.g., assaults at school) to minor offenses (e.g., disruptive behavior in the classroom). In the 2013-2014 school year alone, 2.8 million of the 50 million public school students were suspended out-of-school at least once, and 130,000 students were expelled.[1] The high rate of out-of-school suspension reflects the consequences of policies such as "zero tolerance," which emphasize tough punishment, including social exclusion, as a primary response to crime or misconduct (Skiba and Knesting 2001; Losen and Skiba 2010). However, this disciplinary practice is widely criticized for its inability to improve students' misbehavior and its adverse consequences for suspended students and the broader school community.

Some existing literature suggests that suspension does little to discourage misbehavior and may, in fact, encourage it (Wettach, Owen, and Hoffman 2015; Skiba and Rauch 2015). The literature also finds that suspension lowers academic achievement and raises school dropout rates of offending students (Raffaele Mendez 2003; Arcia 2006; Lee, Cornell, Gregory, and Fan 2011; Skiba and Rauch 2015). Students who are suspended or expelled from school are more likely to be involved in the justice system; this relationship is often referred to as the "school to prison pipeline" (Wald and Losen 2003). A recent study shows that high rates of school suspensions actually harm math and reading scores for non-suspended students (Perry and Morris 2014). Concerns about these adverse effects are amplified by findings of disparities in school disciplinary practices, especially the

[1]Civil Rights Data Collection, 2016, U.S. Department of Education Office for Civil Rights.

disproportionate representation of students of color in school suspension rolls. In the 2013-2014 academic year, while black students represented 15.5 percent of the public school student population, they comprised 39 percent of student out-of-school suspensions.[2] This disparity suggests that the disciplinary practice is particularly harmful to minority groups. Morris and Perry (2016) document that school suspensions produce a racial achievement gap, which accounts for approximately one-fifth of black-white differences in school performance.[3] These findings have motivated the U.S. Department of Justice and Education to release a school discipline guidance package in 2014 to reform discipline policies and practices, several states to enact new legislation, and many schools to consider ongoing school discipline policy reform.[4]

Despite concern expressed in the literature and in the popular press, the effectiveness and consequences of exclusionary school discipline techniques remain controversial. Several issues inherent in identifying "causal effects" in this context, such as endogeneity and selection issues, have not been fully addressed in the literature, which mostly uses descriptive statistics or regressions with limited sets of observables. For example, the negative correlation between out-of-school suspension and suspended students' achievement may reflect the following causal relationships. First, principals may be more likely to suspend less-engaged students or students who commit more serious offenses (even for the same observed type of offense). The lower achievement outcomes of suspended students may not be explained by the consequences of suspension but just simply reflect that these students are "bad apples."[5] Second, students who choose to commit offenses under the threat of suspension may have different personality traits, such as lacking self-control, compared to students who commit offenses under no threat of suspension (the "selection on students'

[2]Civil Rights Data Collection, 2016, U.S. Department of Education Office for Civil Rights.

[3]Black-white academic achievement gap is of great importance for understanding black-white gaps in economic outcomes. For example, Neal and Johnson (1996) document that a test score (AFQT) explains nearly three-quarters of the racial wage gap for young men and all of the gap for young women.

[4]See the following websites for the guidance package: http://www.ojjdp.gov/enews/14juvjust/140109.html; http://www2.ed.gov/policy/gen/guid/school-discipline/index.html. There is new related legislation in several states; for example, California (AB 420, 2014) and Illinois (SB 100, 2015).

[5]I refer to this issue as the "endogeneity issue" caused by principals' decisions.

unobservables").[6] Therefore, the lower achievement outcomes of suspended offending students may be due to the lack of self-control but not the suspension itself. These issues cannot be fully addressed using regressions with limited sets of observables because the engagement levels of students, the severity of offenses and the personality traits are typically not observed by researchers.[7] Furthermore, they cannot be fully addressed with student- or school-level fixed effects (or similar methodologies) because some of the unobservables, such as the severity of offenses or other life-changing circumstances of students, are time-varying factors.[8]

While attempting to carefully address the identification issues, this essay studies the causal effects of exclusionary school discipline on students' in-school behaviors and achievement outcomes, such as end-of-grade test scores, dropout probability and ACT scores.[9] With regard to students' in-school behaviors, I separately identify a "general deterrence effect" and a "specific deterrence effect" of the discipline. As punishments, these discipline techniques may serve as threats to students who intend to commit offenses and deter them from infractions. I refer to this mechanism as "general deterrence effect" following the economics of crime literature.[10] In addition, the experience of a specific punishment may serve as an effective "wake-up call" and decrease students' likelihood of re-offending in the future. I refer to this mechanism as a "specific deterrence

[6] Heckman, Stixrud, and Urzua (2006) discuss how personality traits affect risky behaviors of children.

[7] Even if we observe the type of offense (e.g., fighting), we typically do not observe additional details about the offense (e.g., severity).

[8] A fixed effect approach uses differences in punishment across the same type of offenses for the same student or in the same school to identify the effect. However, the fact that the same type of offenses may be punished differently for the same student or in the same school over time is likely to reflect that some unknown sources that determine the punishment have changed, such as the severity of offenses. There are also other limitations for applying the fixed effect approach in this context. For example, the student fixed effect is not applicable when the student's outcome, such as ACT score, is not observed repeatedly.

[9] Specifically, I focus on the effects of out-of-school suspension (including expulsion) in this essay because it is the major controversy in policy making.

[10] The standard economic model of criminal behavior with the discussion of deterrence is introduced by Becker (1968). There is a large empirical literature that estimates deterrence effects of police or sanctions on criminal activities. Comprehensive reviews include, for example, Levitt and Miles (2006), Durlauf and Nagin (2011) and Chalfin and McCrary (2017).

effect."[11]

I also examine the impact of exclusionary school discipline on students' achievement outcomes. When a principal faces a decision to out-of-school suspend an offending student or not, an important effect under consideration is whether the suspension will lead to worse (or better) achievement outcomes for the offending student.[12] Furthermore, since one reason for suspending offending students is to benefit "well-behaved" students, an important question is whether harsher disciplinary rules (i.e., with higher probability of out-of-school suspension) can benefit achievement outcomes of those "well-behaved" students. In addition, for principals to make disciplinary rules or decisions and governments to decide related policies, it is important to know the total effects of the exclusionary school discipline on achievement outcomes of all students. Therefore, I separately identify the effect of exclusionary school discipline on the achievement outcomes of offending students, students with no offense record, and all students.[13]

Using linked administrative data with detailed misbehavior records for all North Carolina public school students in grades 3-12 from the 2008-2009 to 2014-2015 academic years, this research makes several contributions to the literature. First, I find statistically significant "general deterrence effects" of harsher discipline rules (measured by out-of-school suspension likelihood) on students' first offenses. The effects are heterogeneous for different types of offenses and different student subpopulations. For example, while a 10 percentage point increase in the out-of-school suspension likelihood reduces the mean rate of first offenses for most categories or types of offenses by 7 to 40 percent, this effect is not statistically significant for some types of offenses, such as "fighting." In addition, I find that the effect is generally smaller for students' repeat offenses

[11]The general and specific deterrence effects can be with opposite signs since the exclusionary school discipline may encourage students to commit offenses and the experience of out-of-school suspension may be a "bad lesson" to students and increase their likelihood of re-offending instead.

[12]In North Carolina, as in most U.S. states, local boards of education establish their own disciplinary policies following the broad principles in the state statutes. However, the disciplinary policies often list a broad range of punishments for each type of misbehavior and, thus, school principals or assistant principals may, at their discretion, determine appropriate punishment for each student infraction instance.

[13]A student with no offense record means that she does not have an offense record in my sample period. I use them to represent "well-behaved" students.

and not statistically significant for students who have already had an out-of-school suspension experience. I also find suggestive evidence that the "specific deterrence effect" is either small or not statistically significant.

Second, and contrary to the existing literature, I find that the effects of suspension experience on offending students' achievement outcomes are either small or not statistically significant. For example, an OLS regression with a limited set of controls suggests that a suspension is associated with lower end-of-grade math test scores of offending students by as much as 0.2 standard deviations, but estimates using my preferred empirical strategy suggests that this effect of suspension is not statistically significant. Furthermore, while the OLS regression suggests that suspension experience is positively associated with an offending high school student's dropout probability with increases as high as 18 percentage points, my preferred estimator suggests no statistically significant causal effect. Since the major argument of "school to prison pipeline" theory is that suspension increases the probability of dropping out of high school, and high school dropouts are more likely to be involved with the justice system, this invalidation of the first premise suggests that the "school to prison pipeline" argument is questionable. In addition, these results do not support the argument that the black-white suspension disparity creates a black-white achievement gap.

Finally, I find that harsher disciplinary rules have statistically significant positive but small effects on end-of-grade math scores of students with no offense record.[14] For example, I find that a 10 percentage point increase in the out-of-school suspension likelihood in a school could increase the end-of-grade math scores of middle school students with no offense record by about 0.02 standard deviations. I also find that they have an overall positive effect on all middle school white students' end-of-grade math scores. However, I find that the effects are not statistically significant for other student populations or other achievement outcomes.

My preferred estimation procedure identifies the causal effects by a strategy that combines the instrumental variable (IV) method with a method (two-step grouped-fixed effects) that addresses

[14]This result is contrary to Perry and Morris (2014).

the "selection on students' unobservables" in a large-scale data context. The identification strategy exploits important features of the data; namely, students and principals are followed across academic years and across schools.[15] The IV is a measure of out-of-school suspension propensities of principal teams, which is constructed by principal team members' out-of-school suspension decisions in other schools. It is assumed to affect students' misbehavior or achievement outcomes only through the principal teams' discipline decisions in the school of concern.

A problem with directly applying the IV strategy is that the "selection on students' unobservables" is not addressed and it may jeopardize the validity of IV. Since traditional approaches to address this issue are quite computationally expensive given the sample size (to be detailed in section 1.3), my preferred model uses an empirical framework that combines the IV method with recent approaches (i.e., two-step grouped-fixed effects) that address the unobserved heterogeneity problem. Following the literature (Lin and Ng 2012; Bonhomme and Manresa 2015; Bonhomme, Lamadon, and Manresa 2016b), I model the student unobserved heterogeneity in a flexible yet parsimonious way, which allows it to vary across different groups (types) of students, across different types of misbehaviors and across different academic years. I estimate the causal effects by identifying group memberships of students in the first step, and then applying the IV strategy with an imputation of the group memberships of students in the second step. The step identifying students' group memberships applies the k-means clustering algorithm, which is widely used in machine learning and other related fields (Forgy 1965; Steinley 2006). Since the model also includes school fixed effects, the second step identifies the causal effects by comparing outcomes of otherwise identical students in the same identified unobserved heterogeneity groups in the same schools but assigned to principal teams with different out-of-school suspension propensities.

In addition to the literature mentioned above, the work of Kinsler (2013) is most closely related to my research. Using data on middle school students in the three largest school districts in North Carolina in one academic year (2000-2001), Kinsler (2013) studies the effects (and mechanisms) of

[15]The feature that principals transfer across schools is used to construct the IV. The feature that students are followed across academic years is used to identify the student unobserved heterogeneity.

racial disparities in school discipline on the racial achievement gap in end-of-grade test scores. This work documents significant "general deterrence effects" of out-of-school suspension.[16] It also documents that out-of-school suspension has an overall positive influence on middle school students' end-of-grade test scores. The study attempts to address the endogeneity and selection issues by including permanent unobserved student heterogeneity while jointly estimating structure parameters of student behavior, end-of-grade test score production, and principals' punishment decisions. It assumes, conditional on the observed student characteristics and the permanent unobserved student heterogeneity, that discipline punishments are exogenous; that is, they are predetermined by principals with identical preferences by their forward-looking evaluation of different parties' welfare at the beginning of the academic year. Using this assumption, the study achieves identification without an exclusion restriction that affects principals' punishment decisions but does not affect students' misbehaviors. However, a concern is that there might be unobservables that determine principals' decisions other than permanent unobserved student heterogeneity, such as school-level unobservables, time-varying environmental shocks, student level time-varying unobserved factors or the unobserved severity of offenses. My analysis shows that the exclusion restriction (or IV) I construct, which stems from heterogeneous preferences of principals (in contrast to Kinsler's identical preferences assumption), is important to recover the causal effects.

The rest of the chapter is organized as follows. In section 1.2, I describe the data used in estimation. In section 1.3, I present the empirical framework. Section 1.4 discusses important estimation details. Section 1.5 offers results. Section 1.6 concludes.

1.2 Data Description

1.2.1 Data Sources

This study uses administrative data from North Carolina public schools provided by the North Carolina Education Research Data Center (NCERDC). The data were originally collected by the North Carolina Department of Public Instruction (NCDPI) and the National Center for Education

[16]Comparing the magnitude of the general deterrence effects between my research and Kinsler (2013) is not straightforward because it uses days of suspension to construct the discipline measure.

Statistics (NCES). They include all North Carolina public school students' disciplinary infraction records, academic records and other administrative information. They also include teachers' and other licensed personnels' information, and other school-level statistics.[17]

The students' disciplinary infraction records were collected by the NCDPI through each local education agency's (LEA) superintendent's office (Director/Principal's office in the case of charter schools). They were first reported by the school disciplinary data coordinator, but the principal is ultimately responsible for the data elements. Due to state and federal statutes and state Board of Eduction policies, a record of offense incidents involving the following must be reported: 1.) any of 17 criminal acts committed on a school campus or in connection with a school function;[18] 2.) any act resulting in an out-of-school suspension or expulsion; 3.) any in-school suspension received by an exceptional student;[19] 4.) any of the following acts, regardless of consequences assigned: fighting (or affray), bullying, discrimination, harassment, a violent assault not resulting in serious injury, communicating threats, gang activity, extortion, property damage, and possession or use of tobacco products. In addition to the offense incidents required to be reported by statutes and policies, other routine disciplinary incidents were also recorded and reported for analysis or administrative purpose. Guidelines have been created to ensure consistent reporting.

The academic records were collected by the NCDPI, and include end-of-grade test scores, end-of-course test scores, ACT scores and other academic record information of all North Carolina

[17] Additional information on perceptions of school environments is obtained from a survey (NCTWCS) of all teachers, principals and other licensed personnel in North Carolina public and charter schools conducted by the North Carolina Professional Teaching Standards Commission (NCPTSC) and the Governor's office. Additional information on Positive Behavior Intervention and Support (PBIS) school recognition is collected from the NCDPI website.

[18] The 17 reportable acts are: homicide, assault resulting in serious bodily injury, assault involving the use of a weapon, rape, sexual offense, sexual assault, kidnapping, robbery with a dangerous weapon, robbery without a dangerous weapon, taking indecent liberties with a minor, assault on school personnel, bomb threat, burning of a school building, possession of alcoholic beverage, possession of controlled substance in violation of law, possession of a firearm or powerful explosive, and possession of a weapon. Robbery without a dangerous weapon was removed from this category (moved to category 5) after the 2009-2010 academic year. The 17 (or 16) criminal acts are required to be reported to law enforcement before the 2011-2012 academic year. Since the 2011-2012 academic year, possession of alcoholic beverage, bomb threat, and burning of a school building are no longer required to be reported to police.

[19] Guidelines in later years (e.g., 2014/2015) require reporting any act resulting in-school suspension. I do not find evidence in the data used for this project, that exceptional students are more likely to be reported than other students for in-school suspension.

public school students. In North Carolina, students in grades 3-8 are required to take end-of-grade tests in reading and math, and students in grades 9-12 are required to take end-of-course tests for Algebra I and English 1 when they are enrolled for credit in these courses.[20] In addition, beginning in the 2012-2013 academic year, every 11th grader is required to take the ACT college entrance exam as part of the new North Carolina Standard Course of Study. These test scores are used as academic achievement measures in this essay.[21]

The disciplinary infraction records were matched with the academic records and all other information by NCERDC.[22] Each student and licensed school personnel was assigned an unique identifier (a randomized number). This identifier allows me to follow a student or school personnel (e.g., principals or assistant principals) over time and across schools, which is important for my identification strategy.

1.2.2 Sample Construction

The disciplinary infraction data provided by NCERDC span academic years from 2000-2001 to 2014-2015. Beginning in the 2007-2008 academic year, reporting requirements for offenses have been greater, and the matching rate of infraction data with other data has largely increased.[23] However, in the 2013-2014 academic year, there may be a data imputing problem caused by an upgrade of the data system of the N.C. public schools.[24] Therefore, for most of the empirical

[20]An End-of-Course test in English 2 instead of English 1 has been required since the 2012-2013 academic year. The data also include records of students who were absent or exempt from the tests.

[21]Other academic records provided by NCERDC, such as end-of-grade or end-of-course test scores for writing, computer skills, biology, Algebra 2, Civics and Economics, U.S. History, Chemistry and Physics, and SAT scores are not used in this essay because they are not available for all the academic years (or only available for a subset of the students).

[22]The detail of the matching process is on the website of NCERDC: http://childandfamilypolicy.duke.edu/research/nc-education-data-center/list-files-variables/.

[23]The matching rate is greater than 99 percent since the 2007-2008 academic year.

[24]Since the 2013-2014 academic year, N.C. public schools have upgraded from an NC WISE to a Pearson's PowerSchool product to report suspension data. There were reports that the new system experienced crashes and technical issues.

work, I use data from the 2007-2008 to 2012-2013 academic years. I use the 2013-2014 and 2014-2015 academic year data mostly for robustness checks, except that the ACT score information and dropout information in these academic years are used for the main estimation.[25]

The data for charter schools are not used in this essay because the offense reporting requirement and rates can be different and school administrator information is largely unavailable. In addition, I use student observations in grades 3-12 since several common explanatory variables, such as economically disadvantaged status and limited English proficiency status, are not available for grades K-2.

In the data, there are 1,909,831 distinct public school students in grades 3-12 from the 2007-2008 to 2012-2013 academic years, who contribute 6,559,362 student-year observations.[26] For most of the empirical work, lagged student or school offense statistics are used as explanatory variables; in these cases, the 2007-2008 academic year data are used for explanatory variables but not dependent variables. Further, I drop schools missing one or more years of data. These deletions result in 1,687,330 distinct students and 5,271,039 student-year observations. I use this sample to calculate the summary statistics in the next subsection. Depending on the empirical work, there may be additional data selections. For example, for estimating the effects of suspension on end-of-grade test scores, I use students in grades 3-8. For estimating the effects on ACT scores, I mainly use disciplinary data of students in grade 9-10 from the 2009-2010 to 2012-2013 academic years, with their ACT scores from the 2012-2013 to 2014-2015 academic years. For the effects on dropout probability, I mainly use grade 9 students' disciplinary data with the information that they finally graduate or dropout from high school.[27]

[25]The IV is constructed by using punishment records from the 2007-2008 to 2014-2015 academic years. The 2013-2014 and 2014-2015 academic year data are used to ensure enough observations to construct the IV.

[26]A student-year observation is calculated as one yearly observation per student if the student is ever enrolled and assigned an identifier in the data. The student might have several offenses or transfer to different schools within one year, but they are all regarded as one observation in this calculation. Students who are not matched between disciplinary data and other administrative data (less than 1%) are not in this calculation and not used for this essay.

[27]In the dropout analysis, only first time grade 9 students (no students who repeat grade 9) are included in the sample. As discussed later, I use grade 9 students to avoid the dynamic selection problem. I drop grade 9 students in the 2012-2013 academic year because I only observe the dropout information until the 2014-2015 academic year.

1.2.3 Descriptive Statistics and Data Issues

There are 3,951,754 recorded offense instances in the constructed sample, which were committed by 651,040 distinct students with 1,236,497 student-year observations. That means about 23 percent of student-year observations have at least one offense in the academic year, and about 38 percent of distinct students have at least one offense record in the sample period. In addition, about 56 percent of offending students re-offended in the same academic year.[28] About 20 percent of distinct students (349,611 distinct students) had at least one out-of-school suspension record in the sample period.[29]

The NCDPI classified offenses using about 90 offense types.[30] There were nearly 40 consequence types that were assigned to the offenses. The most commonly used consequence types were out-of-school suspension and in-school suspension, each representing about 30 percent of all consequences in the data. Most of the out-of-school suspensions were short-term ($\leq$ 10 days); only one percent of out-of-school suspensions were long-term ($>$ 10 days).[31] There were only 146 cases of expulsions in the sample. Most other consequence types were less severe punishments than suspensions, such as lunch detention or a warning.[32] Since controversy in policy making surrounds out-of-school suspension, my empirical work simplifies the punishment as either out-of-school suspension (including expulsion) or not, where the "not" category includes in-school

---

Since there might still be censored information for students who repeat grades, I do robustness checks by only using students from the 2008-2009 to 2010-2011 academic years.

[28] The calculation is based on student-year observations but not distinct students.

[29] About 10 percent of student-year observations (577,886 student-year observations) were out-of-school suspended at least once in the academic year.

[30] Each offense instance may be described by multiple offense types. Since only about one percent of offenses have associated with them more than one offense types, I use the first (typically most serious) offense type to represent the offense. The number of offense types and the definition of each type changed across years. They only changed slightly, however, from 2007 to 2015 and most of the changes only involve adding new types.

[31] According to the North Carolina state statute, a long-term suspension must be assigned by the superintendent under a principal's recommendation (115C-390.7).

[32] One percent of offenses were assigned the consequence of "alternative learning program." Of these, about 40 percent were assigned out-of-school suspension at the same time. This analysis ignores "alternative learning program" as a type of punishment.

suspension or other less severe punishments.[33]

To simplify the discussion, I provide two classifications of the offense types based on their similarities.[34] The first classification contains six categories - "violence," "drug," "disrespect," "truancy," "property" and "other offenses." This classification is used to discuss deterrence effects by category. A more detailed classification further divides the "violence" and "other" categories into four sub-categories respectively, resulting in twelve categories. This classification is used for constructing explanatory variables and identifying student unobserved heterogeneity. Appendix A shows the offense types in each category by classification. One concern for estimating the deterrence effects of out-of-school suspension is that additional punishments assigned by the juvenile justice system for criminal acts are not observed. Very few violent offenses (less than 4 percent) and property offenses (less than 0.5 percent) required reporting to law enforcement. Most of the offenses in the "drug" category, however, required such reporting and might have resulted in additional punishment; results from this specification should be interpreted with caution.

Table 1.1 provides summary statistics for different categories of offenses and several types of offenses within categories.[35] The percentages of offenses punished by out-of-school suspension (Column 3-4) show that out-of-school suspension was frequently used even for minor offenses, such as disruptive behavior and excessive tardiness.[36] The generally higher rates of out-of-school suspension for students' second offenses than for their first offenses reflect that escalating punishment rules are commonly used in education practices. This finding motivates my empirical framework that separately specifies the first and second punishment in order to more fully evaluate the deterrence effects.

---

[33]Additional inquiry into the effects of in-school suspension on students' misbehavior or achievement outcomes would complement this analysis.

[34]The classifications also take into account the number of observations.

[35]"Excessive tardiness" and "disruptive behavior" are from "other" category, which account for more than 53 percent of offense cases of the category.

[36]The percentages also show that the sample includes a large portion of offenses that were assigned less severe punishments than out-of-school suspension.

According to the data reporting requirement, almost all types of offenses in the “violence,” “drug” or “property” categories must be reported regardless of the consequences assigned. However, one concern is whether other minor offenses, such as from the category “truancy,” are well reported. To get a sense of the reporting requirement, Column 5 lists the percent of schools with at least one reported offense in the academic year for each type of offense. The percentages suggest that “violence” and “disrespect” offenses are widely reported in many schools.[37] Furthermore, although schools were required to report most of the “drug” and “property” offenses, there was a relative low percentage of schools that actually reported them. This low percentage is likely due to lower offense or catching rates. Note that some less severe offenses, such as “disruptive behavior,” are also widely reported. For other less severe offenses, such as “truancy” and “excessive tardiness,” the percentages are relative low. A further check by different school levels shows that the reporting rate of these offenses changes largely by school levels. More than 84 percent of middle school or high school school-year observations reported at least one “truancy” case, and more than 50 percent of middle school or high school school-year observations reported at least one excessive tardiness case. However, it is likely that many minor offenses were not reported. Therefore, in addition to the econometric effort I will make to address this issue, results from these offenses should be interpreted with caution.

To further illustrate the data, Table 1.2 separately reports sample means of student characteristics for the student-year observations with and without any offense record in the academic year. For the observations with any offense record, I further divide them into those who were not punished by out-of-school suspension in the academic year and those who were punished by out-of-school suspension ever in the academic year. The table shows that the three groups differ considerably along all the dimensions, which indicates possible selection of students into the offender group and into the out-of-school suspended group. While the ratios of white students decrease from the no offense group to the offender group, and from the not out-of-school suspended offender group to

[37] Since, intuitively, the likelihood of being caught committing these offenses should be high, I use these offenses to construct my main instrumental variable.

the out-of-school suspended group, the ratios of black students differ in these groups in an opposite direction. Black student-year observations account for 23 percent of the no offense group and 33 percent of the not out-of-school suspended offender group, but they account for 51 percent of the out-of-school suspended group, which indicates the over-representation of black students in out-of-school suspension rolls. Female students make up a smaller percentage of the offender group and an even smaller percentage of the out-of-school suspended group. Economically disadvantaged students account for a higher percentage of the offender group and an even larger percentage of the out-of-school suspended group.[38] The offender group and the out-of-school suspended group were more likely to be physically or intellectually disabled, less likely to be academically and intellectually gifted, and more likely to be above typical age in the grade or be repeating the grade in the academic year than the no offense group. They also had lower lagged test scores. Grade 9 was the grade with the highest offense and out-of-school suspension percentages.

1.3 Empirical Framework

In this section, I introduce the empirical framework used to estimate the "general deterrence effects," the "specific deterrence effects," and the effects of exclusionary school discipline on students' achievement outcomes. I also discuss the motivation for the empirical models and the econometric issues that must be addressed in order to obtain causal effects.

[38]Economically disadvantaged students are students receiving free or reduced price meals. Eligibility for free or reduced lunch is determined by family size and family income. The most recent criteria can be found on http://www.dpi.state.nc.us/newsroom/news/2015-16/20150814-01.

1.3.1 The Model for Students' Behavior

Let D_{1ist} (D_{2ist}) indicate whether or not student i in school s committed a first (or second) offense in academic year t. I assume the following linear-in-the-parameters model:

$$D_{1ist} = \beta_{10} + \alpha_{11}P^*_{1ist} + \beta_{11}D_{ist-1} + \beta_{12}G_{ist-1} + \beta_{13}X^{std}_{ist} + \beta_{14}X^{sch}_{st} + \beta_{15}X^{std}_{-ist} + \phi^{sch}_{1s} + \epsilon_{1ist} \tag{1.1}$$

$$\begin{aligned} D_{2ist} = {} & \beta_{20} + \alpha_{21}P^*_{2ist} + \alpha_{22}P_{1ist} + \beta_{21}D_{ist-1} + \beta_{22}G_{ist-1} + \beta_{23}X^{std}_{ist} + \beta_{24}X^{sch}_{st} + \beta_{25}X^{std}_{-ist} \\ & + \beta_{26}D^{type}_{1ist} + \beta_{27}N_{1ist} + \phi^{sch}_{2s} + \vartheta_{1ist} + \epsilon_{2ist} \quad \text{if} \quad D_{1ist} = 1 \end{aligned} \tag{1.2}$$

where P^*_{1ist} (P^*_{2ist}) is the potential punishment for the student if she commits the first (second) offense; P_{1ist} is the actual punishment (out-of-school suspension or not) received for her first offense (i.e., if $D_{1ist} = 1$); D_{ist-1} is a vector that captures a student's misbehavior in academic year $t-1$;[39] G_{ist-1} denotes the student's test scores in academic year $t-1$;[40] X^{std}_{ist} is a vector of student's observed characteristics (see variables in Table 1.2); X^{sch}_{st} is a vector of time-varying school observables (see variables in Table 1.3); X^{std}_{-ist} is a vector of the student's peers' observed characteristics;[41] D^{type}_{1ist} is the type of her first offense; and N_{1ist} is a variable that captures the remaining number of in-school days for the student in the academic year after her first offense. ϕ^{sch}_s represents school time-invariant factors, which are not observed by the researcher. ϑ_{1ist} represents

[39] For these variables, I use the student's offense frequencies of each misbehavior category (twelve categories) in the previous academic year.

[40] I use the student's lagged math and reading scores from end-of-grade tests (grades 3-8) or end-of-course tests (grades 9-12) for these variables.

[41] I use same-grade peers' characteristics in estimation. The variables include the ratios of black students (to the whole student population), other minority students, female students, exceptional students, AIG students, students with limited English proficiency, students who repeated grades this academic year, students who are above the typical age in the grade, and economically disadvantaged students, and the means of peers' last year math standard scores and peers' last year reading standard scores.

unobserved (by the researcher) shocks that were not observed by the student when she chose her behavior (D_{1ist}), but were observed by administrators when they assigned the punishment (P_{1ist}); for example, it could be the realized severity of the student's first offense that was not expected by her when she made the first offense decision. The error term ϵ_{ist} captures the student's unobservable (by the researcher) characteristics or other environmental shocks (or misbehavior opportunities) that are observed by the students when they make offense decisions and by administrators when they assign punishments.

The model aims to recover consistent estimates of α_{11}, α_{21}, and α_{22}. The coefficients α_{11} and α_{21} measure the average "general deterrence effects," which describe the effects that result from the threat of a punishment for misbehavior. The coefficient α_{22} measures the average "specific deterrence effect," which describes the role of a previous punishment (as one's punishment experience) on re-offending.[42] The equation 1.1 captures the general deterrence effect of out-of-school suspension on "ever misbehaving or not" in the academic year, and the equation 1.2 isolates the general and specific deterrence effects on "recidivism or not" in the academic year.[43]

The statistical model can be regarded as an approximation to an economic model of a student's decision, in which the student chooses to commit an offense or not (D_{ist}) based on her "tastes" for current behaviors and her expectation on potential punishments.[44] Her "tastes" for current behaviors are determined by her past behaviors (D_{ist-1}, D_{1ist}^{type}, ϑ_{1ist}) through habit formation (or "criminal capital" accumulation), the punishment for her past offense (P_{1ist}) through learning, her past grades (G_{ist-1}), other observables (X_{ist}^{std}, X_{st}^{sch}, X_{-ist}^{std}), and student- and school-level unobservables (ϵ_{ist}, ϕ_s^{sch}). Theoretically, the effectiveness of the potential punishments (the size of the "general deterrence effects") will depend on students' knowledge of or beliefs about potential

[42] While beyond the scope of this analysis, the model can be expanded to include correlated random coefficients (in which students sort into suspension status on the basis of heterogeneous gains from the suspension).

[43] As a robustness check, I also estimate a specification that defines the outcome variables as the student's first (second) offense in middle school or high school, instead of the first (second) offense in an academic year.

[44] A forward-looking student may consider the punishments for the second or further offenses when she chooses the first offense. Therefore, an interesting question for a further study would be how the potential punishment for the second or further offenses deter students' first offense choices.

punishments, their forward-looking ability, their expected probability of being caught, and their expected utility loss from punishments (e.g., physical or emotional unhappiness directly due to out-of-school suspension or indirectly due to consequences of out-of-school suspension).[45] For example, potential punishments may not effectively deter students from committing offenses if students are fully myopic, or they do not know (or do not correctly understand) the disciplinary rules, or they do not care about the consequences of out-of-school suspension.[46]

The variable, "remaining number of in-school days" (N_{1ist}), is calculated by subtracting the days of out-of-school suspension for the first punishment from the total remaining days in the academic year after her first offense. Inclusion of this variable addresses a potential concern when estimating the specific deterrence effect. That is, if a student was out-of-school suspended (or expelled) for the first offense, then she would not be able to re-offend (in-school) during her out-of-school suspension. The "incapacitation effect" might confound the "specific deterrence effect" because a lower re-offending likelihood may be due to less available time for another offense during the academic year and not due to the "wake-up call" effect of the suspension.[47] Using the additional variable to control for the "incapacitation effect," I identify the "specific deterrence effect" by comparing students who have the same remaining number of days to commit the second offense.

For these equations, I also separately estimate the effects by categories (or types) of misbehavior, where for category c misbehavior, for example, the outcomes are defined as "ever committing a category c misbehavior or not in the academic year," and "re-offending a category c misbehavior or not in the academic year." Punishments (or potential punishments) for these specifications

[45]To discuss the policy-relevant effects, I define the deterrence effects as the effects of objective punishments. See Apel (2013) for a detailed discussion of perceptual-deterrence and actual deterrence.

[46]They may even prefer to be punished (i.e., the general deterrence effect is with an opposite sign) if out-of-school suspension provides greater relative enjoyment than staying in the classroom (studying or not studying). This phenomenon is described as "negative reinforcement" in the education literature, which describes that out-of-school suspension may offer students an incentive to misbehave.

[47]A discussion on the relationship between the "specific deterrence effect" and the "incapacitation effect" may be found in, for example, Ehrlich (1981), Marvell and Moody (1994), Tauchen, Witte, and Griesinger (1993) and Levitt (1998).

are defined by the punishments for the corresponding categories.[48] If we regard punishments as "prices," these specifications assume different categories of misbehaviors (demand) operate in different markets and are priced separately.[49] Therefore, an interesting research question that is not explored in this essay is the substitution patterns of general (and specific) deterrence effects (cross-price elasticities) among different types of misbehaviors; for example, whether harsher punishment for violent behaviors would deter students from truancy.

Note that potential punishment variables (P^*_{1ist}, P^*_{2ist}) are not directly observed in the data. While ex-post realized punishments might be regarded as a good approximation for ex-ante potential punishments for those who misbehaved (and have records of resulting punishments), ex-ante potential punishments for those who did not misbehave need to be carefully defined.[50] To address this issue, I use variables $\tilde{P}_{1st}$ and $\tilde{P}_{2st}$, to approximate P^*_{1ist} and P^*_{2ist}, and to define potential punishments for all students (misbehaved and not misbehaved).[51] For misbehavior category c, for example, $\tilde{P}_{1st}$ ($\tilde{P}_{2st}$) is defined by a "normalized" (by offense types) rate of out-of-school suspension assigned to the first offense (the second offense) of c category misbehavior of all the students in school s in academic year t.[52] I refer to these variables as "Disciplinary Punishment Indexes

[48]For these specifications, I include the number of her offenses in other categories before the first offense in the given category as controls for equation 1.2. I also include her same-grade peers' offense rates of other offense categories in the current academic year as controls for both equations 1.1 and 1.2. I carefully choose to not control for the peers' offense rates of the same offense category model in order to allow the general deterrence effect for violent behavior, for example, to include both its direct effects on a students' violent behavior and its indirect effects through the student's peers' violent behavior. I control for offense rates of other categories because one concern for my IV estimates (introduced later) by category of offense is that the IV may change the discipline decisions for other categories as well. This response might be another channel that through which the IV affects the misbehavior outcomes in the discussed category if the discipline rule for the category is not representative for other categories. Robustness checks are done with controls for peers' offense rates in all offense categories to achieve the direct effects (not through peer effects) of the discipline. Robustness checks are also done without controlling for peers' misbehavior rates to address the concern that controlling for the peers' offense rate in other categories may cause an additional endogeneity problem for estimating the deterrence effects.

[49]Cross-market criminal capital accumulation is allowed since past offenses of all categories are controlled for, although cross-market specific deterrence effects are not studied.

[50]Punishments for those who misbehaved but were not caught or reported by the administrators are also not observed.

[51]The reason for using these variables to define potential punishments for all students but not only for non-misbehaved students is to avoid non-classical measurement error problems that may confound estimation results.

[52]I use current year (instead of past year) punishment decisions to construct the proxy variables because punishments

(DPI)" of the school. I discuss the details of their construction in section 1.4.2.

To recover the general and specific deterrence effects, an important issue is the endogeneity of the punishment variables caused by principals' decisions (P_{1ist}, $\tilde{P}_{1st}$ and $\tilde{P}_{2st}$); that is, principals' punishment decisions may depend on unobservables in equation 1.1 and 1.2.[53] A related concern for recovering the specific deterrence effect is the selection of students into committing a first offense ($D_{1ist} = 1$) (the "selection issue"). From equation 1.1, we know that if $\alpha_{11} \neq 0$, we have $Cov(\tilde{P}_{1st}, \epsilon_{1ist} | D_{1ist} = 1) \neq 0$. Therefore, since P_{1ist} is correlated with $\tilde{P}_{1st}$ by construction, we might have $Cov(P_{1ist}, \epsilon_{2ist} | D_{1ist} = 1) \neq 0$ if $Cov(\epsilon_{1ist}, \epsilon_{2ist}) \neq 0$.[54] Additional concerns include: student misbehaviors are not observed if they were not caught or reported and endogeneity of past behaviors, test scores, and remaining number of days variables (D_{ist-1} G_{ist-1}, N_{1ist}) in these equations, which might affect identification of deterrence effects. I discuss how I address these concerns in the subsequent subsections.

1.3.2 Administrators' Punishments and Identification Strategies

A principal or an assistant principal's punishment rules or decisions may depend on misbehavior frequencies and severity and other observable or unobservable characteristics of students or schools since these factors may affect their beliefs on the optimal rules or decisions that maximize the welfare of students and other stakeholders. In addition, it is also possible that some principals' punishment decisions are affected by their prejudice toward students with particular observed or unobserved characteristics. The inclusion of the comprehensive set of controls in the student misbehavior equations (D_{ist-1}, G_{ist-1}, X^{std}_{ist}, X^{sch}_{st}, X^{std}_{-ist}, D^{type}_{1ist}, ϕ^{sch}_s) partially addresses the concern

assigned for the current year offending students should exactly be the potential punishments they would get when they make offense decisions, and these punishments should also better reflect the current year disciplinary policy for other students. The idea of not using past year data to construct the measure also follows the spirit of the "Lucas Critique." The problem in this context could be, for example, if the ineffectiveness of past year lenient punishments had led the principals to change to a harsher discipline rule in new academic year, which was announced to (or expected by) students and changed their behavior, using past punishments would lead to a false conclusion that the lenient discipline changed students' behavior.

[53] The issue is similar for P^*_{ist} and $\tilde{P}_{st}$, assuming that $\tilde{P}_{st}$ is a good approximation for P^*_{ist}. Therefore, I do not distinguish them in the discussion that follows.

[54] The selection problem might also be an issue for estimating the general deterrence effect in equation 1.2 if forward-looking students consider the punishment for the second offenses as well when they choose to initially offend.

that an administrator's punishment decisions depend on the observables and time-invariant school unobservables. However, it does not address the problem that an administrator's punishment decisions may also depend on other unobservables; that is, P_{1ist}, $\tilde{P}_{1st}$, and $\tilde{P}_{2st}$ may be correlated with ϑ_{1ist}, ϵ_{1ist}, and ϵ_{2ist}.[55] The first empirical strategy I use to address this issue is an instrumental variable (IV) method. I specify the following reduced-form equations for the punishment variables:

$$\begin{aligned}\tilde{P}_{1st} = \gamma_{10} + \gamma_{11} Z_{1st} + \gamma_{12} D_{ist-1} + \gamma_{13} G_{ist-1} + \gamma_{14} X_{ist}^{std} + \gamma_{15} X_{st}^{sch} + \gamma_{16} X_{-ist}^{std} \\ + \phi_{1s}^{p,sch} + \epsilon_{1st}^{p}\end{aligned} \tag{1.3}$$

$$\begin{aligned}\tilde{P}_{2st} = \gamma_{20} + \gamma_{21} Z_{1st} + \gamma_{22} Z_{2st} + \gamma_{23} D_{ist-1} + \gamma_{24} G_{ist-1} + \gamma_{25} X_{ist}^{std} + \gamma_{26} X_{st}^{sch} + \gamma_{27} X_{-ist}^{std} \\ + \gamma_{28} D_{1ist}^{type} + \gamma_{29} N_{1ist} + \phi_{2s}^{p,sch} + \epsilon_{2st}^{p} \quad \text{if} \quad D_{1ist} = 1\end{aligned} \tag{1.4}$$

$$\begin{aligned}P_{1ist} = \gamma_{30} + \gamma_{31} Z_{1st} + \gamma_{32} Z_{2st} + \gamma_{33} D_{ist-1} + \gamma_{34} G_{ist-1} + \gamma_{35} X_{ist}^{std} + \gamma_{36} X_{st}^{sch} + \gamma_{37} X_{-ist}^{std} \\ + \gamma_{38} D_{1ist}^{type} + \gamma_{39} N_{1ist} + \phi_{3s}^{p,sch} + \epsilon_{3ist}^{p} \quad \text{if} \quad D_{1ist} = 1\end{aligned} \tag{1.5}$$

where the instrumental variables (Z_{1st}, Z_{2st}) describe the out-of-school suspension propensity for the first or second offense of a principal team, which are constructed as the "normalized" out-of-school suspension rates in other schools in which the team members worked. That is, for each principal team (i.e., principals and assistant principals) in each school in each academic year, if there were members who had principal experience in other schools, I calculate a "normalized" out-of-school suspension rate for these schools. I use it to construct the IV for suspension decisions in the school of concern.[56] The underlying assumption is that principals who were more likely to use out-of-school suspensions in other schools might be more likely to use them in the current school because it may reflect their preferences or beliefs about the effectiveness of out-of-school suspension (or harsher discipline rules). I do not include any out-of-school suspension decisions

[55] Recall that $\tilde{P}_{1st}$ and $\tilde{P}_{2st}$ are constructed by actual punishments P_{1ist} and P_{2ist} for offending students in school s; therefore, they might be correlated with the individual level observables or unobservables.

[56] The IV is defined at the school level but not at the principal level because for each case I do not observe which principal or assistant principal assigned the consequence.

(for any types of offenses in any academic years) in the current school in construction of the IV because they may have direct effects on the students' outcomes in the current school, which would violate the IV assumption. Additional detail about construction of the IV is presented in section 1.4.1. The idea of constructing the IV by using the "punishment propensity" of people is similar in spirit to Kling (2006) and Aizer and Doyle (2015) for the identification of causal effects of incarceration, and Doyle (2007) for the identification of causal effects of foster care.

Suppose $\epsilon^*_{2ist} = \vartheta_{1ist} + \epsilon_{2ist}$, the IV strategy is supposed to address the problem that $Cov(\epsilon^p_{1st}, \epsilon_{1ist}) \neq 0$, $Cov(\epsilon^p_{2st}, \epsilon^*_{2ist}|D_{1ist} = 1) \neq 0$ and $Cov(\epsilon^p_{3st}, \epsilon^*_{2ist}|D_{1ist} = 1) \neq 0$, which is due to the endogeneity or selection issues discussed before. However, the problem in IV estimation is that, if selection plays a role, the exclusion restriction assumptions for the IV might not be satisfied. The exclusion restriction assumptions are $Cov(Z_{1st}, \epsilon_{1ist}|\Omega_{1ist}, \phi^{sch}_s) = 0$, $Cov(Z_{1st}, \epsilon^*_{2ist}|D_{1ist} = 1, \Omega_{2ist}, \phi^{sch}_s) = 0$ and $Cov(Z_{2st}, \epsilon^*_{2ist}|D_{1ist} = 1, \Omega_{2ist}, \phi^{sch}_s) = 0$, where Ω_{1ist} (Ω_{2ist}) is the vector of control variables, which include G_{ist-1}, D_{ist-1}, X^{std}_{ist}, X^{sch}_{st}, X^{std}_{-ist} (and D^{type}_{1ist}, N_{1ist}). However, from equation 1.1, we know that if $\alpha_{11} \neq 0$ and $\gamma_{11} \neq 0$, we have $Cov(Z_{1st}, \epsilon_{1ist}|D_{1ist} = 1, \Omega_{1ist}, \phi^{sch}_s) \neq 0$. Therefore, if $Cov(\epsilon_{1ist}, \epsilon^*_{2ist}) \neq 0$, the assumption $Cov(Z_{1st}, \epsilon^*_{2ist}|D_{1ist} = 1, \Omega_{2ist}, \phi^{sch}_s) = 0$ might not be satisfied.[57] That is, if students react to the threat of punishment when they make first offense decisions, principals with different out-of-school suspension propensities may end up with students with different unobserved characteristics in the offending group ($D_{1ist} = 1$) and, if these unobserved characteristics affect students' recidivism decisions, the IV assumption would not be satisfied for equation 1.2.

To address this problem, I propose an empirical strategy that combines the IV method with a method that recovers student unobserved heterogeneity. The method that addresses student unobserved heterogeneity follows the recent "two-step grouped-fixed effects" literature (Lin and Ng 2012; Bonhomme and Manresa 2015; Bonhomme et al. 2016b). Specifically, I assume that there are a relatively small number of distinct groups of students. The subscript $g_i \in \{1, \ldots, G\}$ represents the student i's group membership, which is estimated from the data, and is assumed to

[57]This argument is not based on correlation properties; it is based on underlying intuition.

not change across time (and across misbehavior categories or types when estimation is for different categories or types of misbehaviors). Using $n = 1, 2$ to represent the equation 1.1 and 1.2, I decompose the error terms ϵ_{nist} as follows:

$$\epsilon_{nist} = \theta_{ng_ist} + v_{nist} \tag{1.6}$$

where the component θ_{ng_ist} captures student i's unobserved heterogeneity.[58] One way to interpret θ_{ng_ist} is to regard it as a product of some permanent student unobserved characteristics and the coefficient on these characteristics (the effects of the student unobserved characteristics); that is, $\theta_{ng_ist} = \rho_{nt}\theta_{g_is}$. I assume that the component θ_{ng_ist} absorbs all the correlation between ϵ_{1ist} and ϵ^*_{2ist}. That is, I assume that $Cov(v_{1ist}, \vartheta_{1ist} + v_{2ist}) = 0$, where v_{nist} may capture, for example, transitory environmental shocks.

Modeling the student unobserved heterogeneity may help to addresses the endogeneity of punishment variables and the selection problem discussed above. In addition, it may also help to address endogeneity of the student's past offense history (D_{ist-1}), past test scores (G_{ist-1}) and "the remaining days variable" (N_{1ist}) for estimating both equation 1.1 and 1.2, if the endogeneity bias is assumed to be due to the correlation between the identified student unobserved heterogeneity and these student-level explanatory variables (i.e., if we assume D_{ist-1} and G_{ist-1} are not correlated with v_{1ist} or v_{2ist}, and we assume N_{1ist} is not correlated with v_{2ist}).[59] I use this empirical strategy to estimate equations 1.1 and 1.2.

The estimation proceeds in two steps. The first step is to partition all students into G groups. Partitions are determined by an application of the k-means clustering algorithm, which is widely used in machine learning and other related fields (Forgy 1965; Steinley 2006). In the second step, the group-specific unobserved heterogeneity (θ_{ng_ist}) is estimated with equations 1.1 and 1.2

[58] When the estimation is for different categories of types of misbehaviors, the θ_{ng_ist} term is also different for different misbehavior categories (or types), i.e., it is $\theta^c_{ng_ist}$, where c represents a specific category (or type).

[59] D_{ist-1}, G_{ist-1}, N_{1ist} are also assumed to be uncorrelated with ϑ_{1ist}. In reality, N_{1ist} might be correlated with ϑ_{1ist}, which is not addressed in this essay.

by imputing the estimated group memberships of students, $\hat{g}_i$. This step amounts to adding the group membership indicators of students into the new estimation. Details about this procedure are discussed in section 1.4.3.

While I show results using the "two-step grouped-fixed effects" without the IV strategy, my preferred model combines the IV strategy with the "grouped-fixed effects" of students. The IV strategy is important for estimating the general and specific deterrence effects even after conditioning on the student unobserved heterogeneity: $\tilde{P}_{1st}$ ($\tilde{P}_{2st}$) might be correlated with v_{1ist} (v_{2ist}) and P_{1ist} might be correlated with ϑ_{1ist} (i.e., principals' punishment rules or decisions might depend on environmental shocks and the severity of students' offenses).[60] After conditioning on $\theta_{n\hat{g}_i st}$, the exclusion restriction assumptions for the IV become $Cov(Z_{1st}, v_{1ist}|\Omega_{1ist}, \phi_s^{sch}, \theta_{\hat{g}_i st}) = 0$, $Cov(Z_{1st}, \vartheta_{1ist} + v_{2ist}|D_{1ist} = 1, \Omega_{2ist}, \phi_s^{sch}, \theta_{\hat{g}_i st}) = 0$, and $Cov(Z_{2st}, \vartheta_{1ist} + v_{2ist}|D_{1ist} = 1, \Omega_{2ist}, \phi_s^{sch}, \theta_{\hat{g}_i st}) = 0$.

The information I use to identify the unobserved heterogeneity is students' misbehavior decisions for different types of offenses across all academic years ($D_{is} = \{D_{is1}, D_{is2}, ..., D_{is\bar{t}}\}$) conditional on observables and school fixed effects. I assume that the student unobserved group is reflected by her misbehavior decisions across time and across different types of misbehaviors. Conditional on all observed factors, students who have different unobserved characteristics, for example, might more (or less) frequently commit offenses across different types of misbehaviors or across time although the punishment rules change across these dimensions.[61] Following Assumption 4 in (Bonhomme et al. 2016b), I assume that v_{nist} are independent across students i, across time t, across n, and across different types of misbehaviors and assume that the groups of students are large and well-separated in the population. The assumption allows the product of the

[60] Principals may change the severity of overall punishments in response to students' behaviors and the realized offense rate within a school/academic year. This reaction creates a problem, called "reverse causality," for estimating the "general deterrence effects," i.e., observed high (low) rates of out-of-school suspension might be the consequences of, but not the reasons for, high (low) misbehavior rates. "Reverse causality" is another way to interpret the correlation between the punishments and the environmental shocks.

[61] I use this information because, intuitively, it closely reflects the unobserved factors that cause the selection issue discussed above. Note that D_{1ist} or D_{2ist} should not be directly used since $D_{2ist} = 1$ has positive probability only when $D_{1ist} = 1$.

number of students' misbehavior types and the number of academic years to grow polynomially more slowly than the number of students. According to (Bonhomme and Manresa 2015), the assumption can be relaxed to allow weak dependence of v_{nist} across types of offenses, academic years or individuals (i.e., assumptions 1 and 2 in Bonhomme and Manresa (2015)).[62] However, the assumption that there are only a relatively small number of distinct groups of students in terms of unobserved heterogeneity is critical for achieving consistent estimates using this estimation strategy. If the population students' unobserved heterogeneity is a continuous variable, the method will generate biased estimates and a bias reduction method should be applied (Bonhomme, Lamadon, and Manresa 2016a).

1.3.3 Students' Achievement Outcomes

For a student who has committed an offense, a principal needs to make a decision on either out-of-school suspending her (i.e., the treatment) or not. An important effect under consideration for such a decision is whether the treatment would lead to worse achievement outcomes for the offending student. To study such an effect, I estimate the following linear-in-the-parameters models:

$$\begin{aligned} Y_{ist} = {} & \beta_{30} + \alpha_{31} P_{ist} + \beta_{31} D_{ist} + \beta_{32} G_{ist-1} + \beta_{33} X^{std}_{ist} + \beta_{34} X^{sch}_{st} + \beta_{35} X^{std}_{-ist} + \beta_{36} D_{-ist} \\ & + \phi^{sch}_{3s} + \vartheta_{ist} + \epsilon_{3ist} \quad \text{if} \quad D_{1ist} = 1 \end{aligned} \tag{1.7}$$

$$\begin{aligned} Y_{ist} = {} & \beta'_{30} + \alpha'_{31} P_{1ist} + \beta'_{31} D_{ist-1} + \beta'_{32} G_{ist-1} + \beta'_{33} X^{std}_{ist} + \beta'_{34} X^{sch}_{st} + \beta'_{35} X^{std}_{-ist} + \beta'_{36} D_{-ist} \\ & + \beta'_{37} D^{type}_{1ist} + \phi'^{sch}_{3s} + \vartheta'_{1ist} + \epsilon'_{3ist} \quad \text{if} \quad D_{1ist} = 1 \end{aligned} \tag{1.8}$$

where P_{ist} is an indicator of whether the student was ever suspended out-of-school in academic year t and P_{1ist} is an indicator of whether the student was suspended out-of-school for her first offense in academic year t; D_{ist} is a vector of her offense frequencies of each misbehavior category in academic year t; D_{-ist} is a vector of her same-grade peers' average offense frequencies of each

[62] This assumption can capture a theoretical model in which environmental shocks are allowed to be weakly correlated with each other across different "misbehavior markets."

misbehavior category in the current academic year; ϑ'_{1ist} represents unobserved (to the researcher) shocks that were not observed by the student when she chose D_{1ist}, but were observed by administrators when they assigned the punishment P_{1ist}, such as the realized severity of the student's first offense that is not expected by her when she makes the first offense decision. ϑ_{ist} represents, for example, unobserved severity of all her offenses in academic year t.

The equations are estimated using students who have ever committed an offense in academic year t ($D_{1ist} = 1$) since these students are students at risk for receiving out-of-school suspension treatment. An out-of-school suspension may directly affect an student's achievement outcomes because, for example, it may reduce the student's classroom learning time. In addition, it may change the student' subsequent misbehavior (through the specific deterrence effect) or her peers' behaviors and, thus, indirectly affects the student's achievement outcomes by such behavioral changes. The coefficient α_{31} captures the direct effect of out-of-school suspension on offending students' achievement outcomes since students' behaviors (D_{ist}, D_{-ist}) are in the controls. The coefficient α'_{31} captures both the direct and indirect effects of out-of-school suspension for students' first offenses on their achievement outcomes.

Y_{ist} in these equations generally represents different types of achievement outcomes. The outcomes I study include end-of-grade math scores of students (G^{Math}_{ist}), ACT composite scores of students (G^{ACT}_{is}), and an indicator of whether the student dropped out or graduated (Q_{is}). When the outcome variable is end-of-grade math scores of students, G^{Math}_{ist}, the scores for students in grades 3-8 are observed at the end of each academic year and thus the model estimates the short-run effect of the punishment within an academic year. The model has a value added structure since the past outcome (G_{ist-1}) is included and multiple observations across academic years are used for each student. ACT composite scores of students (G^{ACT}_{is}) and the indicator of high school dropout or graduate (Q_{is}) are observed only once for each student. Therefore, I estimate, for example, how the punishment to a student in grade 9 affects her final dropout or graduation from high school or her final ACT score.

Following the discussion in the last subsection, to recover the causal effect, α_{31} (α'_{31}) of the punishment variable P_{ist} (P_{1ist}), we must consider its possible correlation with ϑ_{ist} or ϵ_{3ist} (ϑ'_{1ist} or ϵ'_{3ist}). I use the IV strategy to address this endogeneity problem first. The IV, Z_{st} (Z_{1st}), is the out-of-school suspension propensity for all offenses (the first offenses) of a principal team. However, as discussed before, the selection issue of students who enter into the group $D_{1ist} = 1$ might not be addressed by the IV strategy and causes problems for the validity of the IV. An additional problem is that the explanatory variables D_{ist}, G_{ist-1}, and D_{-ist} are endogenous as well. To address these issues, I decompose the error terms $\epsilon_{3ist} = \theta_{3g_ist} + \upsilon_{3ist}$ and $\epsilon'_{3ist} = \theta'_{3g_ist} + \upsilon'_{3ist}$. The grouped-fixed effects θ_{3g_ist} (θ_{3g_ist}) are estimated with other parts of the model in the second step after students' group membership (g_i) is estimated in the first step.[63] My preferred model combines both the IV strategy with the grouped-fixed effects, which allows P_{ist} (P_{1ist}) to be correlated with ϑ_{ist} (ϑ'_{1ist}) even after conditioning on the grouped-fixed effects.[64]

Another important question that the existing literature attempts to answer is how exclusionary school discipline affects "well-behaved" students' achievement outcomes (Perry and Morris 2014). In practice, since exclusionary school discipline removes offending students from the school, this social exclusion prevents (or incapacitates) offenders from committing additional offenses in school for a period of time ("the incapacitation effect"). In addition, the exclusionary school discipline may deter potential offending students from committing offenses or change offending students' subsequent behaviors as discussed before ("general and specific deterrence effects"). Therefore, if students' misbehaviors are harmful to peers' achievement outcomes (i.e., they have harmful "spillover effects"), exclusionary school discipline may lead to better average achievement outcomes for "well-behaved" students if the "general and specific deterrence effects" and the "incapacitation effect" reduce these harmful "spillover effects" of offending students. To

[63]The outcomes used to estimate the group membership follow the outcomes in the last section (student misbehavior frequencies of each type of misbehavior in each academic year). This choice reflects the assumption that the selection problem is caused by some unobservables that affect students' misbehavior choices. An alternative approach could use student misbehavior frequencies and other outcomes, such as student test scores.

[64]In addition to the outlined econometric concerns, selection plagues estimation of the discipline effects in all equations discussed in this section when the outcomes are dropout indicator or ACT scores. I discuss this issue in section 1.5 when presenting the results. I also discuss other limitations in section 1.6.

answer this question, I define a "well-behaved" student as a student who has never had an offense record during all (observed) academic years, and use the notation $D_{1i} = 0$ to represent this group.[65] I estimate the following model:

$$
\begin{aligned}
Y_{ist} = \beta_{40} + \alpha_{41} P^*_{ist} + \beta_{42} G_{ist-1} + \beta_{43} X^{std}_{ist} + \beta_{44} X^{sch}_{st} + \beta_{45} X^{std}_{-ist} \\
+ \phi^{sch}_{4s} + \epsilon_{4ist} \quad \text{if} \quad D_{1i} = 0
\end{aligned}
\tag{1.9}
$$

where α_{41} captures the effect of the harshness of exclusionary school discipline on the achievement outcomes of students who have never had an offense record. The harshness of exclusionary school discipline in school s (P^*_{ist}) is approximated by the "DPI" measure $\tilde{P}_{st}$, which is defined by a "normalized" rate of out-of-school suspension assigned to all offenses in school s in academic year t.

Since harsher exclusionary school discipline may benefit some students and harm other students, one important concern among principals and policy-makers may be the total effects of the disciplinary rules on achievement outcomes of all students. To evaluate the total effect, I estimate the following model:

$$
\begin{aligned}
Y_{ist} = \beta_{50} + \alpha_{51} P^*_{ist} + \beta_{51} D_{ist-1} + \beta_{52} G_{ist-1} + \beta_{53} X^{std}_{ist} + \beta_{54} X^{sch}_{st} + \beta_{55} X^{std}_{-ist} \\
+ \phi^{sch}_{5s} + \epsilon_{5ist}
\end{aligned}
\tag{1.10}
$$

where α_{51} captures the total effect of the harshness of exclusionary school discipline in a school on the achievement outcomes of all students in the school.

The problems for recovering consistent estimates of α_{41} and α_{51} are similar as discussed before. For example, the potential punishment variable (or its proxy variable), P^*_{ist} ($\tilde{P}_{st}$) may be correlated with ϵ_{4ist} and ϵ_{5ist} since punishments may depend on student unobservables or environmental shocks (that may affect "well-behaved" students' achievement outcomes). In addition,

[65] Alternatively, a "well-behaved" student in an academic year t can be defined as a student who has never committed an offense in the academic year t. However, by this definition, estimation of the effects may suffer from a more severe selection of students.

variables D_{ist-1} and G_{ist-1} may be endogenous in these equations.[66]

I use the same empirical strategies to recover consistent estimates of α_{41} and α_{51}. I show the results (for different achievement outcomes) using an IV strategy, where the IV, Z_{st}, is constructed by the "normalized" out-of-school suspension rates for all offenses in other schools in which the principal team members had worked. I also show the results using two-step grouped-fixed effects estimation. My preferred model combines both the IV strategy and the grouped-fixed effects.[67]

1.3.4 Discussion

The proposed approach that solves the unobserved heterogeneity problem (two-step grouped-fixed effects) is related to, but different from, finite mixture models or other "random effect" estimation models, which rely on assumptions that restrict the correlation between unobserved heterogeneity and the covariates. In contrast, it is in close analogy with fixed effects, which leaves the correlation between the unobserved heterogeneity and other covariates unrestricted (Bonhomme and Manresa 2015; Bonhomme et al. 2016b). Since the unobserved heterogeneity is likely to be correlated with several observables discussed above and to be correlated with the school fixed effects or other school observables due to the non-random sorting of students into schools, a "random effect" style approach to solve these problems would require modeling all these mechanisms and jointly estimating the structural equations. The computation would be infeasible given our sample size and the identification requires distribution assumptions and additional exclusion restrictions. Furthermore, traditional student fixed effects do not work in this context because of the following reasons. First, equations 1.2, 1.7 and 1.8 are estimated only for offending students. Students who do not repeatedly commit offenses across years have no counterfactuals to cancel out the fixed effects and directly adding individual indicators would result in an incidental parameters problem.

[66] Although, in estimating equation 1.4, I use the students who had never committed an offense to minimize the selection issue, these students might still be a selected group that depend on the P^*_{ist}; for example, students who never committed an offense might be different in terms of their unobserved heterogeneity in a school that always uses harsh school discipline compared to a school that always uses lenient school discipline.

[67] Note that for equation 1.4, there is no variation of historical misbehavior outcomes. However, to define students' group memberships, since I use the information of misbehavior outcomes conditional on all observables, the variation still exists for the regression residues that are detailed in section 1.4.3.

Second, the equations include lagged outcomes as explanatory variables, which requires strong assumptions to deal with fixed effects. Third, since the school fixed effects are also included, the transfer of students across schools causes a two-way fixed effect problem.[68] The approach used in this essay solves these problems by reducing the dimensionality of the student fixed effects and creating the necessary counterfactuals among students in the same group.

The approach also effectively uses the information on students' numbers of offenses of different types, and offers a flexibility that allows the unobserved heterogeneity to change across type of offenses and across time. This flexibility also relaxes the assumption for the validity of the IV, which is discussed in section 1.4.1.

A problem I have not discussed yet is that student misbehaviors are not observed in the data if they were not caught or not reported to (or by) the administrators. While it might not be an issue for misbehaviors such as violent offenses because the catching rate should be high and reporting is required according to state and federal statutes and state Board of Education policies, it might be an issue for some minor misbehaviors. In this case, in equations 1.1 and 1.2, for example, when we have $D_{ist} = 1$ observed in the data, the actual event is $D_{ist} = 1$ and $C_{ist} = 1$ and $R_{ist} = 1$, where C refers to be caught and R refers to be reported. Therefore, to recover the marginal effects of interest (the general and specific deterrence effects) in equation 1.1 and 1.2, we need to assume that the unobserved catching or reporting probabilities are not correlated with the punishment variables, $\tilde{P}_{1st}$, $\tilde{P}_{2st}$ or P_{1ist}, conditional on all other explanatory variables and school fixed effects.[69] The identification strategy adopted in this essay relaxes the assumption, which only requires that the

---

[68]For some of the outcomes, such as high-school dropout or graduate, there are no repeated outcomes for a student to cancel out the traditional fixed effect.

[69]To see this, as an example, suppose we regard P^*_{1st} as either 1 or 0 and ignore the reporting issue to simplify the discussion (the idea is similar with the reporting issue and with continuous P^*_{1st}), we have the following equation: $Pr(D_{ist} = 1 \cap C_{ist} = 1 | P^*_{1st} = 1, \Omega_{1ist}) - Pr(D_{ist} = 1 \cap C_{ist} = 1 | P^*_{1st} = 0, \Omega_{1ist}) = Pr(C_{ist} = 1 | D_{ist} = 1, P^*_{1st} = 1, \Omega_{1ist}) * Pr(D_{ist} = 1 | P^*_{1st} = 1, \Omega_{1ist}) - Pr(C_{ist} = 1 | D_{ist} = 1, P^*_{1st} = 0, \Omega_{1ist}) * Pr(D_{ist} = 1 | P^*_{1st} = 0, \Omega_{1ist})$, where $\Omega_{1ist} = \{D_{ist-1}, G_{ist-1}, X^{std}_{ist}, X^{sch}_{st}, X^{std}_{-ist}, \phi^{sch}_{1s}\}$. Therefore, if $Pr(C_{ist} = 1 | D_{ist} = 1, P^*_{1st}, \Omega_{1ist}) = Pr(C_{ist} = 1 | D_{ist} = 1, \Omega_{1ist})$, we have the above equation equal to $Pr(C_{ist} = 1 | D_{ist} = 1, \Omega_{1ist}) * (Pr(D_{ist} = 1 | P^*_{1st} = 1, \Omega_{1ist}) - Pr(D_{ist} = 1 | P^*_{1st} = 0, \Omega_{1ist}))$, from which we can calculate the general deterrence effect in terms of percent change: $Pr(C_{ist} = 1 | D_{ist} = 1, \Omega_{1ist}) * (Pr(D_{ist} = 1 | P^*_{1st} = 1, \Omega_{1ist}) - Pr(D_{ist} = 1 | P^*_{1st} = 0, \Omega_{1ist}))/Pr(D_{ist} = 1 \cap C_{ist} = 1 | \Omega_{1ist}) = (Pr(D_{ist} = 1 | P^*_{1st} = 1, \Omega_{1ist}) - Pr(D_{ist} = 1 | P^*_{1st} = 0, \Omega_{1ist}))/Pr(D_{ist} = 1 | \Omega_{1ist})$.

unobserved catching or reporting probabilities are not correlated with the IV conditional on the above variables and the student unobserved heterogeneity. There is an additional issue related to this; that is, for some students, the observed first offense might actually be her second or third offense. In this case, the estimated general deterrence effect α_{11} (α_{21}) might capture the weighted effects of general deterrence for the first (second) offense and further offenses.

When the treatment effects are heterogeneous, the IV strategy captures local average treatment effects (LATE) of the specific deterrence on re-offending students' misbehavior or of suspension on offending students' achievement outcomes. Since the IV is capturing the suspension propensity of administrator teams, we can consider a simplified example with two types of administrator teams: lenient and tough. The tough team prefers to use out-of-school suspension more than the lenient team even for the same type of offense committed by identical students in the same school. The differences in re-offending rates and achievement outcomes between these two teams identifies the local average treatment effects (LATE) of the specific deterrence on re-offending or of suspension on offending students' achievement outcomes (Imbens and Angrist 1994). The LATE measures the effects of punishments on students who would not be punished by the lenient team but would be punished by the tough team (the "compliers"). Since my actual IV variable is a continuous variable, my IV strategy identifies a weighted average of the compliers induced by each marginal change in the IV values (Heckman and Vytlacil 2005). An important condition for the identification of the LATE is monotonicity. It suggests that any student who is out-of-school suspended by the lenient team would also be out-of-school suspended by the tough team, and any student who is not out-of-school suspended by the tough team would not be out-of-school suspended by the lenient team. For my continuous IV, the monotonicity should be satisfied for every marginal change of the IV. While the assumption cannot be fully tested, I show some indirect evidence in the next section to check the assumption.

1.4 Estimation Details

1.4.1 Instrument Construction and Validity

For a student in school s in academic year t, the value of the IV is defined as:

$$Z_{st} = \left(\frac{1}{n(J_{st})}\right) \sum_{j \in J_{st}} \left[\left(\frac{1}{\sum_{k \neq s} \sum_{\tau \neq t} \sum_m \sum_r d_{jk\tau} n_{k\tau mr}}\right) \sum_{k \neq s} \sum_{\tau \neq t} d_{jk\tau} \left(\sum_m \sum_r n_{k\tau mr} (\bar{P}_{k\tau mr} - \bar{P}_{mr})\right)\right] \tag{1.11}$$

where j denotes jth principal (or assistant principal); J_{st} is the set of principals in school s in academic year t;[70] $n(J_{st})$ is the number of principals in school s in academic year t; $d_{jk\tau}$ is one if principal j has worked as a principal in school k in academic year τ, and zero otherwise; $n_{k\tau mr}$ is the total number of m type offenses among students' rth offenses in the school k in academic year τ, where rth refers to first, second, or third offenses. The summation $\sum_{k \neq s} \sum_{\tau \neq t} \sum_m \sum_r d_{jk\tau} n_{k\tau mr}$ represents the total number of offenses in other schools in which principal j has worked. Let $\bar{P}_{k\tau mr}$ be the out-of-school suspension rate for the m type misbehavior among students' rth offenses in school k in academic year τ, and $\bar{P}_{mr}$ be the total out-of-school suspension rate for the m type misbehavior among students' rth offenses in the sample. Thus, $\bar{P}_{k\tau mr} - \bar{P}_{mr}$ normalizes the punishments by the most important features (r and m) of offenses and captures the relative harshness of punishments. With this normalization the IV is less likely to reflect the types of offenses that the principals faced in other schools and, thus, more likely to reflect their preferences for or beliefs about the harshness of punishments.[71]

[70] For most of the academic years, I do not observe the exact dates that a principal works in a school. I define principal teams by academic year, which means that principals defined to be on one team could have worked in non-overlapping periods of the academic year. I use at most two principals and at most five assistant principals per school/year. If, in the academic year, there were more principals in the school, I use those who have worked the longest time in the school in the academic year. (I observe an approximate length of time that each principal works in a school in an academic year.)

[71] The IV is analogous to the average of all principal team members' average regression residuals, where the out-of-school suspension indicator for each offense (in other schools) is regressed on the $r \times m$ offense indicators. An alternative approach to construct the IV is to run a regression (normalization) that also includes all other explanatory variables and the school fixed effects to eliminate their effects on the IV. I do not take this approach in order to avoid the "noise" introduced by the functional form of the regression, which weakens the effectiveness of the IV. Therefore, the IV may also reflect the information found in these explanatory variables rather than the principals' "beliefs" or "preferences." Because of this, I include all of these explanatory variables and school fixed effects in the estimation to validate the IV.

As reported in Table 1.1, some categories (or types) of offenses were detected or reported in only a small number of schools. I find that these categories (or types) of offenses weaken the performance of the IV. To achieve better performance, I construct the IV using those data for which the observed punishment is applied to well detected and reported offenses, namely the "violence" and "disrespect" category of offenses.[72] The IV is missing in some schools in some academic years (12 percent of all student-year observations) since there were no principals or assistant principals who had transfer experience. These student-year observations are not used for the corresponding estimation.

In addition to the IV defined above (henceforth called the main IV), I find that instrumental variables constructed for specific categories (or types) of offenses may have better first stage performance for these categories (or types) of offenses. Therefore, I use these instrumental variables for corresponding specifications.[73] For estimating equation 1.2, I use two instrumental variables consisting of the out-of-school suspension rates for only the first offense ($r = 1$) and only the second offense ($r = 2$), respectively.

Conditional on the control variables, the student unobserved heterogeneity and the school fixed effects, estimation requires that the IV affects the student's behaviors or achievement outcomes only through the principal teams' punishment decisions in the current school. The sorting of principals into schools may be a concern. For example, tough principals may be more (or less) likely to be selected into schools with worse qualities or with more disruptive students. Since I control for time-varying unobserved heterogeneity of students, school fixed effects and time-varying observed characteristics of schools, the concern stems from time-varying unobserved school factors that affect both students' misbehavior decisions and the assignment of the principal teams in the current academic year.

[72]The IV is constructed using data from the 2008-2009 academic year to the 2014-2015 academic year. I only use offenses with $r <= 3$ to eliminate outliers.

[73]Using multiple IVs in these situations is typically less effective than using the best one.

Because the concern cannot be directly tested, I explore indirect evidence, such as the correlation between the IV and the time-varying observed characteristics of the school, to determine the magnitude of the problem. I regress the main IV on time-varying observed characteristics of the school and school fixed effects (Table 1.3). The coefficient column shows that there are only three regressors that are statistically significant at the 5 percent significance level. These are "other disciplinary infraction cases" "total number of classroom teachers," and "PBIS Exemplar school."[74] The magnitudes of the coefficients are small. The F-statistics of jointly testing the significance of all the time-varying observed characteristics is 1.55 with p-value 0.0069. The p-value increases to 0.0974 if the test does not include "other disciplinary infraction cases," and "total number of classroom teachers" variables. The correlation between school time-varying observables and the IV suggests that there is no strong evidence that time-varying unobserved characteristics are correlated with the IV. Since the regressors include many school quality measures in the academic year of concern, the result also suggests that it is less likely that an administrator's out-of-school suspension propensity reflects her other abilities that could directly affect students' behavior or achievement outcomes. However, the fact that there are some significant correlations indicates that including the time-varying observed school characteristics is important for the validity of the IV. In robustness checks, I find, for other constructed IVs used in separate estimations for different categories (or types) of misbehaviors, the correlation between the IV and the time-varying observed school characteristics is generally smaller.

I expect the effect of the IV on the out-of-school suspension decision to be positive, since "tougher" principal teams (proxied by their out-of-school suspension rate in other schools) should be more likely to use out-of-school suspension in the current school. As discussed in the last section, the monotonicity assumption of the IV suggests that the effect (of each margin) of the IV on the punishment for each offense should be non-negative. That is, the "tougher" principal team would be more likely (or equally likely) to use out-of-school suspension for any offenses. While

[74]PBIS means positive behavior intervention and support.

the assumption cannot be directly tested, indirect evidence, such as the effects of the IV on the out-of-school suspension decisions for different types of offenses, could be used to infer the plausibility of the assumption. To check the monotonicity assumption and the first stage performance of the IV, I run the following OLS regressions:

$$P_{1ist} = \gamma_0 + \gamma_1 Z_{ist} + \gamma_2 \Omega_{ist} + \phi_s^{sch,\gamma} + \theta^{\gamma}_{\hat{g}_i st} + \epsilon^{\gamma}_{ist} \tag{1.12}$$

where P_{1ist} is the punishment assigned for the first offense of student i in school s in academic year t, and Z_{ist} is the main IV;[75] Ω_{ist} is the vector of control variables, which include G_{ist-1}, D_{ist-1}, X^{std}_{ist}, X^{sch}_{st}, X^{std}_{-ist}.[76] To infer the monotonicity of the IV, I separately run the first stage regression for each type of offense to check the sign of γ_1, which I expect to be positive. Table 1.4 shows the coefficients (γ_1) for each of the regressions with the type of offense listed in the first column. I include only the types of offenses with more than 4000 observations, since all of the coefficients for the types of offenses with less than 4000 observations are not statistically significant. The coefficient column shows that the IV performs well for most type of offenses, especially for the types of offenses in the "violence" and "disrespect" categories. At the 5 percent significance level, the IV has positive effects for twelve types of offenses. There are two types of offenses with negative effects at the five percent significance level. These variables are "excessive tardiness" and "late to class." As discussed in the data section, it is very likely that consequences for these categories were not well reported. The reporting issue might be the reason that the IV has negative effects on the rate of out-of-school suspension in these categories. Therefore, to improve the plausibility of the monotonicity assumption, I do not include students with these two types of offenses in the offender group in the related specifications. I separately estimate the deterrence effect for the "excessive tardiness" offense by using the IV constructed using suspension decisions

[75] I only use the punishment for first offenses because a large part of my first stage regressions only involve the punishment for the first offense. I find that the results are generally consistent by using the first, the second, and the third offenses of students in the academic year.

[76] I also try other specifications of the first stage equations according to the different specifications discussed in the empirical strategy section. The results are consistent.

for "excessive tardiness" offenses only, which has a positive first stage coefficient.[77] In a robustness check, I find evidence that instrumental variables constructed using the same category offenses only have better performance in terms of the monotonicity for these categories. This finding is one motivation for separately estimating the deterrence effects by offense categories.[78]

The last row of the table shows the coefficient, γ_1, from the regression that uses observations of all types of misbehaviors, which informs the first stage regression for some of my specifications.[79] The F-statistic for testing $\gamma_1 = 0$ in the regression is 86, which is well above the rule of thumb for testing weak instruments.[80]

1.4.2 Disciplinary Punishment Index Construction

Potential punishments for misbehaviors (P^*_{1ist}, P^*_{2ist}, P^*_{ist}) are not observed for each student. As mentioned in section 1.3, I use "normalized" rates of out-of-school suspension within a school/year ($\tilde{P}_{1st}$, $\tilde{P}_{2st}$ or $\tilde{P}_{st}$) to approximate them. I refer to these proxies as "Disciplinary Punishment Indexes (DPI)." The DPI for the rth (1st or 2nd) offense of a student in school s in academic year t is defined by:

$$\tilde{P}_{rst} = \frac{1}{\sum_m n_{rmst}}[\sum_m n_{rmst}(\bar{P}_{rmst} - \bar{P}_{rm})] \tag{1.13}$$

where n_{rmst} denotes the number of type m offenses among students' rth offenses in school s in academic year t; $\sum_m n_{rmst}$ calculates the total number of rth offenses in school s in academic year t; $\bar{P}_{rmst}$ is the out-of-school suspension rate for type m offenses among students' rth offenses in school s in academic year t; and $\bar{P}_{rm}$ is the out-of-school suspension rate for m type offenses among rth offenses in the sample. Similar to the construction of the IV, the normalization allows the measure to reflect the severity of the punishments in the school but not the severity of offenses

---

[77] I did not find a valid IV for the "late to class" offenses.

[78] I also do robustness checks for the effects of suspension on achievement outcomes by separately estimating the effects using different categories of misbehaviors.

[79] The estimation is with "type of misbehavior" as an additional control variable. The estimation results only change slightly between including or not including "excessive tardiness" and "late to class" offenses.

[80] Since first stage estimations are different across my different empirical models, I report these first stage results separately with my other estimation results in the following subsections or in the appendix.

in the school.

$\tilde{P}_{st}$ is constructed in an analogous manner by using the out-of-school suspension decisions for all the first, second and third offenses of students in an academic year.[81] The DPI for each category of offense and some selected types of offenses are also constructed by using the out-of-school suspension decisions for same category (or type) offenses of students. In addition, I also separately construct DPI for black and white students by only using the punishments for black and white students, respectively.[82]

The DPI is used as a proxy variable for the potential punishment for each student in the school in the academic year. For the schools without offenses, the DPI is missing. This unobservability suggests that my estimation results for general deterrence effects might not be applicable to the schools with no offense cases in the academic year. The DPI may not be a good proxy if it is constructed by only few observations. Therefore, in most of the estimation specifications, I drop schools with less than five offenses.

Note that $\tilde{P}_{1st}$, $\tilde{P}_{2st}$ and $\tilde{P}_{st}$ approximate the actual punishments for would-be offenders with measurement errors. That is,

$$P^*_{1ist} = \tilde{P}_{1st} + e_{1ist} \tag{1.14}$$

$$P^*_{2ist} = \tilde{P}_{2st} + e_{2ist} \tag{1.15}$$

$$P^*_{ist} = \tilde{P}_{st} + e_{ist} \tag{1.16}$$

To achieve consistent estimates, my preferred model assumes that the instrumental variables are not correlated with the measurement errors (e_{1ist}, e_{2ist}, e_{ist}), conditional on the controls, school fixed effects and student grouped fixed effects.

[81] As discussed in the last section, the offenses of type "excessive tardiness" and "late to class" are not included in this calculation.

[82] The DPI for all offenses or the DPI at the category level typically have a mean close to zero, and standard deviations from 0.2 to 0.4. Note that the range of DPI could be larger than 1, but the DPI values that are outside a range of 1 are rare cases.

1.4.3 Student Unobserved Heterogeneity

As mentioned in section 1.3, I use a two-step approach to capture the student unobserved heterogeneity. The first step is to partition all I students into G groups. In the second step, the group-specific fixed effects are estimated with all other parts of the models by imputing the estimated group memberships of students into the models.

Specifically, in the first step, the partition problem is:

$$\min_{g(1),\ldots,g(I),\theta_1,\ldots,\theta_G} \sum_{i=1}^{I}\sum_{t=1}^{\bar{t}}\sum_{c=1}^{C}(D_{ist}^{c} - \hat{\beta}_t^c \Omega'_{ist} - \hat{\phi}_{st}^{sch,c} - \theta_{g_i st}^{c})^2 \tag{1.17}$$

where $g(i)$ indicates that student i is assigned to group g; later, I use g_i to denote the assigned group of student i. The vector $\theta_g = (\theta_{g1}^1, \ldots, \theta_{g1}^C, \ldots, \theta_{g\bar{t}}^1, \ldots, \theta_{g\bar{t}}^C)$ contains the mean value of each feature in each academic year among students in group g. The mean value of the cth feature in academic year t among students in group g is $\theta_{g_i t}^c$.

The cth feature of student i in academic year t is defined by the expression $D_{ist}^c - \hat{\beta}_t^c \Omega'_{ist} - \hat{\phi}_{st}^{sch,c}$. It requires that I regress the cth outcome of student i in academic year t, D_{ist}^c (the frequency of the cth category offense of student i in academic year t), on a vector of control variables, Ω'_{ist}, and school-year-category fixed effects in order to obtain estimated values, $\hat{\beta}_t^c$ and $\hat{\phi}_{st}^{sch,c}$.[83] The regressions are run separately for each category of offense, c, in each academic year, t. The control variables, Ω'_{ist}, include G_{ist-1}, D_{ist-1}, X_{ist}^{std} and X_{-ist}^{std}.[84] The expression for the cth feature, $D_{ist}^c - \hat{\beta}_t^c \Omega'_{ist} - \hat{\phi}_{st}^{sch,c}$, equals the regression residuals, which are used to partition students.

The regressions are similar to the reduced form of equations 1.1 and 1.2, but the dependent variables (the frequencies of each category of offense of a student within an academic year) aggregate the information of the "first offense" and "second offense" and gains further efficiency

[83] If student i has transfer experience within academic year t, her school, s, is defined by the one in which most of her offense records come from.

[84] School time-varying observables are not included since school time-varying effects are captured by $\hat{\phi}_{st}^{sch,c}$.

by using information on third and more offenses of this student in the academic year.[85] The regression residuals reflect the additional factors that influence a student's misbehaviors (i.e., factors other than observed characteristics of the student and observed and unobserved characteristics of the school). Common patterns among these additional factors are used to identify a student's group membership. That is, they reflect students' unobserved types.

The partition problem is solved by a k-means algorithm that proceeds by alternating between "assignment" and "updating." The "assignment" step assigns group memberships for all students; the "updating" step updates the new mean, θ, of each group. The group membership of student i is solved by:

$$\hat{g}_i = \underset{g \in \{1,\dots,G\}}{\operatorname{argmin}} \sum_{t=1}^{\bar{t}} \sum_{c=1}^{C} (D^c_{ist} - \hat{\beta}^c_t \Omega'_{ist} - \hat{\phi}^{sch,c}_{st} - \theta^c_{gt})^2 \tag{1.18}$$

The partition task is designed to classify students by $C \times \bar{t}$ features captured by the regression residuals. The fact that data are not available for all students in all academic years creates a complication. Only using students with all years of data may create a selection problem if, for example, students who were offenders or who were suspended have fewer years of data. Therefore, I group the students by the number of academic years they contribute, and separately classify students within the same "observed years of data" group into more groups defined by the algorithm. The classification may be not fully efficient because it forces students with different numbers of academic years of data to be classified into different groups. However, following the spirit of Bonhomme and Manresa (2015), the fact that it generates more groups reduces the biases of the common parameters that I estimate.[86]

[85] Note that the school-year-category fixed effect soaks up all information in the IV in the reduced form estimation of equations 1.1 and 1.2.

[86] My sample of students are followed for five academic years. In a robustness check, I find that a partition task that only uses students with three or more academic years of data produces similar final estimation results to the discussed task. Bonhomme et al. (2016b) suggest that, based on a set of assumptions, different moments and specifications could be used as features for the partition when the moments provide information about the student's group membership. Therefore, I could also add, for example, academic achievement outcomes of students as additional features. Bonhomme et al. (2016a) also show that, with bias reduction, the clustering method could achieve satisfying identification without finite group assumption of the heterogeneity.

I classify each sample of students (with the same observed years of data) into ten groups.[87] Since there are five years of data in the sample, I have fifty groups of students. Because the k-means algorithm is sensitive to the choice of initial value, I use 500 random initial values to select the classification with the best performance. About 45 percent of students contribute data all five academic years; I use this sample to illustrate the classification results. In total, there are twelve categories of offenses used for the classification; therefore, for students with five academic years of data, their classification is based on sixty ($C \times \bar{t}$) features.

In Figure 1.1, I show trends of the frequencies of each category of offense across years for each group of students. Each graph represents each of the twelve categories of offenses. Each line color represents each of the ten student groups. Students in different groups are observed to behave quite differently, which shows the efficiency of the classification. The group represented by the yellow line has low numbers for each category of offense, suggesting that it is a rarely misbehaved group of students; in contrast, some other groups of students misbehave frequently across each category of offense or across years. These groups might capture the groups of students that are not sensitive to punishments. In addition, there are groups of students with offense frequencies that fluctuate across academic years (or across offense categories).

1.5 Results

In this section, I present estimation results. In addition to the results from my preferred estimation strategy, I also show results from other specifications that incrementally address the endogeneity and selection issues so that we can learn about the source of any bias.

1.5.1 Results for Deterrence Effects

Table 1.5 reports estimates of α_{11} in equation 1.1 – the "general deterrence effects" for the first offense.[88] While the rows show the results for different categories of misbehavior, the columns compare the results from different estimation strategies. I begin with results for "all offenses" (row

[87] As a robustness check, I also classify each sample of students into five groups or fifteen groups, which shows that the major results are robust to these specifications.

[88] The estimates for other coefficients of equation 1.1 are reported in the appendix.

1), which should be interpreted with caution for the following reasons.[89] First, the estimates may reflect the effects for some types of misbehaviors that mostly contribute to the variations of the DPI.[90] Second, the estimates may reflect the effects for offense types punished by discipline decisions that are mostly affected by the IV.[91] In rows 2-6, I report the results for different categories of misbehavior.[92]

The first method (OLS) is an OLS model with all control variables and school fixed effects. The second method (OLS&GFE) adds controls for student unobserved heterogeneity; I use the term "grouped fixed effects" (GFE) to represent the student unobserved heterogeneity, which follows the terminology used by Bonhomme and Manresa (2015). The estimates from these two methods suggest that the "general deterrence effect" is statistically significant for most of the categories. The OLS&GFE estimates are slightly, but statistically significantly, different from the OLS estimates for "all offenses," "violence" and "disrespect," which shows that the additional controls for student unobserved heterogeneity may address some biases in estimation. However, as discussed before, the bias caused by some types of unobserved factors (e.g., environmental shocks) might not be addressed by these controls. Therefore, I further address these issues by instrumenting for the DPI variable, the key explanatory variable. The third method (2SLS) instruments for the DPI variable, but does not include controls for student unobserved heterogeneity; the fourth method

[89]For "all offense" sample, the key explanatory variable, the DPI, is constructed by the suspension decisions for students' first offenses in an academic year; the offense could be any type of misbehavior. The dependent variable is the indicator that the student committed first offense (any type) in the academic year. As discussed in section 1.4.1, the offense types "excessive tardiness" and "late to class" are not used for constructing the DPI or the offense indicator.

[90]Since the DPI reflects weighted deviations of the schools' suspension rates (for a type of misbehavior) from the average suspension rates in all schools, some serious offenses that all schools assign suspensions to, for example, would contribute little to the variation. In addition, the DPI (for all offenses) for different schools is calculated by different types of misbehaviors due to the different reporting rates, catching rates and offense rates for them across schools, which means that the estimates should be interpreted with caution.

[91]Using "all offenses" is also less likely to satisfy the monotonicity assumption. Although my endogenous variable is continuous, the underlying intuition of monotonicity generalizes to this case.

[92]The DPI measure for a category of misbehavior is constructed by using only the suspension decisions for students' first offenses of the category of misbehavior. The dependent variables are whether or not the student committed the first offense of the category of misbehavior in the academic year.

(2SLS&GFE) adds controls for student unobserved heterogeneity. The first stage (adjusted) F-statistics for these estimates are all well above the rule of the thumb for testing for weak instruments, and are reported in the appendix (Stock and Yogo 2005). Note that 2SLS and 2SLS&GFE estimates all suggest higher "general deterrence effects," which are statistically different from the OLS and OLS&GFE estimates.[93] The results may suggest that 2SLS and 2SLS&GFE estimation methods provide additional reductions in bias. Another possible explanation is that these IV estimates may capture some "local" effects for some types of offenses or some subpopulations of students if the IV only affects the discipline decisions for them. The 2SLS&GFE estimates are different from 2SLS estimates (but the differences are not statistically significant), which shows that including "grouped-fixed effects" may further address some estimation issues.

The last column shows means of the dependent variables, which capture average (reported) offense rates in the estimation sample. By dividing the coefficients by the means, my preferred estimates (from the model 2SLS & GFE) suggest that a 10 percentage point increase in the out-of-school suspension likelihood index could reduce the mean rate of a student committing any type of offense in a year by about 15.6 percent, "violence" offenses by about 13 percent, "disrespect" offenses by about 11.5 percent, "truancy" offenses by about 22 percent, and "drug" offenses by about 18 percent.[94] The estimate is not statistically significant for the "property" category. A possible explanation might be that the low offense rate of this category makes the DPI measure less precise and identification of the effect more difficult (although the empirical framework may address part of the measurement error problem).

---

[93]Since one way to interpret the correlation between the environmental shocks and the punishments is through "reverse causality," and harsher discipline rules are more likely to be used for the school/year with higher offense rates or more severe offenses, intuitively, OLS and OLS&GFE estimates might understate the improvement achieved by harsher discipline rules.

[94]Although I use the average reported offense rate to calculate these percentages, interpretation of them may generalize to percentages of real offense rate reduction (observed and unobserved) under the assumption that my empirical strategy effectively addresses the related econometric issues. As discussed in the data section, the results for the drug category should be interpreted with caution because there might be additional punishments for these types of offenses. A robustness check shows that the estimate is not statistically significant for the drug category after controlling for the same-grade peers' drug offense rate, which captures the direct general deterrence effects (without effects that work through peer effects). The robustness check is also done for other categories.

To further explore the sources of the effects, I estimate α_{11} for different types of misbehaviors, in which the DPI for a type of misbehavior is constructed by using the suspension decisions for the specific offense type only. In Table 1.6, I show the results for types with the highest offense rates because the estimate is typically not statistically significant for types with lower offense rates. The estimates indicate significant "general deterrence effects" for most of the offense types. The changing patterns of estimates across estimation methods are consistent with the patterns in Table 1.5. My preferred estimation method (2SLS&GFE) shows that the "general deterrence effects" are heterogeneous across different types of misbehaviors. They are high for offense type "aggressive behavior," "insubordination," and "inappropriate language" – a 10 percentage point increase in the out-of-school suspension likelihood could reduce the mean offense rate for these types by about 30-40 percent. The percent reduction is about 12 percent for "skipping class," 20 percent for "disruptive behavior," 7.5 percent for "disrespect to faculty," and 16 percent for "excessive tardiness."[95] It is not statistically significant for the "fighting" offense. The heterogeneity in response may stem from the differences in motivation, nature, or characteristics of offenders for these offenses. Another possible explanation is that the heterogeneity may be partly due to differences in average out-of-school suspension rates assigned for these offenses, which are shown in Table 1.1.[96] Finally, these results should be interpreted with caution as the offense rate for each offense type is relatively low, which makes the DPI measure less precise and the identification of the effect more difficult.

To further explore the potential heterogeneity of the general deterrence effect, I present estimates stratified by student observables (Table 1.7). The DPI measures are constructed by using the suspension decisions for the students with the corresponding observables. I present only the results for the well reported "violence" and "disrespect" categories because DPI measures are less precise

[95]The percent reduction is calculated by dividing the coefficients by the means of dependent variables.

[96]A 10 percentage point increase in the suspension rate might mean a much harsher discipline when the current average suspension rate is low; in addition, the reported average suspension rate may also represent the "severity" of these types of offenses, which may suggest that the heterogeneity partly contributes to the severity of offenses. However, the correlation between the average suspension rate and the general deterrence effect is not quite clear.

and the estimates are more likely to suffer from a weak instruments problem after the sample is conditioned on a particular student characteristic.

The differences between OLS&GFE and 2SLS&GFE estimates are consistent with the patterns discussed in Table 1.5 and Table 1.6. My preferred estimates (2SLS&GFE) find statistically significant "general deterrence effects" for most of the student subpopulations, and the effects are heterogeneous. First, I find that "general deterrence effects" are not statistically significant (and with small coefficients) for high school violent behavior, and elementary school disrespectful behavior.[97] Second, I find that the "general deterrence effects" might be higher (in terms of the percent reduction of mean rate) for white students than black students – a 10 percentage point increase in the suspension likelihood index reduces the mean offense rate of "violence" and "disrespect" behaviors for white students by about 16 percent and 24 percent, and by 11 percent and 9 percent for black students. In addition, I find that, in terms of the percent reduction of mean rate, the effects are smaller for female students than male students, smaller for economically disadvantaged students than those not economically disadvantaged, smaller (and not statistically significant) for "violence" offenses of students with lagged math scores below (or equal) the average than those with lagged math scores above the average.[98]

Table 1.8 reports estimates of the general deterrence effect (α_{21}) and the specific deterrence effect (α_{22}) for students' second offenses in an academic year (equation 1.2), which are estimated separately by offense category.[99] The table includes results for the better reported "violence" and "disrespect" categories, and for the "truancy" category. I show the results for other categories or types of offenses in Table A1 in the appendix. Most of the OLS and OLS&GFE estimates

[97] A robustness check by controlling for offense rates of all types of peers' misbehaviors also finds that the effects are relative low (but statistically significant) for the disrespectful behavior of high school students.

[98] The percent reduction is calculated by dividing the coefficients by the means of dependent variables (the fourth number in each cell). Again, the heterogeneity should be interpreted with caution because of the following reasons. First, some of the differences are not statistically significant. Second, some of them are mostly due to the differences in the mean offense rates. Third, the DPI variables may less precisely reflect the punishment severity in the school because they are constructed with less observations.

[99] The DPI variable, $\tilde{P}_{2st}$, is constructed by only using students' second offenses of the category in the academic year.

suggest statistically significant positive "general deterrence effects" for students' second offenses. However, the estimates for α_{22} suggest that the "specific deterrence effect" is negative for "violence" offenses and it is positive but small for "disrespect" and "truancy" offenses. The OLS&GFE estimate suggests that for "disrespect," for example, the suspension experience itself could only reduce the mean rate of the student's re-offending by less than 2 percent. The 2SLS and 2SLS&GFE estimates suggest that both "general deterrence effects" and "specific deterrence effects" are not statistically significant.[100] However, one problem with these estimates is that the collinearity between α_{21} and α_{22} is severe because the IV method reduces the variation of $\tilde{P}_{2st}$ and P_{1ist} in the second stage estimation. This reduced variation inflates the standard errors and the magnitude of these estimates; therefore, the inference is less precise and the results should be interpreted with caution.[101]

I further explore the efficacy of "general deterrence," α_{21}, by estimating equation 1.2 separately for student-observations that were out-of-school suspended for their first offenses ($p_{1ist} = 1$) and those that were not out-of-school suspended for their first offenses ($p_{1ist} = 0$).[102] Table 1.9 shows the estimates of α_{21} using these specifications for the "violence," "disrespect," and "truancy" categories. I show the results for other categories and types of offenses in Table A2 of the appendix.

My preferred estimates (2SLS&GFE) suggest that the "general deterrence effects" are not statistically significant for students who were out-of-school suspended for their first offenses in the academic year. They are also not statistically significant for the "violence" and "truancy" category offenses for students who were not suspended.[103] In addition, I find that the effect is statistically significant for the second "disrespect" offenses of students who were not suspended for their first offenses. However, the estimate suggests a smaller effect compared to the effect for students' first

[100]F-tests also show that they are jointly not statistically significant.

[101]Robustness checks, by instrumenting only one of these endogenous variables, yield similar results.

[102]By this specification, I could estimate the "general deterrence effects" without adding P_{1ist} as an explanatory variable in the regression, which solves the collinearity problem.

[103]The large coefficients and standard errors for the "truancy" category for these estimates are because the first stage performances of the IV are relative weak, with (adjusted) F-statistics of about 7 and 10.

"disrespect" offenses – a 10 percentage point increase in the suspension likelihood index reduces the mean rate of re-offending of "disrespect" offenses by about 8 percent.

1.5.2 Results for the Effects on Achievement Outcomes

Table 1.10 reports estimates of α_{31} and α'_{31}– the effects of out-of-school suspension experience on offending students' achievement outcomes. The outcomes (dependent variables) include end-of-grade math score, a dropout indicator (finally graduate or drop out from high school) and ACT composite score.[104] The "dropout" dependent variable is whether or not the grade 9 student finally drops out from high school, which may happen in any future grades.[105] I estimate how it is affected by the student's suspension experience in grade 9. In robustness checks, I find that the results are robust to the specifications that use students' suspension experience in grade 9-10 or grade 9-12. I mainly focus on interpreting the results for grade 9 because adding student observations in higher grades may create additional dynamic selection problems, which may make interpretation of the results less clear.[106] For estimation using ACT composite scores as the outcome, I use student-year observations in grade 9-10 who have their final ACT composite scores.[107] In a robustness check, I find that the results are similar when I use student observations in grade 9 only.

Columns 2-6 report results using different estimators (for α_{31}) where the key explanatory variable is an indicator of whether or not the student was ever suspended out-of-school during the academic year. To compare my methods/results with those from the exiting literature, I begin with OLS regressions that include controls for school-level observables and school fixed effects only. The results in Column 2 are from a regression using all students (not only offending students);

---

[104]The end-of-grade math scores are normalized to have a zero mean and standard deviation of one among students who took the same tests across the state.

[105]In North Carolina, students can drop out of high school at 16 years old.

[106]The problem is that the offending students in grade 10 who did not drop out at the end of grade 9 may be a selected group, such as a relatively better behaved group in the offending student population. Although by assumption, my preferred estimation strategy may address the problem for estimating α_{31} and α'_{31} for observed offending students in grade 10 (or higher grade), but the concern that they may be the relatively better behaved group, for example, suggests that extrapolating these estimates to all offending students should be done with caution.

[107]The choice is mainly based on the data availability, which was discussed in section 1.2.

it compares the achievement outcomes of students who had experienced suspension to a control group that includes both students who had an offense record but were not suspended and students who did not have an offense record in the academic year. A suspension experience decreases grade 3-8 students' end-of-grade math scores by about 0.2 standard deviations. Suspension experience in grade 9 increases the dropout probability by about 18.5 percentage points. Suspension experience in grade 9-10 also lowers ACT composite scores by about 1.1 points. One problem with interpretation of these effects is that, among students in the control group, those who did not commit an offense are not at risk of suspension. To measure the causal effect of suspension on outcomes, we may desire to focus on students facing the possibility of being suspended. Beginning with Column 3, I limit the estimation sample to students who had an offense record in the academic year and label them "offending students."[108] Column 3 reports the results with the same controls as column 2 but uses the "offending students" sample. The results indicate smaller negative effects of suspension for all three achievement outcome measures, which suggests that using a more representative sample may solve some selection issues. To further reduce the endogeneity and selection biases, I add all student-level controls discussed in section 1.4 and report the results in the "OLS with full controls" column. The results indicate much smaller negative effects of suspension – a suspension experience reduces students' end-of-grade math scores by about 0.013 standard deviations, and ACT composite scores by about 0.21 points. It increases the dropout probability by about 7.2 percentage points. These effects fall even more (to 0.012 standard deviations, 0.17 points, and 5 percentage points) after adding an additional control for student unobserved heterogeneity ("OLS & GFE" column).

When I address the estimation biases by instrumenting for the suspension decision (2SLS and 2SLS&GFE), I find that the effects for all three outcomes are not statistically significant.[109] The

[108]Although the sample of offending students does not include students who committed an offense but did not get a record, it should be more representative of the "at risk for punishment" students than the "all students" sample.

[109]The first stage (adjusted) F-statistics for the IV are 65 and 74 for the 2SLS and 2SLS&GFE estimates of end-of-grade math test scores, 26 for the 2SLS and 2SLS&GFE estimates of dropout indicator, and 17 for the 2SLS and 2SLS&GFE estimates of ACT composite scores.

estimates for end-of-grade math and dropout probability also change signs. The results should be interpreted as the effects of suspension on achievement outcomes of students on the margin of suspension - the compliers. The large standard errors of the estimates also suggest using caution in interpretation. However, the results from all estimates consistently show that the negative correlation between suspension experience and students' end-of-grade math scores or dropout probability become smaller once I further address the endogeneity and selection biases. These findings suggest that the correlation is largely (or completely) explained by the correlation between observed or unobserved factors of students (or offenses) and these achievement outcome measures.

Some additional information suggest that the results for ACT scores should be undertaken with caution. First, since the ACT is required for students in grade 11, offending students who drop out before the test do not generate observed test scores. Therefore, the students who have the scores may be a selected group and extrapolating the estimates to all offending students should be done cautiously. An additional concern is that the suspension decision may induce students with different unobserved factors to drop out at different rates, which may imply that the offending students who have both ACT scores and suspension experience have different unobserved factors than offenders who had ACT scores but did not experience suspension.[110] Because the student unobserved heterogeneity is mostly identified by students' misbehavior information but not dropout decisions (which cannot be repeatedly observed), the unobserved heterogeneity in this dimension might not be captured. However, my estimation for the dropout may indicate that this concern is not severe because the suspension experience might not significantly affect students' dropout probability.

In the last two columns of Table 1.10, I provide the estimates (for α'_{31}) with the key explanatory variable being an indicator of whether the student was suspended for her *first* offense in that academic year.[111] As introduced in section 1.3, the specification controls for students' offense frequencies in the previous academic year instead of in the academic year of concern. These estimates

[110]In this case, the difference of ACT scores between those who have or did not have suspension experience may be a reflection of these unobserved factors but not the "causal effects" of suspension.

[111]I find that the results are robust to only using students who were suspended for their first "violence" and "disrespect" offenses.

may capture both the direct effect and the indirect effect of suspension on achievement outcomes resulting from the offending students' behavior changes in the current academic year. Estimates from the preferred model (2SLS& GFE) suggest that the effects are not statistically significant, which is consistent with the previous findings.[112]

One interesting question is whether the effects of suspension on achievement are different for black and white students. Table 1.11 provides results from models estimated separately for black and white students. Each of the four models (Columns 2-5) uses the full set of control variables and the indicator of whether the student was ever suspended out-of-school during the academic year as the key explanatory variable. The OLS&GFE estimates suggest that the negative effects of suspension on all three achievement outcomes are larger for white students than for black students. The 2SLS and 2SLS&GFE estimates suggest that the effects of suspension experience on end-of-grade math scores and dropout probability are not statistically significant for either black or white students.[113] The 2SLS estimate suggests that the negative effect on white students' ACT scores is about 3.3 points, which is statistically significant at the 10% confidence level. The 2SLS&GFE point estimate shows that the negative effect on white students' ACT scores is about 2.9 points, but it is not statistically significant.[114]

While these results illustrate the effects of suspension experience on offending students' achievement outcomes, I further explore the effects of exclusionary school discipline on achievement outcomes of students with no offense record and the overall effects on all students' achievement.

[112]The large standard errors and point estimates for ACT scores are likely due to a weak instruments problem. The (adjusted) F-statistics of the IV in the first stage regression is only about 3.

[113]The first stage (adjusted) F-statistic for the IV is 50 for black students' end-of-grade math scores, and 7 for white students' end-of-grade math scores. In addition, the first stage (adjusted) F-statistic for the IV is 6 for white students' dropout probability and 13 for black students' dropout probability. The IV is relatively weak for white students; however, I find that the results are robust by using different specifications for white students.

[114] A robustness check using student unobserved heterogeneity that does not change across time suggest the estimate for white students' ACT scores is statistically significant at 10% confidence level. The same robustness checks are done for all other estimates, and show that other results are robust. The large standard error for 2SLS and 2SLS&GFE estimates for black students' ACT scores is because of the weak instruments problem; the first stage (adjusted) F-statistic for the IV is only about 1. The F-statistic for white students' ACT scores are about 22;

As discussed before, these effects may encompass students' and peers' behavior changes resulting from the "general or specific deterrence effects" or from the "incapacitation effect." Table 1.12 reports estimates of α_{41} and α_{51}.[115] Again, I separately report OLS, OLS&GFE, 2SLS, and 2SLS&GFE estimates in different columns.[116] Estimates from my preferred model (2SLS&GFE) show that the effect is not statistically significant for end-of-grade math scores, dropout probabilities and ACT composite scores of "all students." In addition, the effect is not statistically significant for dropout probabilities and ACT composite scores of students with no offense record.[117] However, I find that the effect is positive and statistically significant for the end-of-grade math scores of middle school students with no offense record. A 10 percentage point increase in the suspension likelihood index could increase the end-of-grade math scores of "well behaved" middle school students by about 0.02 standard deviations.

To further explore whether the overall effects are heterogeneous by racial group, Table 1.13 reports estimation results for white and black students by splitting the sample. The 2SLS&GFE estimates suggest that harsher school discipline has a positive overall effect on middle school white students' end-of-grade math scores. A 10 percentage point increase in the suspension likelihood index increases the end-of-grade math scores of middle school white students by about 0.028 standard deviations. I do not find such evidence for middle school black students' end-of-grade math scores.[118]

---

[115] A student with no offense record means that she does not have any offense record in my sample period. I only include students with no offense record at least for three academic years to make sure that they are "well behaved." For the achievement outcome measures "dropout" and "ACT composite scores," I also drop students with any offense record in 2013-2014 and 2014-2015 academic year.

[116] All of the estimates are with a full set of control variables. The key explanatory variable, $\tilde{P}_{st}$, is constructed by using "normalized" suspension rates for students' first, second, and third offenses in the academic year. As discussed before, "excessive tardiness" and "late to class" are not included into this calculation.

[117] The larger negative coefficients for ACT composite scores suggest that one should interpret the results for ACT scores with caution.

[118] The (adjusted) F-statistics for the IV in the first stage regression is about 7 for middle school black students' end-of-grade math scores.

1.6 Conclusions

Exclusionary school discipline techniques are widely criticized for their inability to improve students' behavior and for their adverse effects on students' achievement outcomes. However, I find that disciplinary rules exhibiting a higher out-of-school suspension likelihood could significantly deter students from committing first offenses. Estimates from my preferred estimation method also suggest that the adverse effects of out-of-school suspension experience on offending students' end-of-grade test scores and high school dropout probability are not statistically significant. Moreover, contrary to Perry and Morris (2014), I find that disciplinary rules with higher out-of-school suspension likelihood could improve the end-of-grade math scores of "well behaved" middle school students. The results imply that policies that reduce or remove suspension options from schools should carefully consider these benefits of the disciplinary practice. Particularly, policies that focus on reducing out-of-school suspension rates for minority groups may increase offense rates of these minority groups. The results also suggest that, contrary to the existing literature (e.g., Morris and Perry (2016)), there is no evidence that the suspension disparity between white and black students creates crucial black-white achievement gap.

However, my estimates also suggest that the disciplinary practice is less effective or ineffective for repeat offending students, especially for students who have had suspension experience. Since repeat offenses account for a large portion of all infractions, the results suggest that it is important to find a more effective approach to deal with them. I also find suggestive evidence that the disciplinary practice might be less effective or ineffective for some types of (first) offenses among certain student subpopulations, such as high school students' violent behavior. Therefore, alternative approaches to deal with these offenses are also important.

When interpreting the results, several caveats should be kept in mind. First, when heterogeneous effects are in play, my preferred method may be more likely to capture weighted average effects for "compliers" (the students who were, or would be, punished differently because they were assigned to principal teams with different out-of-school suspension propensities).[119] Nevertheless, positive general deterrence effects are consistently found for different categories (or types) of misbehaviors and for different students' subpopulations. The results also consistently show that the documented negative effects of suspension experience on end-of-grade math scores or dropout probability are largely due to endogeneity or selection.

Second, since students' out-of-school misbehaviors are not observed in my data, the analysis does not capture the potential effects of exclusionary school discipline in that dimension. Therefore, my estimates may understate the costs of suspensions (or overstate the benefits of them) if they lead to increases in these unobserved misbehaviors.[120]

In addition, since I simplify the punishment as "out-of-school suspension" or not and most of the out-of-school suspensions are short-term, the results for the effects of suspension experience on achievement outcomes are more likely to capture the effects from comparing short-term out-of-school suspensions with other less severe punishments. Therefore, the results may not completely capture the achievement loss of students who were suspended for a long time, and they might also not represent the effects of out-of-school suspension compared to no punishment at all.

[119]The effects on these students may be particularly interesting for policy making, as they may be the student group for which different policies suggest different punishments.

[120]One example is that, compared to the alternative punishments (e.g., no punishment, detention, in-school suspension), out-of-school suspension might be more likely to lead students to commit offenses off-campus. Then, my estimation does not capture this part of social costs. If these off-campus offenses substitute the students observed (detected and reported) in-school offenses, my estimates might also overstate the deterrence effects for in-school misbehaviors.

Table 1.1: Summary Statistics for Selected Categories of Offenses

Offense	Total Number of Incidents	Percent of Offending Students who receive OSS for 1st offense	for 2nd offense	Percent of Schools with any reported offense among academic years
Offense Category				
Violence	584,566	63.6	64.3	93.0
Disrespect	902,248	29.4	34.9	86.4
Truancy	343,896	17.8	24.7	48.7
Drug	75,072	59.1	58.1	42.0
Property	63,075	53.1	58.2	71.0
Other	1,981,559	17.2	17.2	91.2
Offense Type				
Fighting	198,547	84.9	83.4	77.3
Aggressive Behavior	189,695	44.9	47.2	75.8
Disrespect to Faculty	206,862	34.5	40.8	67.5
Insubordination	435,622	27.5	30.5	62.1
Inappropriate Language	259,337	35.5	41.9	77.3
Skipping Class	204,914	14.4	20.9	38.1
Disruptive Behavior	753,064	23.6	25.2	82.3
Excessive Tardiness	309,018	8.0	8.9	24.3

Note: Column 2 lists the total number of incidents for each category (or type) of offense in the sample. Column 3 (Column 4) lists the percentages of students in the academic year who were punished by out-of-school suspension (or expulsion) for the first (second) offense in each category. Column 5 lists percentages of schools with any reported offense in each category among academic year. The total number of school-year observations is 11,425. Type "fighting" and "aggressive behavior" are in the "violence" category; type "disrespect to faculty," "insubordination," and "inappropriate language" are in the "disrespect" category; type "skipping class" is in the "truancy" category; type "disruptive behavior" and "excessive tardiness" are from the "other" category.

Table 1.2: Sample Means of Student Characteristics

	Student-Year Sample with No Offense Record	Any Offense Record No OSS	Any Offense Record Any OSS
Race			
White	0.574	0.503	0.325
Black	0.228	0.326	0.514
Hispanic	0.118	0.107	0.094
Asian	0.030	0.010	0.006
Multi-Racial	0.035	0.038	0.037
American Indian	0.014	0.014	0.023
Other Race	0.001	0.000	0.001
Disability			
No Disability	0.875	0.829	0.778
Physical Disability	0.056	0.078	0.088
Intellectual Disability	0.069	0.093	0.133
Other Dichotomous Characteristics (omitted: alternative group)			
Female	0.531	0.390	0.313
Economically Disadvantaged	0.453	0.603	0.735
Limited English Proficiency	0.062	0.049	0.049
Academically and Intellectually Gifted - Reading	0.143	0.074	0.032
Academically and Intellectually Gifted - Math	0.155	0.081	0.037
Old in the Grade	0.118	0.187	0.301
Repeating Grade in the Academic Year	0.016	0.039	0.104
Mean of Lagged Scores			
Lagged Normalized Math Score	0.043	-0.048	-0.202
Lagged Normalized Reading Score	0.064	-0.041	-0.198
Lagged Score Missing Indicator	0.216	0.152	0.178
Grade level			
Grade 3	0.122	0.047	0.036
Grade 4	0.120	0.055	0.048
Grade 5	0.115	0.063	0.060
Grade 6	0.098	0.113	0.110
Grade 7	0.093	0.122	0.127
Grade 8	0.092	0.124	0.135
Grade 9	0.098	0.141	0.188
Grade 10	0.091	0.126	0.129
Grade 11	0.086	0.113	0.097
Grade 12	0.086	0.096	0.070
Observations (student years)	4,034,542	658,611	577,886

Note: This table separately reports summary statistics for the student-year samples without offense records, with offense records but without out-of-school suspension records, and with both offense records and any out-of-school suspension records. For grade 10-12 students, End-of-Course Test English 1 is used for the calculation of lagged reading scores; End-of-Course Test Algebra 1 is used for the calculation of lagged math scores. Lagged test scores are normalized to have a zero mean and standard deviation of one among the students who took the same tests across the state.

Table 1.3: Instrumental Variable and Time Varying School Observed Variables

Dependent Variable: Principal Team's Exclusive OSS Tendency (IV)	Coefficient	Standard Error
Violent crime cases last year (N)	0.0009	(0.0006)
Students involved in misbehavior last year (N)	-0.0001	(0.0001)
Students assigned OSS or expulsion last year (N)	0.0000	(0.0001)
Assault, robbery or sexual offense cases (N)	-0.0004^{*}	(0.0002)
Threat or possession of a weapon cases (N)	-0.0003	(0.0004)
Disorderly conduct or harassment cases (N)	-0.0001	(0.0001)
Other violent cases (N)	-0.0000	(0.0001)
Drug related cases (N)	0.0000	(0.0002)
Disrespect cases (N)	-0.0000	(0.0000)
Disruptive behavior cases (N)	0.0000	(0.0000)
Truancy cases (N)	0.0000	(0.0000)
Tardiness cases (N)	0.0000	(0.0000)
Property cases (N)	0.0004	(0.0003)
Other rule violation cases (N)	0.0000	(0.0000)
Other disciplinary infraction cases (N)	0.0001^{**}	(0.0000)
Minor (average OSS days < 0.55) misbehavior cases (N)	-0.0000	(0.0000)
Moderate ($0.55 \leq$ average OSS days ≤ 1) misbehavior cases (N)	-0.0000	(0.0000)
Major (average OSS days > 1) misbehavior cases (N)	-0.0001	(0.0000)
Ratio of black students	0.0056	(0.0327)
Ratio of Hispanic students	-0.0435	(0.0455)
Ratio of other minority students	-0.1419^{*}	(0.0755)
School mean of normalized math score last year	-0.0030	(0.0161)
School mean of normalized reading score last year	-0.0001	(0.0141)
Proportion of students – math scores 2 sd below state average last year	-0.0584	(0.1001)
Proportion of students – reading scores 2 sd below state average last year	0.0799	(0.1011)
Title I eligible school	0.0046	(0.0075)
School-wide title I	0.0131^{*}	(0.0075)
Ratio of teachers licensed in the school for more than 5 years	0.0132	(0.0168)
Ratio of female personnels	0.0579	(0.0420)
Ratio of black personnels	0.0474	(0.0479)
Ratio of non-white non-black personnels	0.0507	(0.0866)
Magnet School Indicator	-0.0044	(0.0167)
Total student number	0.0000	(0.0000)
Students who are economically disadvantaged %	0.0069	(0.0284)
Total full-time equivalent classroom teachers	0.0002	(0.0006)
Total number of classroom teachers	-0.0012^{**}	(0.0005)
Fully licensed teachers %	-0.0153	(0.0432)
Teachers with experience 4-10 years %	-0.0063	(0.0296)
Teachers with experience more than 11 years %	-0.0613^{*}	(0.0316)
Teachers with Advanced Degrees %	-0.0220	(0.0314)
Teacher Turnover Rate %	0.0119	(0.0262)
Average daily school attendance %	0.0089	(0.2181)
Students per Instructional Computer (N)	-0.0009	(0.0008)
Books per Student (N)	-0.0001	(0.0002)
Average age of Books in library or media center	-0.0000	(0.0000)
Classes taught by highly qualified teachers %	-0.0531	(0.0475)
Adequate yearly progress target met %	-0.0120	(0.0120)
Classrooms connected to the Internet %	-0.0539^{*}	(0.0288)
PBIS - Green Ribbon School	0.0059	(0.0072)
PBIS - Model School	0.0029	(0.0079)
PBIS - Exemplar School	0.0229^{**}	(0.0114)
One or more school variables were missing	-0.0026	(0.0100)
Number of school-year observations	9210	

Robust standard errors are in the parentheses.

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

Note: Instead of per student numbers, raw numbers of offense cases are used because total number of students is in the controls.

Table 1.4: Monotonicity of the Instrumental Variable

Type of Offense	Coefficient	Standard Error	Observations
Assault on student	0.133***	(0.039)	16153
Assault on student (no weapon, no serious injury)	0.107*	(0.056)	10244
Fighting	0.022***	(0.009)	156787
Aggressive behavior	0.062***	(0.014)	133864
Bullying	0.143***	(0.031)	30541
Gang activity	0.122*	(0.071)	5552
Disorderly conduct	0.057	(0.038)	23297
Communicating threats	0.061*	(0.033)	20902
Harassment - verbal	0.023	(0.044)	14629
Harassment - sexual	0.077	(0.050)	11803
Disrespect of faculty/staff	0.058***	(0.014)	134443
Inappropriate language/disrespect	0.034***	(0.011)	181037
Insubordination	0.072***	(0.010)	234486
Disruptive behavior	0.020***	(0.007)	368301
Possession of marijuana	−0.062	(0.045)	10707
Possession of a weapon (not firearms or explosives)	0.055	(0.039)	13295
Possession of tobacco	−0.008	(0.041)	11969
Use of tobacco	−0.004	(0.030)	19997
Theft	0.022	(0.026)	35716
Property damage	−0.075*	(0.042)	16329
Inappropriate items on school property	0.033	(0.043)	16779
Skipping class	0.019*	(0.010)	131950
Truancy	0.145***	(0.044)	21672
Leaving class without permission	0.044*	(0.023)	33621
Leaving school without permission	0.004	(0.033)	22644
Skipping school	−0.054*	(0.031)	27448
Late to class	−0.075***	(0.012)	67915
Excessive tardiness	−0.119***	(0.009)	128301
Excessive display of affection	0.045	(0.042)	10145
Honor code violation	0.248***	(0.030)	16427
Dress code violation	−0.016	(0.020)	45371
Falsification of information	−0.037	(0.055)	8683
Being in an unauthorized area	−0.012	(0.032)	21765
Cell phone use	−0.027*	(0.015)	71203
Bus misbehavior	−0.003	(0.006)	131022
Other School Defined Offense	0.093***	(0.019)	86391
Other	0.065***	(0.019)	64154
Misuse of school technology	0.019	(0.047)	11770
For all types	0.037***	(0.004)	1077758

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

Note: This table reports OLS regression results of coefficients γ_1 for different types of offenses. The dependent variable is the out-of-school suspension indicator. The key explanatory variable is the (main) instrumental variable. Standard errors are reported in parentheses and clustered at the student level.

Table 1.5: General Deterrence Effects for First Offense by Offense Category

Offense Category	OLS	OLS &GFE	2SLS	2SLS &GFE	Sample Size	Mean of Dependent Var.
All Offenses	-0.244^{***}	-0.248^{***}	-0.372^{***}	-0.397^{***}	4545364	0.250
	(0.001)	(0.001)	(0.034)	(0.028)		
Violence	-0.041^{***}	-0.039^{***}	-0.141^{***}	-0.116^{***}	4372421	0.083
	(0.001)	(0.001)	(0.027)	(0.022)		
Disrespect	-0.064^{***}	-0.071^{***}	-0.114^{***}	-0.121^{***}	4003540	0.104
	(0.001)	(0.001)	(0.017)	(0.014)		
Truancy	-0.061^{***}	-0.061^{***}	-0.139^{***}	-0.154^{***}	2984461	0.068
	(0.001)	(0.001)	(0.030)	(0.029)		
Drug	-0.001^{**}	-0.001	-0.038^{**}	-0.040^{***}	2321145	0.022
	(0.001)	(0.001)	(0.015)	(0.014)		
Property	-0.001^{**}	-0.000	-0.009	0.020	2870740	0.015
	(0.000)	(0.000)	(0.065)	(0.031)		

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

Note: This table reports estimates of α_{11} in equation 1.1 for different categories of misbehaviors (in different rows), estimated separately. Columns 2-5 show the results using different estimation methods. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include the control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. The sample does not include schools with less than 5 observations of students who have committed first offenses in an offense category in the academic year; that's why the sample size changes across different categories. Using a subset of schools with all categories of offenses results in a substantial loss of observations. In the appendix, I compare the summary statistics of observed characteristics for each above sample to show that there is no significant selection problem for the sample of each category. The mean of the dependent variable (offense indicator) is reported in the last column. Standard errors are reported in parentheses and clustered at the student level.

Table 1.6: General Deterrence Effects for First Offense by Offense Type

Offense Category	OLS	OLS &GFE	2SLS	2SLS &GFE	Sample Size	Mean of Dependent Var.
Fighting	-0.009^{***}	-0.009^{***}	-0.019	-0.017	4009898	0.038
	(0.001)	(0.001)	(0.032)	(0.028)		
Aggressive Behavior	-0.018^{***}	-0.018^{***}	-0.242^{***}	-0.166^{***}	3493914	0.037
	(0.000)	(0.000)	(0.050)	(0.032)		
Disrespect to Faculty	-0.024^{***}	-0.026^{***}	-0.029^{***}	-0.031^{***}	3239179	0.040
	(0.001)	(0.001)	(0.010)	(0.009)		
Insubordination	-0.051^{***}	-0.053^{***}	-0.232^{***}	-0.228^{***}	3365506	0.068
	(0.001)	(0.001)	(0.038)	(0.033)		
Inappropriate Language	-0.016^{***}	-0.018^{***}	-0.168^{***}	-0.144^{***}	3833696	0.046
	(0.001)	(0.001)	(0.020)	(0.017)		
Skipping Class	-0.037^{***}	-0.037^{***}	-0.054	-0.056^{*}	2807283	0.047
	(0.001)	(0.001)	(0.034)	(0.034)		
Disruptive Behavior	-0.085^{***}	-0.091^{***}	-0.242^{*}	-0.203^{***}	3813508	0.096
	(0.001)	(0.001)	(0.128)	(0.074)		
Excessive Tardiness	-0.071^{***}	-0.072^{***}	-0.180^{***}	-0.124^{**}	1513515	0.073
	(0.001)	(0.001)	(0.059)	(0.061)		

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

Note: This table reports estimates of α_{11} in equation 1.1 for different types of misbehaviors (in different rows), estimated separately. Columns 2-5 show the results using different estimation methods. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include the control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. The sample does not include schools with less than 3 observations of students who have committed first offenses in an offense type in the academic year; that's why the sample size changes across different types. Offense type "fighting" and "aggressive behavior" are from the "violence" category; offense type "disrespect to faculty,""insubordination," and "inappropriate language" are from the "disrespect" category; offense type "skipping class" is from the "truancy" category; offense type "disruptive behavior" and "excessive tardiness" are from the "other" category. The mean of the dependent variable (offense indicator) is reported in the last column. Standard errors are reported in parentheses and clustered at the student level.

Table 1.7: General Deterrence Effects for First Offense by Student Characteristics

Sample	Violence		Disrespect	
	OLS &GFE	2SLS &GFE	OLS &GFE	2SLS &GFE
Elementary School	-0.027***	-0.108***	-0.030***	-0.024
	(0.001)	(0.014)	(0.001)	(0.023)
	1146718	1146718	854611	854611
	0.069	0.069	0.048	0.048
Middle School	-0.046***	-0.152**	-0.076***	-0.158***
	(0.002)	(0.062)	(0.001)	(0.045)
	1461682	1461682	1433382	1433382
	0.124	0.124	0.111	0.111
High School	-0.013***	-0.004	-0.070***	-0.167***
	(0.002)	(0.047)	(0.001)	(0.019)
	1880429	1880429	1879421	1879421
	0.056	0.056	0.117	0.117
White Students	-0.024***	-0.079***	-0.050***	-0.168***
	(0.001)	(0.025)	(0.001)	(0.021)
	2241355	2241355	2040484	2040484
	0.054	0.054	0.072	0.072
Black Students	-0.054***	-0.175***	-0.101***	-0.176***
	(0.002)	(0.059)	(0.002)	(0.032)
	1229951	1229951	1150650	1150650
	0.149	0.149	0.180	0.180
Female	-0.016***	-0.003	-0.047***	-0.048***
	(0.001)	(0.050)	(0.001)	(0.015)
	1896881	1896881	1744554	1744554
	0.055	0.055	0.077	0.077
Male	-0.050***	-0.173***	-0.083***	-0.174***
	(0.001)	(0.033)	(0.001)	(0.021)
	2279396	2279396	2106486	2106486
	0.113	0.113	0.135	0.135
Econ Disadvantage	-0.042***	-0.164***	-0.084***	-0.145***
	(0.001)	(0.036)	(0.001)	(0.024)
	2231892	2231892	2060519	2060519
	0.121	0.121	0.146	0.146
Not Econ Disadvantage	-0.018***	-0.070**	-0.043***	-0.114***
	(0.001)	(0.029)	(0.001)	(0.017)
	2040170	2040170	1854851	1854851
	0.045	0.045	0.060	0.060
Lagged math ≤ 0	-0.034***	-0.133	-0.086***	-0.194***
	(0.002)	(0.082)	(0.002)	(0.045)
	832128	832128	794255	794255
	0.135	0.135	0.184	0.184
Lagged math > 0	0.028***	-0.079***	-0.054***	-0.078***
	(0.001)	(0.026)	(0.001)	(0.016)
	2654784	2654784	2454586	2454586
	0.073	0.073	0.082	0.082

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

Note: This table reports estimates of α_{11} in equation 1.1 for the “violence” and “disrespect” categories by student characteristics, estimated separately. OLS&GFE: a model with all control variables, school fixed effects and student unobserved heterogeneity; 2SLS&GFE: additionally instruments the DPI variable. Lagged math score is normalized to have a zero mean and standard deviation of one among the students who took the same tests across the state. Standard errors are reported in parentheses and clustered at the student level. The third number in each cell is the sample size. The fourth number in each cell is the mean of dependent variable (offense indicator). The sample for each estimation does not include schools with less than 3 students’ first offense observations from the category in the subpopulation in the academic year.

Table 1.8: General and Specific Deterrence Effects for Second Offense

Offense Category	OLS	OLS &GFE	2SLS	2SLS &GFE	Sample Size	Mean of Dependent Var.
Violence						
α_{21}	-0.028^{***}	-0.034^{***}	-2.437	-2.489	351829	0.280
	(0.005)	(0.005)	(3.594)	(4.660)		
α_{22}	0.008^{***}	0.004^{**}	2.515	2.582		
	(0.002)	(0.002)	(3.901)	(5.081)		
Disrespect						
α_{21}	-0.064^{***}	-0.069^{***}	1.951	2.122	407292	0.419
	(0.005)	(0.005)	(2.498)	(3.028)		
α_{22}	-0.000	-0.008^{***}	-1.780	-1.915		
	(0.002)	(0.002)	(2.096)	(2.572)		
Truancy						
α_{21}	-0.008	-0.011^{*}	1.228	-0.529	195544	0.314
	(0.007)	(0.006)	(5.099)	(2.598)		
α_{22}	-0.002	-0.006^{*}	-1.670	0.751		
	(0.003)	(0.003)	(7.187)	(3.755)		

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

Note: This table reports estimates of the general deterrence effect (α_{21}) and the specific deterrence effect (α_{22}) in equation 1.2 for different categories of misbehaviors, estimated separately. Columns 2-5 show the results using different estimation methods. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include a control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. The sample for each category does not include schools with less than 5 students' first offense observations in the category or without second offense observations in the category in the academic year; that's why the sample size changes across different categories. The mean of the dependent variable is reported in the last column. Standard errors are reported in parentheses and clustered at the student level.

Table 1.9: General Deterrence Effects for Second Offense by OSS Experience

Offense Category	Sample with OSS Experience		Sample without OSS Experience	
	OLS & GFE	2SLS & GFE	OLS & GFE	2SLS & GFE
Violence	-0.054***	-0.117	-0.068***	-0.264
	(0.006)	(0.117)	(0.009)	(0.206)
	226697	226697	125132	125132
	0.260	0.260	0.310	0.310
Disrespect	-0.057***	-0.206	-0.121***	-0.348***
	(0.008)	(0.410)	(0.007)	(0.116)
	119209	119209	288115	288115
	0.404	0.404	0.425	0.425
Truancy	0.014	0.854	-0.022***	-1.231
	(0.013)	(0.883)	(0.008)	(0.909)
	33936	33936	161203	161203
	0.301	0.301	0.315	0.315

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

Note: This table reports estimates of α_{21} (general deterrence effects on students' second offense) in equation 1.2 by separately using the student-observations that were out-of-school suspended (OSS) for their first offenses ($p_{1ist} = 1$) and the student-observations that were not out-of-school suspended for their first offenses ($p_{1ist} = 0$). The third number in each cell is number of observations. The third number in each cell is the sample size. The fourth number in each cell is the mean of the dependent variable (the offense indicator). Standard errors are reported in parentheses and clustered at the student level.

Table 1.10: Effects of Suspension Experience on Offending Students' Achievement Outcomes

	All Students	Only Offending Students						
Achievement		Ever OSS or Not in the Year					OSS/Not for 1st Offense	
Outcome	OLS Limited Controls	OLS Limited Controls	OLS Full Controls	OLS & GFE	2SLS	2SLS & GFE	OLS & GFE	2SLS & GFE
End of Grade Math	-0.206^{***}	-0.143^{***}	-0.013^{***}	-0.012^{***}	0.068	0.051	-0.025^{***}	0.125
(Grade 3-8)	(0.003)	(0.003)	(0.002)	(0.002)	(0.187)	(0.178)	(0.003)	(0.293)
	3145909	548893	548893	548893	548893	548893	548893	548893
Dropout	0.185^{***}	0.154^{***}	0.072^{***}	0.050^{***}	-0.077	-0.216	0.038^{***}	-0.356
(Grade 9)	(0.002)	(0.003)	(0.003)	(0.003)	(0.186)	(0.187)	(0.003)	(0.319)
	165046	114642	114642	114642	114642	114642	114642	114642
ACT Composite Score	-1.112^{***}	-0.999^{***}	-0.211^{***}	-0.178^{***}	-2.741	-2.068	-0.120^{***}	-5.956
(Grade 9-10)	(0.026)	(0.029)	(0.024)	(0.024)	(1.824)	(1.807)	(0.029)	(5.979)
	115719	97419	97419	97419	97419	97419	97419	97419

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

Note: This table reports estimates of α_{31} and α'_{31} – the effects of out-of-school suspension on offending students' achievement outcomes. The achievement outcomes include end-of-grade test math score (normalized), a dropout indicator (finally graduate or dropout from the high school) and ACT composite score (points). The estimation for dropout and ACT composite scores uses suspension experience of students in grade 9 and students in grade 9-10 respectively. While the second column shows the results using the "all students" sample, other columns show the results by limiting the sample to students with an offense record – students at risk of suspension. The key explanatory variable for Columns 2-6 is an indicator of whether the student was ever suspended in the academic year. For Columns 7-8, it is an indicator of whether the student was suspended for the first offense in the academic year. OLS (Limited Controls): a model with only school-level control variables and school fixed effects; OLS (Full Controls): a model with all control variables (student-level and school-level) and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include the control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. Standard errors are reported in parentheses and clustered at the student level. The third number in each cell is the sample size.

Table 1.11: Effects of Suspension Experience on Achievement by Race

Achievement Outcome by Race	OLS	OLS & GFE	2SLS	2SLS & GFE	Sample Size
End of Grade Math Test Score					
White	-0.018***	-0.017***	-0.328	-0.242	212389
(Grade 3-8)	(0.003)	(0.003)	(0.702)	(0.526)	
Black	-0.008**	-0.008**	0.300	0.305	241130
(Grade 3-8)	(0.003)	(0.003)	(0.223)	(0.227)	
Dropout					
White	0.074***	0.050***	-0.289	-0.383	47751
(Grade 9)	(0.005)	(0.005)	(0.422)	(0.413)	
Black	0.066***	0.046***	0.113	-0.033	47750
(Grade 9)	(0.004)	(0.004)	(0.260)	(0.249)	
ACT Composite Score					
White	-0.205***	-0.178***	-3.345*	-2.886	43385
(Grade 9-10)	(0.042)	(0.042)	(1.813)	(1.839)	
Black	-0.180***	-0.149***	0.029	1.565	38894
(Grade 9-10)	(0.033)	(0.033)	(6.906)	(6.008)	

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

Note: This table reports estimates of α_{31} for white and black students – the effects of out-of-school suspension experience on white or black offending students' achievement outcomes. The achievement outcomes include end-of-grade test math score (normalized), a dropout indicator (finally graduate or dropout from the high school) and an ACT composite score (points). The estimation for dropout and ACT composite scores uses suspension experience of students in grade 9 and students in grade 9-10 respectively. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include the control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. Standard errors are reported in parentheses and clustered at the student level.

Table 1.12: Effects of Discipline on Achievement of Well-Behaved Students and All Students

Achievement Outcome by Student Group	OLS	OLS & GFE	2SLS	2SLS & GFE	Sample Size
End of Grade Math Score (Grade 3-8)					
All Students	0.003	-0.003	-0.014	-0.029	2711053
	(0.003)	(0.003)	(0.054)	(0.050)	
Students with no offense record	-0.001	-0.003	0.103*	0.074	1225598
	(0.003)	(0.003)	(0.061)	(0.055)	
End of Grade Math Score (Grade 6-8)					
All Students	0.002	0.001	0.168	0.116	1433766
	(0.003)	(0.003)	(0.126)	(0.111)	
Students with no offense record	-0.001	0.002	0.222**	0.202**	649423
	(0.004)	(0.004)	(0.105)	(0.101)	
Dropout					
All Students (Grade 9)	0.003	0.001	0.068	-0.041	400444
	(0.004)	(0.003)	(0.087)	(0.084)	
Students with no offense record	0.005	0.004	0.019	0.050	153684
	(0.003)	(0.003)	(0.049)	(0.045)	
ACT Composite Score					
All Students (Grade 9-10)	0.025	0.076*	-0.545	-0.138	444065
	(0.046)	(0.046)	(0.988)	(1.046)	
Students with no offense record	-0.025	-0.008	-1.320	-1.394	197375
	(0.067)	(0.067)	(1.008)	(1.042)	

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

Note: This table reports estimates of α_{41} and α_{51} – the effects of out-of-school suspension on "never offending" students or all students' achievement outcomes. The achievement outcomes include an end-of-grade test math score (normalized), a dropout indicator (finally graduate or dropout from the high school) and an ACT composite score (points). The estimation for dropout and ACT composite scores uses suspension experience of students in grade 9 and students in grade 9-10 respectively. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include the control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. Standard errors are reported in parentheses and clustered at the student level.

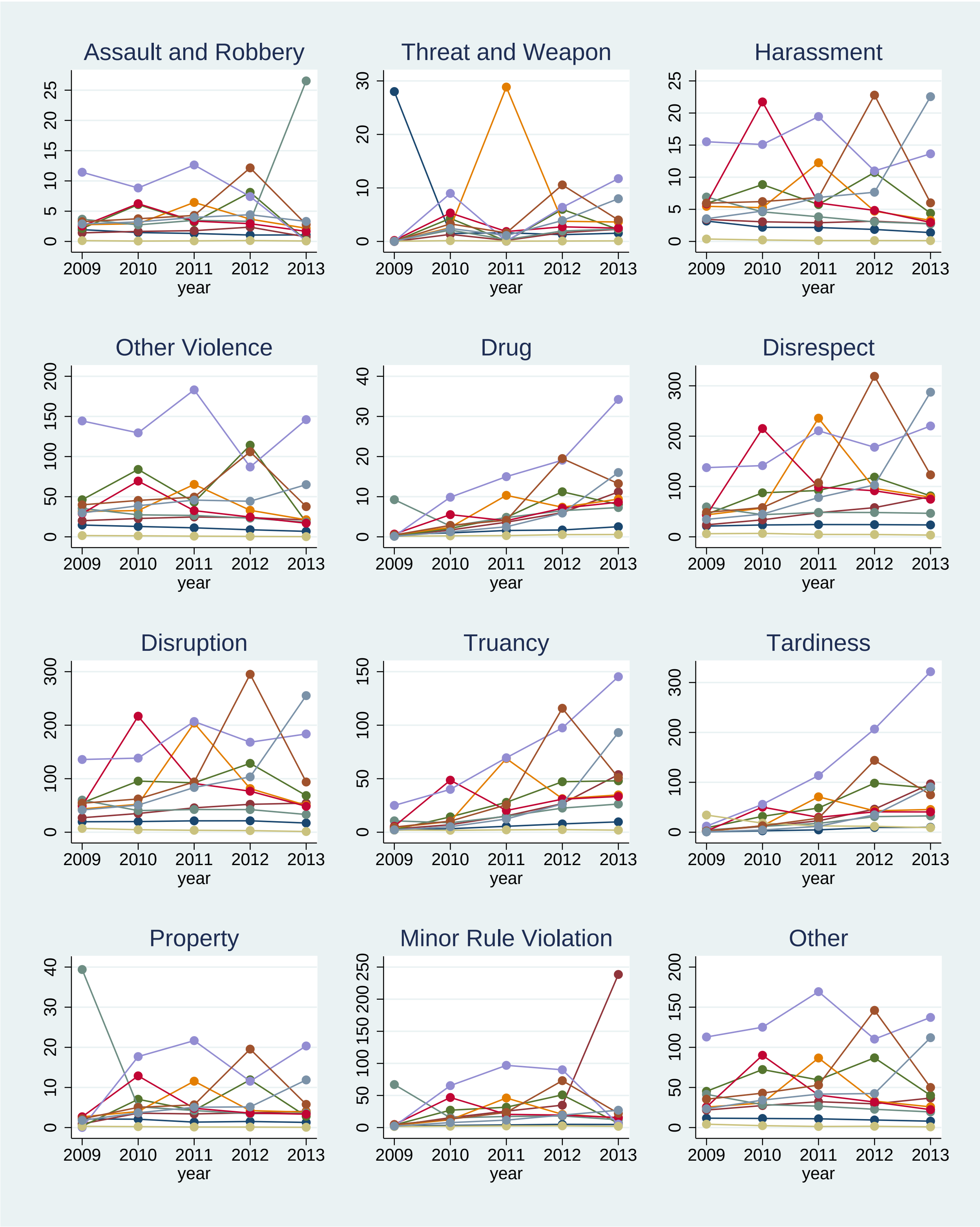
Table 1.13: Overall Effects of Discipline on Achievement by Race

Achievement Outcome	White Students		Black Students	
	OLS & GFE	2SLS & GFE	OLS & GFE	2SLS & GFE
End of Grade Math Test Score				
(Grade 3-8)	0.000	0.084	0.004	-0.239
	(0.003)	(0.054)	(0.005)	(0.153)
	1420134	1420134	735245	735245
(Grade 6-8)	0.002	0.282***	0.004	2.058
	(0.004)	(0.099)	(0.007)	(1.891)
	765241	765241	392171	392171
Dropout				
(Grade 9)	-0.003	-0.017	0.007	0.107
	(0.004)	(0.060)	(0.007)	(0.344)
	217395	217395	114156	114156
ACT Composite Score				
(Grade 9-10)	0.147**	-1.059	-0.142*	1.343
	(0.063)	(0.799)	(0.074)	(2.254)
	256441	256441	115083	115083

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

Note: This table reports estimates of α_{51} for black and white students, estimated separately. The second and third columns show the results for white students; the fourth and fifth columns show the results for black students. The estimation for dropout and ACT composite scores uses the suspension experience of students in grade 9 and students in grade 9-10 respectively. OLS&GFE: a model with all control variables, school fixed effects and student unobserved heterogeneity; 2SLS&GFE: additionally instruments the DPI variable. The third number in each cell is sample size. Standard errors are reported in parentheses and clustered at the student level.

Figure 1.1: Estimated Student Group and Offense Numbers



Note: This figure shows the frequencies (in units of hundreds) of each category of offense across years for each estimated group of students. The ten colors represent ten different estimated groups of students.

CHAPTER 2

THE EQUITY OF EXCLUSIONARY SCHOOL DISCIPLINE

2.1 Introduction

Exclusionary school discipline, which refers to school suspension or expulsion, is a disciplinary practice that punishes misbehaving students by isolating them from their classroom environments. It is widely used in U.S. public schools; for example, in the 2013-2014 school year alone, 2.8 million of the 50 million public school students were suspended out-of-school at least once.[1] However, in the past few years, the popular press and policy makers express an increasing concern about racial disparity in its implementation. For example, in the 2013-2014 school year, black students represented 39 percent of student out-of-school suspensions, but they only comprised 15.5 percent of the public school student population. While 5 percent of white boys and 2 percent of white girls were suspended out-of-school at least once, the numbers for black boys and black girls were 18 percent and 10 percent.[2]

The racial disparity in school discipline is troubling especially because some literature documents that school suspension is harmful to suspended students - it lowers their academic achievement and raises school dropout rates (Raffaele Mendez 2003; Wald and Losen 2003; Arcia 2006; Lee et al. 2011; Skiba and Rauch 2015). Therefore, racial disparity in school discipline may create black-white academic achievement gaps, which may further create black-white gaps in future economic and life outcomes.[3] This evidence has motivated the U.S. Departments of Justice and

[1]Civil Rights Data Collection, 2016, U.S. Department of Education Office for Civil Rights.

[2]Civil Rights Data Collection, 2016, U.S. Department of Education Office for Civil Rights.

[3]Morris and Perry (2016) show that racial disparity in exclusionary school discipline may account for approximately 20 percent of black-white differences in school performance. Black-white academic achievement gap is of great importance for understanding black-white gaps in economic outcomes. For example, Neal and Johnson (1996) document that a test score (AFQT) explains nearly three-quarters of the racial wage gap for young men and all of the

Education to release a school discipline guidance package in 2014 to reform discipline policies and has motivated several states to enact new legislation.[4]

There is new evidence that the negative causal effects of school suspension on students' academic achievement and school dropout rates are actually small or not statistically significant (Li 2017). In addition, school suspension can statistically significantly deter students from committing offenses (Kinsler 2013; Li 2017). However, even if there are no negative effects of school suspension on suspended students, it is still important to study the racial disparity in school suspension. An important question is whether black students are more likely to be suspended than white students for the same type of offenses in the same circumstance, which is generally regarded as discrimination towards black students regardless of the consequences of the punishments. A large amount of literature documents that the racial disparity in out-of-school suspension persists even after accounting for offense types and other observable characteristics of students and observable environmental measures of schools (e.g., Skiba, Michael, Nardo, and Peterson (2002), Mendez and Knoff (2003), Wallace Jr, Goodkind, Wallace, and Bachman (2008), Skiba, Chung, Trachok, Baker, Sheya, and Hughes (2014), Anyon, Jenson, Altschul, Farrar, McQueen, Greer, Downing, and Simmons (2014)). However, using administrative data on North Carolina public school students in one school year (2000-2001), Kinsler (2011) found that the racial disparity documented in the literature is primarily generated by cross-school disparity of principals' punishments. It disappears entirely when black and white student suspensions are compared within schools. This evidence suggests that the disparity can be explained by a higher proportion of black students in the schools with more severe discipline. Furthermore, Kinsler (2013) documents that the cross-school disparity of principals' punishments is consistent with their punishment strategies that focus on maximizing students' academic achievements. Therefore, the evidence suggests that racial discrimination plays no role in out-of-school suspension decisions.

---

gap for young women.

[4]There is new related legislation in several states, for example, California (AB 420, 2014) and Illinois (SB 100, 2015).

Using rich administrative data on North Carolina public school students from the 2008-2009 to 2014-2015 school years, this study provides new evidence on the racial disparity of school suspension between black and white students, as well as punishment disparities among students from families with different economic statuses. Compared to the 2000-2001 school year data used in Kinsler (2011), the data in this study document many more types of students' offenses. The students' offenses are classified into about 90 types, which allows for a more detailed study of punishment disparities for different types of offenses. The panel feature of the data in this study also provides important controls for students' offense histories in estimation.

Similarly to Kinsler (2011), I find that the average punishment disparity (for "out-of-school suspension or not") between black and white students for all types of offenses disappears (and even slightly favors black students) when suspensions are compared within the same school in the same academic year (i.e., when offenses are more likely to be punished by the same principals).[5] However, when estimation is conducted separately for each type of offenses, I find that racial disparities exist but the direction depends importantly on the types of offenses when suspensions are compared within the same school in the same academic year.[6] While black students were more likely to be suspended for fighting, sexual harassment, aggressive behavior, dress code violation, theft, and excessive display of affection, white students were more likely to be suspended for communicating threats, verbal harassment, inappropriate language, insubordination, disrespect toward faculty, truancy, leaving class without permission and skipping class. I do not find evidence that these racial disparities are correlated with principals' or assistant principals' race. In addition, I find that Economically Disadvantaged students are consistently more likely to be suspended out-of-school for almost all types of offenses, regardless of the comparison restricted to be within schools or not.

The rest of the chapter is organized as follows. Section 2.2 describes the data for this research.

[5]The estimation is conditional on types of offenses.

[6]The results are similar when the comparison is only restricted to be within the same school.

Section 2.3 discusses the empirical framework and provides estimation results. Section 2.4 concludes.

2.2 Data Description

The administrative data of North Carolina public schools are provided by the North Carolina Education Research Data Center (NCERDC). They were originally collected by the North Carolina Department of Public Instruction (NCDPI) and the National Center for Education Statistics (NCES). The data include students' disciplinary infraction records, demographic information, and academic records, and information on teachers and schools.

The disciplinary infraction records span from the 2000-2001 to 2014-2015 academic years. I use data from the 2007-2008 to 2014-2015 academic years since the matching rate of infraction data with other data has largely increased since 2007-2008 academic year and reporting requirements for offenses have been greater.[7] Furthermore, since lagged student offense records are used as control variables, the 2007-2008 academic year data are not used to construct dependent variables in estimation. In addition, I use student observations of grades 3-12 since several explanatory variables, such as economically disadvantaged status and limited English proficiency status, are not available for grades K-2.

There are 5,249,004 recorded offense instances committed by 800,484 distinct grade 3-12 students from the 2008-2009 to 2014-2015 academic years.[8] Table 2.1 reports sample means of student characteristics for all offenses, for offenses that were not punished by out-of-school suspension, and for offenses that were punished by out-of-school suspension. The table shows that black students account for 47 percent of total offenses, but they account for 56.8 percent of offenses that were punished by out-of-school suspension (including expulsion). While Economically Disadvantaged students comprised 72.2 percent of total offenses, they represented 77 percent of offenses

[7]The matching rate of infraction data with other data is higher than 99 percent since the 2007-2008 academic year. The infraction records are not used for this project if they are not matched with other data.

[8]The sample does not include offense instances missing some important information, such as type of offense. These offense instances are less than 1 percent of total offense instances.

that were punished by out-of-school suspension.[9] For all offending students, their average lagged math and reading scores are below average by about 0.181 standard deviations. For suspended students, their average lagged math scores are below average by 0.247 standard deviations and their reading scores are below average by 0.245 standard deviations.

The NCDPI classifies offenses using about 90 offense types. Table 2.2 provides percentages of offenses that were punished by out-of-school suspension for several of the most common offense types.[10] The table also reports percentages of offenses that were punished by out-of-school suspension for white students, black students, Economically Disadvantaged students and Non-economically Disadvantaged students.[11] Note that for all types of offenses, the percentages of out-of-school suspensions were higher for black students than white students. For almost all types of offenses, the percentages of out-of-school suspensions were higher for Economically Disadvantaged students than Non-economically Disadvantaged students (except for "possession of marijuana").

2.3 Estimation and Results

Student infractions are commonly caught by teachers or other school personnel, and referred to principals' offices. Then, school principals or assistant principals determine appropriate punishments for them.[12] In this section, I estimate whether students' race or economically disadvantaged status affects principals' punishment decisions on "out-of-school suspension (including expulsion) or not." Note that the analysis does not answer the question of whether there is racial bias in

---

[9]Economically disadvantaged students are students receiving free or reduced price meals. The eligibility for free or reduced lunch is determined by family size and family income. The most recent criteria can be found on http://www.dpi.state.nc.us/newsroom/news/2015-16/20150814-01.

[10]Each offense instance may be described by multiple offense types. The statistics in Table 2.2 only use offenses with a single offense type. The offenses described by multiple offense types are only about one percent of total offenses.

[11]Those offenses that were not punished by out-of-school suspension (including expulsion) were typically assigned less severe punishments, such as in-school suspension, lunch detention, or a warning.

[12]According to the North Carolina state statute, a long-term suspension must be assigned by the superintendent under a principal's recommendation (115C-390.7). Only about 1 percent of out-of-school suspensions were long-term ($>$ 10 days).

teachers' referral decisions.

Let Suspension_{oist} indicate whether or not student i's oth (reported) offense in school s in academic year t was punished by out-of-school suspension. I use the following OLS model for estimation:

$$\text{Suspension}_{oist} = \beta_0 + \beta_1 X_{ist} + \beta_2 \text{Record}_{oist} + \text{Type}_o + \phi_{st}^{sch} + \epsilon_{oist} \tag{2.1}$$

where X_{ist} is a vector of student observable characteristics (variables in Table 2.1); Record_{oist} is a vector of control variables that describe student i's offense records before offense o in academic year t in school s (variables in Appendix B); Type_o is a categorical variable specifying the type of offense; ϕ_{st}^{sch} represents school-year fixed effects;[13] ϵ_{oist} is the error term.

I begin by estimating equation 2.1 for all offenses in the sample. That is, while I control for offense type, I do not allow the other coefficients in the model to vary by offense type. Table 2.3 reports estimated coefficients of vector β_1. To compare with the existing literature, I also report estimated coefficients without controlling for school-year fixed effects ϕ_{st}^{sch} (in the left panel).[14] The results indicate that without controlling for school-year fixed effects, black students were 8.6 percentage points more likely to be suspended than white students (holding all other characteristics the same). After controlling for school-year fixed effects, the racial disparity almost disappears and white students were even 0.3 percentage points more likely to be suspended that black students. This result is largely consistent with Kinsler (2011), which indicates that the black-white disparity in "out-of-school suspension or not" is mostly from cross-school variation in punishment. In addition, it shows that when the comparison is within the same school in the same academic year, the punishment may slightly favor black students on average.

[13] For each of the specifications in this essay, I also estimate a model using school fixed effects instead of school-year fixed effects. All of the results are consistent.

[14] I control for year indicators for the specification without school-year fixed effects.

Furthermore, I find that, without controlling for school-year fixed effects, other minority students, such as Hispanic, Asian, Multiracial and American Indian, were more likely to be out-of-school suspended than white students, but the punishment differences become much smaller or not statistically significant after controlling for school-year fixed effects. In addition, I find, that without controlling for school-year fixed effects, Economically Disadvantaged students are 1.9 percentage points more likely to be suspended than Non-Economically Disadvantaged students. The result is similar after controlling for school-year fixed effects, which suggests that Economically Disadvantage students were more likely to be suspended than Non-economically Disadvantaged students even if the comparison is within-school. I also find that, with or without school-year fixed effects, disabled students were less likely to be out-of-school suspended; students who were old in the grade or who were repeating grade in the year were more likely to be suspended;[15] students with better past academic achievements or who were academically and intellectually gifted were less likely to be out-of-school suspended.

I also estimate equation 2.1 using students' first offenses only, and report the estimates for β_1 in Table 2.4.[16] The results show that, without controlling for school-year fixed effects, black students were 9.6 percentage points more likely to be suspended than white students for their first offenses. The racial impact drops significantly to 0.4 percentage points after controlling for school-year fixed effects, but remains statistically significant. Since standard errors for these estimates are small, the results suggest that the black-white disparity is mostly due to cross school variation. In addition, Economically Disadvantaged students were more likely to be out-of-school suspended (2.2 percentage points or 1.7 percentage points) with or without controlling for school-year fixed effects.

To explore whether the disparities change across school levels and time, I separately estimate equation 2.1 using offenses in each school level (i.e., elementary, middle and high school) in each

[15] "Old in the grade" is defined as not typical age in the grade. For example, most of the students in grade 4 are 9 or 10 years old. "Old in the grade" for grade 4 students are those who are older than 10.

[16] First offenses of students are defined by observed first offenses of students in the sample. Students in the sample have at least one lagged year offense records.

school year. Table 2.5 reports estimated coefficients on the black student indicator and Economically Disadvantaged student indicator. School fixed effects are included in these regressions. The results indicate that in elementary or high schools, for most academic years, there is no statistically significant punishment differences between black and white students. In middle schools, since the 2009-2010 academic year, black students were statistically significantly less likely to be out-of-school suspended than white students; the punishment disparities are 0.7 to 1 percentage points. The results also indicate that Economically Disadvantaged students were consistently more likely to be out-of-school suspended than Non-economically disadvantaged students across all school levels and across all school years.

The three sets of results above measure average punishment disparities for all types of offenses. An interesting question is whether the punishment disparities are heterogeneous for different types of offenses. To answer this question, I separately estimate equation 2.1 for different types of offenses. Estimation results for the marginal effects of being a black student (relative to a white student) and for the marginal effects of being an Economically Disadvantage student are separately reported in Table 2.6 and Table 2.7. In Table 2.6, the results suggest that without controlling for school-year fixed effects, black students are consistently (i.e., across all offense types) more likely to be out-of-school suspended than white students. However, after controlling for school-year fixed effects, the racial disparities depend on the type of offense. While black students were more likely to be suspended for fighting, sexual harassment, dress code violation, theft and excessive display of affection, white students were more likely to be suspended for communicating threats, verbal harassment, inappropriate language, insubordination, disruptive behavior, disrespect toward faculty, truancy, leaving class without permission and skipping class. The estimated punishment disparities (with statistical significance) range from 0.5 percentage points to 4.7 percentage points (in absolute value). The results that white students were more likely to be suspended for some subjective offenses, such as inappropriate language, insubordination, and disrespect toward faculty, are contrary to some exiting literature (e.g., Smith and Harper (2015)). In Table 2.7, the results suggest that, without or with controlling for school-year fixed effects, Economically Disadvantaged students are more likely to be out-of-school suspended for most types of offenses (for other types of offenses, the punishment differences are not statistically significant). The estimated punishment disparities range from 0.1 percentage points to 3.9 percentage points.[17]

To explore the heterogeneity in racial disparity for students of different genders, I estimate equation 2.1 for different types of offenses for male and female students separately, and report the analogous results in Table 2.8. The results are generally consistent with Table 2.6 for both male and female students, although some estimates become statistically insignificant.

I explore whether there are new findings of different impacts of race by offense type when I restrict the sample to the first offenses of students.[18] Table 2.9 reports the impacts of being a black student or being Economically Disadvantaged on out of school suspension probabilities. Compared to the estimates in Table 2.6 and Table 2.7, the signs for the estimates in Table 2.9 are mostly the same, although some estimates become statistically insignificant due to the smaller sample size. The differences include, black students are (statistically significantly) more likely to be suspended for disorderly conduct and aggressive behavior.[19]

An interesting check is whether the disparities in discipline are different by principals' race. If we assume, for example, the estimated black-white disparities are due to racial biases (or overcorrection of racial biases), this check may serve as an evidence for whether black and white principals have different degrees of or directions for racial biases (or overcorrection of racial biases). I estimate equation 2.1 with two additional interaction terms that interact a black or white principal

[17]A concern regarding results in Table 2.6 is that the heterogeneous black-white disparities across offense types may reflect unobserved differences among schools in which different types of offense occur. Since using a subset of schools with all types of offenses results in a substantial loss of observations and statistically insignificant results, I use a subset of schools that reported six or nine of the most common types of offenses (i.e., the largest sample sizes) among those with statistically significant results in Table 2.6. I report the estimated black-white disparities in Table B1 and B2 in the appendix. I found that the pattern of heterogeneity of black-white disparities across types remains. The only difference (among statistically significant results) is that black students are found more likely (rather than less likely) to be suspended for disruptive behavior.

[18]First offense is defined by a student's first offense of any type of misbehavior.

[19]I also find same evidence in a robustness check when using all offenses (not only first offenses) and school fixed effects (instead of school-year fixed effects).

indicator with both the black or white student indicator and the economically disadvantaged student indicator. I report the coefficients for the interaction terms in Table 2.10. The results show that there is no evidence that the estimated black-white disparities or disparities by economically disadvantaged status are systematically different between offenses that were punished by black principals and white principals. Note that these results do not rule out the possibility that black and white principals have similar directions of racial biases for different types of offenses.

In addition, I use a black or white assistant principal indicator instead of the principal indicator in Table 2.10 to run the same regressions and report the results for interaction terms in Table 2.11.[20] The results are generally consistent, except that for some type of offenses, I find that the disparities between economically disadvantaged and not disadvantaged students are smaller when the offenses were punished by black assistant principals (compared to white assistant principals).

2.4 Conclusion

Previous literature widely documents that black students face out-of-school suspension with higher probability than white students, even after controlling for a wide range of covariates, such as offense types, socioeconomic status and measures of school quality. The disparity is often characterized as a result of racial bias. Using administrative data on North Carolina public students in recent years, I find that after controlling for school-year fixed effects, the estimated punishment disparities are heterogeneous for different types of offenses. While black students were more likely to be suspended for fighting, sexual harassment, dress code violation, theft and excessive display of affection, white students were more likely to be suspended for communicating threats, verbal harassment, inappropriate language, insubordination, disruptive behavior, disrespect toward faculty, truancy, leaving class without permission and skipping class. In addition, I find that Economically Disadvantaged students are consistently more likely to be out-of-school suspended for different types of offenses, regardless of conditioning on school-year fixed effects or not.

A possible explanation for the findings that white students are more likely to be suspended

[20]When multiple assistant principals worked in one school in one academic year, I use the majority race to define the race of the assistant principals since I do not observe who assigned a punishment for each of the offense.

for some subjective offenses, such as verbal harassment, inappropriate language, insubordination, disrespect toward faculty, is that principals' suspension practices may favor black students for these types of offenses, which might be driven by efforts that reduce discrimination toward black students in school discipline or by greater acceptance of different within-race social or cultural norms.

One caveat is that, although detailed offense types were used to classify offenses, they may not fully capture the severity of offenses. The estimated disparities might be due to differences in unobserved severity of offenses (within offense types) committed by students in different racial groups or economically disadvantaged status. For example, they may reflect differences between the severity of fighting committed by black students and the severity of fighting committed by white students. Without more detailed data to describe offenses, it is hard to separate this possibility from the possibility that principals exhibit bias when assigning punishment.

In our data, suspension decisions are observed only if reported by school administrators, and they are made by those administrators only if students were caught and referred by teachers to the principals' office. Hence, the estimated disparities might reflect differences in unobserved heterogeneity of students in different racial groups or economically disadvantaged status, which is caused by the selection of students into observed punishment. Future work may attempt to check or address this concern using a different empirical strategy with more detailed data. Additional exploration of racial bias in the catching process or teacher's referral decisions could be made with the appropriate data.

Table 2.1: Sample Means of Offending Students' Characteristics

	All Offenses	Offenses Punished by Not OSS	OSS
Race			
White	0.373	0.414	0.278
Black	0.470	0.427	0.568
Hispanic	0.094	0.097	0.087
Asian	0.005	0.005	0.004
Multi-Racial	0.039	0.040	0.036
American Indian	0.019	0.016	0.025
Other Race	0.001	0.000	0.001
Disability			
No Disability	0.776	0.785	0.755
Physical Disability	0.088	0.087	0.089
Intellectual Disability	0.136	0.128	0.155
Other Dichotomous Characteristics			
(omitted: alternative group)			
Economically Disadvantaged	0.722	0.701	0.770
Female	0.300	0.310	0.277
Old in the Grade	0.261	0.237	0.317
Limited English Proficiency	0.043	0.043	0.044
Academically and Intellectually Gifted - Reading	0.034	0.040	0.023
Academically and Intellectually Gifted - Math	0.040	0.046	0.027
Repeating Grade in the Academic Year	0.098	0.083	0.132
Mean of Lagged Scores			
Lagged Normalized Math Score	-0.181	-0.152	-0.247
Lagged Normalized Reading Score	-0.180	-0.153	-0.245
Lagged Score Missing Indicator	0.165	0.158	0.180
Grade level			
Grade 3	0.033	0.032	0.032
Grade 4	0.040	0.039	0.042
Grade 5	0.048	0.046	0.051
Grade 6	0.110	0.111	0.109
Grade 7	0.133	0.133	0.132
Grade 8	0.133	0.130	0.140
Grade 9	0.200	0.190	0.219
Grade 10	0.135	0.137	0.130
Grade 11	0.098	0.104	0.086
Grade 12	0.072	0.077	0.059
Observations	5,249,004	3,668,573	1,580,431

Note: This table reports sample means of student characteristics for all offenses, offenses that were not punished by out-of-school suspension, and offenses that were punished by out-of-school suspension. For grade 10-12 students, End-of-Course Test English 1 is used for the calculation of lagged reading scores; End-of-Course Test Algebra 1 is used for the calculation of lagged math scores. Lagged test scores are normalized to have a zero mean and standard deviation of one among the students who took the same tests across the state.

Table 2.2: Percentages of Out-of-School Suspension for Selected Offense Types

		Percent Out-of-School Suspended among				
Offense Type	Number of incidents	All Students	Black	White	ED	NED
Possession of marijuana	15,429	84.5	85.5	83.2	84.3	84.8
Communicating threats	31,724	72.6	75.9	67.6	72.8	71.7
Disorderly conduct	41,957	49.0	54.0	37.2	51.1	41.3
Fighting	263,069	84.4	86.2	80.4	84.6	83.4
Harassment - verbal	22,451	36.1	40.6	31.3	37.1	33.5
Harassment - sexual	18,234	68.1	72.2	61.2	68.8	66.6
Aggressive behavior	265,033	44.7	51.2	35.3	46.7	38.3
Honor code violation	26,085	12.4	14.4	11.1	13.4	11.2
Dress code violation	79,824	16.7	23.7	8.0	18.9	11.2
Inappropriate language/disrespect	333,039	35.6	41.3	28.3	37.0	31.5
Insubordination	547,936	28.1	32.6	21.0	29.2	24.8
Theft	58,793	57.8	65.4	47.6	58.7	54.7
Disruptive behavior	964,658	24.1	28.8	17.1	25.5	19.7
Assault on student	27,423	77.5	80.9	73.5	78.3	74.8
Disrespect of faculty/staff	254,482	35.5	38.1	31.0	35.9	34.3
Excessive display of affection	15,800	15.7	29.9	9.6	17.9	11.6
Excessive tardiness	392,700	7.9	10.8	5.0	9.3	5.6
Truancy	41,943	21.1	25.9	12.8	22.4	18.6
Leaving class without permission	54,386	14.6	18.4	9.0	16.1	10.9
Skipping Class	267,308	16.2	21.2	10.8	17.7	12.8
Late to class	204,124	2.9	6.0	1.4	3.7	1.7
Skipping school	50,002	30.3	36.6	25.3	32.9	25.9
All offenses	529,004	30.1	36.4	22.4	32.1	24.9

Note: This table reports percentages of offenses that were punished by out-of-school suspension for several most common offense types (and all offenses). The last four columns report percentages of offenses that were punished by out-of-school suspension for black students, white students, Economically Disadvantaged students (ED) and Non-economically Disadvantaged students (NED).

Table 2.3: Estimation Results: Out-of-school Suspension for All Offenses

	(1)		(2)	
Dependent Variable: OSS or Not	Coeff.	SE	Coeff.	SE
Black	0.086^{***}	(0.003)	-0.003^{***}	(0.001)
Hispanic	0.037^{***}	(0.002)	0.003^{***}	(0.001)
Asian	0.020^{***}	(0.004)	0.001	(0.003)
Multiracial	0.037^{***}	(0.002)	0.004^{***}	(0.001)
American Indian	0.141^{***}	(0.014)	0.006^{**}	(0.002)
Other Race	0.170^{***}	(0.016)	0.010	(0.009)
Female	-0.005^{***}	(0.001)	0.001	(0.001)
Economically Disadvantaged	0.019^{***}	(0.001)	0.015^{***}	(0.001)
Physical Disabled	-0.015^{***}	(0.002)	-0.016^{***}	(0.002)
Intellectual Disabled	-0.007^{***}	(0.002)	-0.010^{***}	(0.002)
Old in the grade	0.035^{***}	(0.002)	0.022^{***}	(0.001)
Limited English Proficiency	-0.002	(0.002)	-0.002	(0.001)
AIG Reading	-0.011^{***}	(0.003)	-0.007^{***}	(0.001)
AIG Math	-0.015^{***}	(0.002)	-0.011^{***}	(0.002)
Lagged Normalized Math Score	-0.003^{***}	(0.001)	-0.002^{***}	(0.000)
Lagged Normalized Reading Score	-0.005^{***}	(0.001)	-0.004^{***}	(0.000)
Lagged Score Missing Indicator	0.027^{***}	(0.002)	0.018^{***}	(0.001)
Repeating Grade in the Academic Year	0.070^{***}	(0.003)	0.027^{***}	(0.001)
Constant	0.507^{***}	(0.024)	0.438^{***}	(0.021)
Observations	5249004		5249004	
Adjusted R^2	0.244		0.338	
School-Year Fixed Effect	No		Yes	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for the effects of student characteristics (β_1) (except for the categorical variable "grade") in equation 2.1 using the sample of all offenses. Results without (1) or with school-year fixed effects (2) are separately reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.4: Estimation Results: Out-of-school Suspension for First Offense

	(1)		(2)	
Dependent Variable: OSS or Not	Coeff.	SE	Coeff.	SE
Black	0.096***	(0.003)	0.004***	(0.001)
Hispanic	0.047***	(0.003)	0.009***	(0.002)
Asian	0.011*	(0.005)	−0.003	(0.004)
Multiracial	0.034***	(0.003)	0.004*	(0.002)
American Indian	0.114***	(0.011)	0.005	(0.004)
Other Race	0.194***	(0.018)	0.040**	(0.012)
Female	−0.006***	(0.001)	0.002**	(0.001)
Economically Disadvantaged	0.022***	(0.002)	0.017***	(0.001)
Physical Disabled	−0.007**	(0.002)	−0.013***	(0.002)
Intellectual Disabled	−0.014***	(0.002)	−0.016***	(0.003)
Old in the grade	0.027***	(0.002)	0.020***	(0.001)
Limited English Proficiency	−0.007*	(0.003)	−0.003	(0.002)
AIG Reading	−0.013***	(0.003)	−0.005*	(0.002)
AIG Math	−0.012***	(0.002)	−0.011***	(0.002)
Lagged Normalized Math Score	−0.002*	(0.001)	0.001	(0.001)
Lagged Normalized Reading Score	−0.010***	(0.001)	−0.010***	(0.001)
Lagged Score Missing Indicator	0.031***	(0.002)	0.021***	(0.001)
Repeating Grade in the Academic Year	0.096***	(0.008)	0.025***	(0.004)
Constant	0.779***	(0.034)	0.600***	(0.040)
Observations	848205		848205	
Adjusted R^2	0.307		0.424	
School-Year Fixed Effect	No		Yes	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for the effects of student characteristics (β_1) (except for the categorical variable "grade") in equation 2.1 only using the sample of first offenses of students. Results without (1) or with school-year fixed effects (2) are separately reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.5: Estimation Results: Out-of-School Suspension by School Level and Year

	2009	2010	2011	2012	2013	2014	2015
	Coefficients for Black Student Indicator (ref: white student)						
Elementary	0.008	0.010**	−0.003	−0.001	0.001	0.006	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	79238	95161	102221	98233	96369	74393	86399
Middle	−0.002	−0.009***	−0.010***	−0.008***	−0.008***	−0.008**	−0.007**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
	262317	318351	327655	326797	297112	211332	230492
High	0.005*	−0.001	−0.003	−0.002	−0.002	−0.002	−0.004
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
	362350	431128	440785	447487	414206	261705	285126
	Coefficients for Economically Disadvantaged Indicator						
Elementary	0.017***	0.018***	0.018***	0.022***	0.020***	0.019***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
	79238	95161	102221	98233	96369	74393	86399
Middle	0.008***	0.017***	0.014***	0.018***	0.021***	0.018***	0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
	262317	318351	327655	326797	297112	211332	230492
High	0.006***	0.016***	0.014***	0.016***	0.010***	0.016***	0.012***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	362350	431128	440785	447487	414206	261705	285126

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for being a black student (relative to a white student) or being an Economically Disadvantaged student, separately estimated for different school levels in different school years (for example, “2009” represents school year “2008-2009”). School fixed effects are used for the estimation. Results for elementary, middle and high schools use samples spanning grades 3-5, 6-8 and 9-12 respectively. The third number in each cell is number of observations. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school level.

Table 2.6: Estimation Results: Black White Differences in OOS by Offense Type

	(1)		(2)		
Offense Type	Coeff.	SE	Coeff.	SE	Observations
Possession of marijuana	0.019^{*}	(0.008)	-0.001	(0.007)	15429
Communicating threats	0.053^{***}	(0.007)	-0.015^{*}	(0.007)	31724
Disorderly conduct	0.112^{***}	(0.011)	0.006	(0.005)	41957
Fighting	0.065^{***}	(0.004)	0.011^{***}	(0.002)	263069
Harassment-Verbal	0.080^{***}	(0.010)	-0.020^{*}	(0.009)	22451
Harassment-Sexual	0.112^{***}	(0.010)	0.025^{*}	(0.010)	18234
Aggressive behavior	0.117^{***}	(0.005)	0.003	(0.003)	265033
Honor Code Violation	0.027	(0.017)	-0.004	(0.004)	26085
Dress code violation	0.116^{***}	(0.010)	0.009^{**}	(0.003)	79824
Inappropriate language/disrespect	0.094^{***}	(0.005)	-0.015^{***}	(0.002)	333039
Insubordination	0.096^{***}	(0.006)	-0.007^{***}	(0.002)	547936
Theft	0.150^{***}	(0.006)	0.033^{***}	(0.006)	58793
Disruptive behavior	0.093^{***}	(0.004)	-0.005^{***}	(0.001)	964658
Assault on student	0.068^{***}	(0.008)	-0.000	(0.007)	27423
Disrespect of faculty/staff	0.063^{***}	(0.006)	-0.024^{***}	(0.002)	254482
Excessive Display of Affection	0.174^{***}	(0.012)	0.047^{***}	(0.010)	15800
Excessive tardiness	0.043^{***}	(0.007)	-0.001	(0.001)	392700
Truancy	0.104^{***}	(0.019)	-0.013^{*}	(0.005)	41943
Leaving Class without permission	0.068^{***}	(0.006)	-0.013^{**}	(0.004)	54386
Skipping Class	0.084^{***}	(0.006)	-0.005^{*}	(0.002)	267308
Late to class	0.035^{***}	(0.005)	0.000	(0.001)	204124
Skipping School	0.084^{***}	(0.015)	0.003	(0.006)	50002
School-Year Fixed Effect	No		Yes		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for being a black student (relative to a white student), separately estimated for different types of offenses. Results without (1) or with school-year fixed effects (2) are separately reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.7: Estimation Results: Economically Disadvantaged Students' Differences in OSS by Offense Type

	(1)		(2)		
Offense Type	Coeff.	SE	Coeff.	SE	Observations
Possession of marijuana	-0.013	(0.007)	0.006	(0.006)	15429
Communicating threats	0.005	(0.007)	0.016*	(0.006)	31724
Disorderly conduct	0.059***	(0.009)	0.016**	(0.005)	41957
Fighting	0.015***	(0.002)	0.011***	(0.002)	263069
Harassment-Verbal	0.017*	(0.008)	0.006	(0.008)	22451
Harassment - sexual	-0.002	(0.010)	0.007	(0.009)	18234
Aggressive behavior	0.043***	(0.004)	0.023***	(0.002)	265033
Honor Code Violation	0.010	(0.009)	0.004	(0.003)	26085
Dress code violation	0.016***	(0.005)	0.006*	(0.002)	79824
Inappropriate language/disrespect	0.029***	(0.003)	0.018***	(0.002)	333039
Insubordination	0.023***	(0.003)	0.014***	(0.001)	547936
Theft	0.038***	(0.005)	0.039***	(0.005)	58793
Disruptive behavior	0.020***	(0.002)	0.013***	(0.001)	964658
Assault on student	0.031***	(0.007)	0.021**	(0.006)	27423
Disrespect of faculty/staff	0.015***	(0.004)	0.018***	(0.002)	254482
Excessive Display of Affection	0.007	(0.007)	0.003	(0.005)	15800
Excessive tardiness	0.013***	(0.002)	0.002*	(0.001)	392700
Truancy	0.007	(0.009)	0.007	(0.004)	41943
Leaving Class without permission	0.010*	(0.004)	0.004	(0.003)	54386
Skipping Class	0.013***	(0.002)	0.003	(0.002)	267308
late to Class	0.003**	(0.001)	0.001	(0.001)	204124
Skipping School	0.025***	(0.007)	0.012**	(0.004)	50002
School-Year Fixed Effect	No		Yes		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for the effects of being Economically Disadvantaged, separately estimated for different types of offenses. Results without (1) or with school-year fixed effects (2) are separately reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.8: Estimation Results: Black White Differences in OOS by Students' Gender

	Female Students			Male Students		
Offense Type	Coeff.	SE	Sample	Coeff.	SE	Sample
Possession of marijuana	0.019	(0.020)	2514	−0.004	(0.008)	12915
Communicating threats	0.012	(0.018)	8902	−0.026**	(0.009)	22822
Disorderly conduct	−0.000	(0.012)	11681	0.003	(0.007)	30276
Fighting	0.015***	(0.004)	76696	0.008***	(0.002)	186373
Harassment-Verbal	0.002	(0.020)	6331	−0.022	(0.011)	16120
Harassment - sexual	−0.002	(0.055)	1318	0.024*	(0.011)	16916
Aggressive behavior	0.005	(0.006)	63804	0.003	(0.003)	201229
Honor Code Violation	−0.008	(0.007)	9886	−0.002	(0.006)	16199
Dress code violation	0.004	(0.003)	35181	0.011*	(0.005)	44643
Inappropriate language/disrespect	−0.020***	(0.004)	90140	−0.015***	(0.003)	242899
Insubordination	−0.009**	(0.003)	164518	−0.006**	(0.002)	383418
Theft	0.039**	(0.013)	16765	0.030***	(0.007)	42028
Disruptive behavior	−0.002	(0.003)	242226	−0.006***	(0.001)	722432
Assault on student	0.010	(0.018)	7270	−0.001	(0.009)	20153
Disrespect of faculty/staff	−0.031***	(0.005)	79967	−0.022***	(0.003)	174515
Excessive Display of Affection	0.025	(0.014)	7471	0.059***	(0.013)	8329
Excessive tardiness	0.001	(0.002)	157007	−0.003*	(0.001)	235693
Truancy	−0.023**	(0.007)	15979	−0.010	(0.007)	25964
Leaving Class without permission	−0.015*	(0.007)	18462	−0.014**	(0.005)	35924
Skipping Class	−0.009**	(0.003)	96808	−0.004	(0.003)	170500
late to Class	−0.000	(0.002)	83510	0.000	(0.001)	120614
Skipping School	−0.003	(0.010)	17060	0.008	(0.008)	32942

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for the effects of being a black student, separately estimated for female students and male students. Separate estimates for different types of offenses are reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.9: Estimation Results: Out-of-school Suspension for First Offense by Offense Type

	Black Indicator		ED Indicator		
Offense Type	Coeff.	SE	Coeff.	SE	Observations
Possession of marijuana	0.010	(0.017)	-0.001	(0.016)	3040
Communicating threats	-0.013	(0.022)	0.014	(0.020)	5693
Disorderly conduct	0.042^{**}	(0.014)	0.034^{**}	(0.012)	6654
Fighting	0.015^{***}	(0.004)	0.009^{**}	(0.003)	71033
Harassment-Verbal	0.010	(0.022)	-0.009	(0.017)	4874
Harassment - sexual	0.007	(0.030)	-0.002	(0.027)	3318
Aggressive behavior	0.009^{*}	(0.005)	0.015^{***}	(0.004)	60960
Honor Code Violation	-0.009	(0.007)	0.003	(0.004)	9201
Dress code violation	0.003	(0.004)	-0.000	(0.003)	14325
Inappropriate language/disrespect	-0.003	(0.005)	0.016^{***}	(0.004)	49253
Insubordination	-0.001	(0.004)	0.012^{***}	(0.003)	58889
Theft	0.042^{***}	(0.012)	0.034^{***}	(0.009)	14100
Disruptive behavior	0.007^{*}	(0.003)	0.013^{***}	(0.002)	127621
Assault on studen	-0.001	(0.020)	0.036^{*}	(0.017)	5626
Disrespect of faculty/staff	-0.010	(0.006)	0.026^{***}	(0.005)	27732
Excessive Display of Affection	0.061^{**}	(0.022)	-0.008	(0.010)	4383
Excessive tardiness	0.002	(0.001)	0.001	(0.001)	49329
Truancy	-0.011	(0.009)	-0.001	(0.006)	7700
Leaving Class without permission	-0.010	(0.009)	-0.005	(0.007)	6656
Skipping Class	-0.008^{*}	(0.004)	0.002	(0.003)	38846
late to Class	-0.002	(0.002)	0.003^{*}	(0.001)	23351
Skipping School	-0.018	(0.010)	0.024^{**}	(0.008)	8416

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results for the effects of being a black student or being an Economically Disadvantaged student, using a sample of students' first offenses. Separate estimates for different types of offenses are reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.10: Estimation Results: Black White and ED Differences in OOS by Principal's Race

	Black Stud.×Black Prin.		ED Stud.×Black Prin.		
Offense Type	Coeff.	SE	Coeff.	SE	Sample
Possession of marijuana	0.006	(0.016)	−0.018	(0.016)	11838
Communicating threats	−0.002	(0.016)	−0.014	(0.017)	26514
Disorderly conduct	0.009	(0.013)	0.006	(0.013)	33898
Fighting	0.002	(0.005)	−0.004	(0.004)	214477
Harassment-Verbal	−0.034	(0.024)	0.010	(0.021)	18466
Harassment - sexual	−0.015	(0.025)	0.000	(0.024)	14688
Aggressive behavior	−0.009	(0.006)	0.000	(0.006)	218065
Honor Code Violation	−0.014	(0.013)	−0.010	(0.012)	20943
Dress code violation	0.009	(0.010	0.014	(0.008)	65071
Inappropriate language/disrespect)	−0.002	(0.006)	−0.005	(0.005)	274782
Insubordination	0.001	(0.005)	0.002	(0.004)	451649
Theft	0.009	(0.014)	0.007	(0.013)	46966
Disruptive behavior	−0.005	(0.003)	0.001	(0.003)	795732
Assault on student	0.016	(0.017)	−0.003	(0.015)	22332
Disrespect of faculty/staff	0.005	(0.006)	−0.003	(0.006)	210974
Excessive Display of Affection	0.032	(0.034)	0.024	(0.023)	12138
Excessive tardiness	−0.001	(0.003)	0.002	(0.003)	309605
Truancy	0.005	(0.016)	−0.001	(0.013)	33305
Leaving Class without permission	−0.025	(0.013)	0.019^{*}	(0.009)	44467
Skipping Class	0.006	(0.006)	−0.005	(0.005)	209684
late to Class	0.005	(0.005)	0.007	(0.005)	169864
Skipping School	0.010	(0.015)	0.002	(0.013)	37207

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS estimates for the coefficients on an interaction term that interacts the black or white student indicator with a black or white principal indicator and an interaction term that interacts economically disadvantaged student indicator with the black or white principal indicator. Separate estimates for different types of offenses are reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table 2.11: Estimation Results: Black White and ED Differences in OOS by Assistant Principal's Race

	Black Stud.×Black A.P.		ED Stud.×Black A.P.		
Offense Type	Coeff.	SE	Coeff.	SE	Sample
Possession of marijuana	0.013	(0.020)	−0.003	(0.019)	12070
Communicating threats	0.009	(0.018)	−0.020	(0.017)	26832
Disorderly conduct	−0.005	(0.013)	−0.027*	(0.013)	34578
Fighting	0.008	(0.005)	−0.008*	(0.004)	216832
Harassment-Verbal	−0.016	(0.029)	0.021	(0.024)	18698
Harassment - sexual	−0.010	(0.027)	0.031	(0.024)	14901
Aggressive behavior	0.008	(0.006)	−0.013*	(0.006)	220730
Honor Code Violation	0.013	(0.013)	0.005	(0.013)	21355
Dress code violation	0.005	(0.010)	0.007	(0.008)	66838
Inappropriate language/disrespect	0.000	(0.006)	−0.009	(0.005)	278903
Insubordination	−0.001	(0.004)	−0.008*	(0.004)	460288
Theft	0.001	(0.015)	−0.002	(0.013)	47661
Disruptive behavior	0.003	(0.004)	0.004	(0.003)	807995
Assault on student	−0.021	(0.018)	−0.014	(0.015)	22637
Disrespect of faculty/staff	0.008	(0.006)	−0.008	(0.006)	213456
Excessive Display of Affection	−0.017	(0.024)	0.014	(0.016)	12246
Excessive tardiness	0.003	(0.003)	0.002	(0.003)	316800
Truancy	−0.006	(0.017)	−0.015	(0.011)	34240
Leaving Class without permission	−0.024	(0.014)	0.023	(0.012)	45352
Skipping Class	0.009	(0.006)	0.000	(0.005)	213676
late to Class	−0.004	(0.005)	0.013*	(0.007)	170989
Skipping School	0.013	(0.016)	0.023	(0.014)	37732

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS estimates for the coefficients on an interaction term that interacts the black or white student indicator with a black or white assistant principal indicator and an interaction term that interacts the economically disadvantaged student indicator with the black or white assistant principal indicator. Separate estimates for different types of offenses are reported. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

CHAPTER 3

THE INFLUENCE OF ENDOGENOUS BEHAVIORS AMONG SOCIAL PAIRS: SOCIAL INTERACTION EFFECTS OF SMOKING (WITH DONNA GILLESKIE)

3.1 Introduction

In economics, social interactions reflect the interdependence among individuals when the characteristics and behaviors of one influences the preferences, beliefs, and constraints of another (Durlauf and Ioannides 2010). The impacts of social contacts on individual smoking behavior are of particular interest to public health and health economics researchers (e.g., Clark and Etilé 2006; Christakis and Fowler 2008; Cutler and Glaeser 2010). A common empirical hurdle in this application, and in the study of social interactions generally, is identification of endogenous social interaction effects, i.e., the causal effects of social contacts' smoking behaviors on an individual's smoking behavior (Manski 1993). The interest in quantifying endogenous social interaction effects stems from a desire to better understand smoking behavior as well as to prescribe effective anti-smoking policy. The existence of endogenous social interaction effects suggests that a policy intervention with a direct impact on an individual's smoking behavior will also have an indirect impact on the individual's social contacts' behaviors and may spread through the social network (Cutler and Glaeser 2010). Reliable measures of endogenous social interaction effects and hence, the "social multiplier" effects, provide policymakers with an important component involved in cost-effective analyses of anti-smoking policies.

Studies that attempt to measure endogenous social interaction effects on smoking behavior often rely on strong assumptions due to the difficulty of disentangling endogenous social interaction effects from other factors (e.g., Christakis and Fowler 2008). In particular, to identify how social contacts' smoking behaviors affect an individual's smoking behavior, researchers must address three important considerations that may lead one to misinterpret an observed correlation as causal.

These considerations are: 1.) the simultaneous influence of an individual's smoking behavior and her social contact's behavior on each other; 2.) an observed peer relationship that developed because of individual characteristics that also influence smoking (non-smoking), such as they both value health less (more); and 3.) dependence of smoking behaviors on other (exogenous) factors, such as cigarette prices or smoking advertisements. These problems are documented in the literature as *simultaneity*, *homophily* and *confounding*, respectively (e.g., Christakis and Fowler 2008; Bramoullé, Djebbari, and Fortin 2009; De Giorgi, Pellizzari, and Redaelli 2010).

We identify endogenous social interaction effects on smoking behavior while carefully addressing these considerations. To address the simultaneity problem, we model an individual's observed smoking behavior as the optimal solution to a discrete choice problem that depends on her own marginal utility of smoking and her observed social contacts' smoking behaviors, which are being chosen by those individuals at the same time. According to this game theoretical model, an individual and her (one) social contact's smoking behaviors are jointly described as a 2-player simultaneous move game (e.g., between the individual and his/her spouse, between two friends, between two siblings, or between a parent and an adult child) with complete information. The game is assumed to have a Pure Strategy Nash Equilibrium, and the player's decisions are assumed to lead to one of the Pure Strategy Nash Equilibria. We address, in estimation, the multiple equilibria problem by estimating the probability of each equilibrium.[1]. Identification is achieved using exclusion restrictions, which differ across different types of relationships, for the individual's smoking behavior and her social contacts' smoking behavior. For example, for spouses, we use information that affects smoking behaviors of each member of the couple before their marriage and assume that this information has no effect on their smoking behavior after marriage, once we control for own past behaviors. Furthermore, to disentangle the endogenous social interaction effects from homophily and confounding, our model of individual smoking behaviors over multiple periods allows that behavior to depend on one's previous smoking behaviors, permanent individual unobserved heterogeneity, and unobserved time-varying factors.

[1]The idea follows Bajari, Hong, and Ryan (2010).

In addition, we study how social contacts' health shocks affect an individual's smoking behavior. The endogenous health shocks (i.e., a cardiovascular disease event) are modeled as a function of individuals' past smoking behaviors, one's history of health shocks, other observed individual characteristics and unobserved heterogeneity. We use high-normal blood pressure (i.e., systolic blood pressure between 130 and 139 mm Hg or diastolic blood pressure between 85 and 89 mm Hg) as an exclusion restriction for onset of cardiovascular disease. The relationship between high-normal blood pressure and a cardiovascular disease event was not well known during much of the time span of our data; therefore, we assume that neither the knowledge of nor the experience of high-normal blood pressure influenced individuals' smoking behavior directly.

The data we use to measure social interaction effects are from the Framingham Heart Study (FHS) with its complementary social networks data (FHS-Net). A unique feature of these data is that researchers followed individuals and identified social contact relationships, such as spouses, friends, siblings, or parents, over time, which is important for our identification strategy. While previous users of these data have studied social interactions in the context of smoking (Christakis and Fowler 2008), our research is distinct from their research because we explicitly address, in estimation, concerns about simultaneity and homophily in order to identify endogenous social interaction effects.

We summarize our findings here. First, we find statistically significant endogenous social interaction effects for some paired contacts, namely spouses and friends. Our results suggest that a wife who does not smoke decreases her husband's propensity to smoke by 7.8 percentage points relative to a smoking wife; a husband who does not smoke decreases his wife's propensity to smoke by 8.5 percentage points relative to a smoking husband; and a friend who does not smoke decreases her friend's propensity to smoke by 4.2 percentage points. A sibling's or parent's smoking behavior does not have statistically significantly effects, but the parent of a non-smoking child (of adult age) is less likely to smoke.[2] We find that the estimates for endogenous social interaction effects exhibit a large positive bias when we fail to control for simultaneity or homophily. For example,

[2]The marginal effects are simulated assuming the focal individual (or ego) smoked last period.

the marginal effect of a wife on a husband is inflated by about 50 percent if simultaneity is not addressed. With no controls for homophily due to unobserved heterogeneity (even after controlling for past smoking behavior), the estimates would suggest statistically significant (yet upwardly biased) marginal effects (7.5 and 4.4 percentage points) of a parent's smoking behavior on a child's behavior or a sibling's smoking behavior on another sibling's behavior when a causal effect does not exist.

In addition, we find that the effects of social contacts' cardiovascular disease shocks on individual smoking behavior are not statistically significant for each of the types of relationships we study. These results are consistent with Khwaja, Sloan, and Chung (2006), who find no statistically significant effects of spousal health shocks on own smoking, and consistent with Darden and Gilleskie (2016), who find no statistically significant effects of parents' health shocks on children.

The rest of the chapter is organized as follows. Section 3.2 describes the basic empirical model that solves the 2-person simultaneous move game. In Section 3.3, we describe the FHS and FHS-Net data and replicate the findings in Christakis and Fowler (2008). Section 3.4 provides and discusses our empirical results.

3.2 Basic Empirical Model

3.2.1 Basic Setup

In this section we describe a model of smoking behavior with social interactions. Consider a set of individuals $i = 1, 2, \ldots, I_t$ at time t. Individuals simultaneously choose smoking actions; each individual i chooses a smoking action y_{it} from the action space Y_{it}. For simplicity, we let $Y_{it} = \{-1, 1\}$, where $y_{it} = -1$ denotes the action "Not Smoke" and $y_{it} = 1$ denotes the action "Smoke" in period t. Each period, the individuals are embedded in a social network, which indicates different types of relationships (or ties) between individuals; ties are fixed (for biological relationships) or pre-determined (for non-biological relationships) before individuals choose a period t smoking action. We do not explicitly model the network formation in this paper to keep the model parsimonious.[3]

At each period t, an individual i chooses her smoking action to maximize her current period utility. While we allow past behavior to influence current utility and we assume that current behavior is maximized each period based on updated information, we avoid modeling expectations of future unknowns by assuming individuals are myopic.[4] We use $-i$ to denote the set of i's social contacts (i.e., the individuals linked with i), and we use y_{-it} to denote i's social contacts' smoking actions at time t. A utility function, $V_i(y_{it}, y_{-it}, \Omega_{it}, \Omega'_{-it}, \mu_{it}, \epsilon_{it}^{y_{it}})$, describes the individual's payoff of each smoking action. Following Brock and Durlauf (2001), utility is decomposed into three additive parts such that

$$V_i(y_{it}, y_{-it}, \Omega_{it}, \Omega'_{-it}, \mu_{it}, \epsilon_{it}^{y_{it}}) = U_i(y_{it}, \Omega_{it}, \mu_{it}) + S_i(y_{it}, y_{-it}, \Omega'_{-it}) + \epsilon_{it}^{y_{it}} \tag{3.1}$$

where $U_i(y_{it}, \Omega_{it}, \mu_{it})$ captures the individual's private utility; $S_i(y_{it}, y_{-it}, \Omega'_{-it})$ captures her social utility (i.e., utility impacted by social interactions); and $\epsilon_{it}^{y_{it}}$ represents choice-based unobserved (by the econometrician) utility, which might be unobserved private or social utility. Information available at the point of decisionmaking is denoted by the vector $\Omega_{it} = [A_{it}, H_{it}, X_{it}, \xi_i]$. The variable A_{it} measures individual i's smoking stock at time t and, in our empirical estimation, we use y_{it-1} and ever smoked before last period to summarize A_{it}.[5] The vector H_{it} is a set of i's observed endogenous characteristics that affect her utility gain (or loss) from smoking at time t; in the empirical analysis, H_{it} represents i's current health status or occurrence of an unhealthy event.

[3]See Gilleskie and Zhang (2010), Badev (2013) and Goldsmith-Pinkham and Imbens (2013) who jointly model network formation as an identification strategy. In the empirical estimation section, we explain how our empirical results are not biased by the effects of endogenous network formation (homophily) given our carefully controlled empirical estimation strategy.

[4]Dynamic games that involve forward-looking decisionmaking are beyond the scope of this paper.

[5]Theoretically and biologically, the additive stock encompasses the amount of cotinine (i.e., metabolized nicotine) in the body at time t. Economists have approximated this stock by one's observed history of smoking behavior. For example, Darden, Gilleskie, and Strumpf (2017) use years of experience, years of continuous duration, years of cessation, and age of initiation. Given the lag between observed behaviors in our data, we choose to summarize one's history of smoking behavior by the last observed smoking status and ever smoked before last period.

The vector X_{it} is the set of observed exogenous individual and environmental covariates. This vector includes demographic variables as well as exogenous variables that may affect an individual's utility gain (or loss) from smoking at time t or an individual's stochastic health evolution (to be described below).[6] The information vector also includes individual time-invariant characteristics (ξ_i) that are unobserved by the econometrician but are known by individuals. Additionally, some time-varying environmental factors, such as anti-smoking policies, are captured by μ_{it} and are unobserved by the econometrician but known by individuals. Similar to Ω_{it}, Ω'_{-it} is a subset of i's social contacts' information variables and may affect i's social utility of choosing to smoke or not.

The endogenous characteristic H_{it} (e.g., an unhealthy event, in our case) is a function of an individual's exogenous characteristics and health and smoking stocks entering period t and is observed prior to making the current period smoking decision. It represents the flow of health in period t. One's health status entering period t is

$$H_{it} = F_i(H^e_{it}, A_{it}, y_{-it-1}, X^H_{it}, \xi_i, \epsilon^H_{it}) \tag{3.2}$$

and depends on the individual's health stock (H^e_{it}) which is summarized (in estimation) by an indicator of whether or not the individual has "ever had a chronic health condition". This health status in t also depends on one's own lagged smoking behavior and smoking history (A_{it}) as well as the most recent smoking behavior of one's social contacts (y_{-it-1}). We also allow health to depend on observed exogenous variables and unobserved individual time-invariant characteristics. Conditional on these variables, health evolution is uncertain; ϵ^H_{it} is an unanticipated health shock to individual i at time t.

Timing in this model is summarized as follows:

1. entering period t, an individual knows the smoking history (or smoking stock, A_t) and the

[6]Theory suggests avenues through which particular variables (e.g., cigarette prices) may impact behavior. Other variables may impact health production exclusively. When relevant, we use vectors X^S_{it} and X^H_{it} to differentiate these exogenous variables, assuming that either vector always includes the individual demographic variables.

health stock (H_t^e) of herself and her potential social contacts.[7] She also knows the exogenous characteristics (X_t) and individual unobserved heterogeneity (ξ) of all players;

2. after experiencing the uncertain health shock (ϵ_t^H), the health statuses of all individuals (H_t) for period t are realized;

3. theoretically, non-biological relationships of the social network are formed at this point;

4. unobserved utility of each smoking action ($\epsilon_t^{y_t}$) and environmental factors (μ_t) are realized;

5. knowing one's social contacts who may directly impact utility, all individuals simultaneously choose a smoking action.

Note that, at the point of decisionmaking, the stochastic terms (ϵ_t and μ_t) and permanent heterogeneity (ξ) are unobserved by the econometrician, but observed by the individuals. Therefore, individuals play complete information games within their network each period. A pure-strategy Nash Equilibrium at time t is a profile of actions $y_t \in Y, Y = Y_{1t} \times Y_{2t} \times \ldots \times Y_{I_t}$, such that for all i at time t, and for $y'_{it} \neq y_{it}$, $V_i(y_{it}, y_{-it}, \Omega_{it}, \Omega'_{-it}, \mu_{it}, \epsilon_{it}^{y_{it}}) \geq V_i(y'_{it}, y_{-it}, \Omega_{it}, \Omega'_{-it}, \mu_{it}, \epsilon_{it}^{y'_{it}})$. When we take the model to the data for estimation, we assume observed smoking decisions in the data are pure-strategy Nash Equilibrium outcomes. Multiple equilibria are a typical feature of this game. We discuss our identification strategy for addressing multiple equilibria in Section 3.2.6.

3.2.2 Model Specification

We specify an individual's private utility of smoking as:

$$U_i(y_{it}, \Omega_{it}, \mu_{it}) = (\alpha_{1i} + \alpha_{2i}A_{it} + \alpha_{3i}H_{it} + \alpha_{4i}X_{it}^S + \rho_i^S\xi_i + \mu_{it})\mathbf{1}[y_{it} = 1] + \alpha_{5i}A_{it} + \alpha_{6i}A_{it}^2 \quad (3.3)$$

where the indicator function, $\mathbf{1}[y_{it} = 1]$, takes on the value one when an individual chooses to smoke and zero when an individual chooses not to smoke. The specification has some standard

[7] In the empirical model, we include the health status of the social contact as a determinant of own utility. Individual smoking behavior may respond to the smoking behavior of a spouse (or relative or friend) as well as to their health (or death). Such responses may indicate learning (about the risks of smoking) or altruism (Darden and Gilleskie 2016).

features of the addictive goods literature (Becker and Murphy 1988). The coefficient α_{1i} captures the unconditional (average) utility gain of smoking, which is forsaken if the individual chooses not to smoke (i.e., withdrawal effects); α_{2i} describes how one's history of smoking affects utility gain (or loss) from smoking (i.e., reinforcement effect); α_{3i} and α_{4i} are coefficients on endogenous variables H_{it} and exogenous covariates X^S_{it} that shift preferences for smoking; ρ^S_i is a coefficient (i.e., factor loading) on the unobserved factor $\xi_i = [\xi_{i1}, \xi_{i,2}, \ldots, \xi_{i,M}]$ where the distribution of ξ_i is estimated discretely by M mass points (explained later); and α_{5i} and α_{6i} capture the current impact of one's history of smoking independent of the current smoking action (i.e., tolerance effects).[8]

To compare with the existing literature and to keep the estimation strategy parsimonious, we separately estimate social interaction effects for spouses, friend pairs, sibling pairs, and parent-child pairs. We parameterize an individuals' social utility as:

$$S_{ij}(y_{it}, y_{jt}, \Omega'_{jt}) = \frac{1}{2}(\beta_{1ij} y_{it} y_{jt} + \beta_{2ij} y_{it} H_{jt} + \beta_{3ij} y_{it} X^S_{jt}) \tag{3.4}$$

where coefficient β_{1ij} captures effects on i from social contact j's current smoking action (i.e., "endogenous" or peer effects) and coefficient vectors β_{2ij} and β_{3ij} capture effects on i from social contact j's health status given i's action, reflecting either learning or altruism effects and exogenous characteristics (i.e., exogenous effects), respectively. Recall that the smoking action indicator, $y_{kt}, k = i, j$, takes on the values 1 and -1 rather than the typical 0/1 indicator function. Theoretically, the sign of β_{1ij} should not be restricted; $\beta_{1ij} > 0$ means that i's and j's actions are strategic complements, and $\beta_{1ij} < 0$ means that they are strategic substitutes.[9] Since existing empirical results indicate a positive sign (i.e., conformity), we assume that this coefficient is non-negative in our empirical estimation part. Identification could also be achieved when the coefficient is negative, but this scenario is beyond the scope of this paper.

[8] The estimated coefficients have an i subscript, suggesting that there is individual heterogeneity in returns (i.e., estimated as random coefficients, perhaps). In practice, we allow the coefficients to vary by observed position in a relationship (i.e., husband or wife), and restrict the notation appropriately in Section 3.2.3.

[9] See Bulow, Geanakoplos, and Klemperer (1985), or Bramoullé, Kranton, and D'amours (2014).

3.2.3 Smoking Outcomes for Relationship Pairs

Suppose the utility gain (or loss) of choosing to smoke is

$$y_{it}^{*} = V_i(y_{it} = 1, y_{-it}, \Omega_{it}, \Omega_{-it}, \mu_{it}, \epsilon_{it}^{1}) - V_i(y_{it} = -1, y_{-it}, \Omega_{it}, \Omega_{-it}, \mu_{it}, \epsilon_{it}^{-1}).$$

Then, the observed smoking behavior of individual i at time t can be explained by $P(y_{it} = 1) = P(y_{it}^{*} > 0)$. Specifically, for each pairwise relationship where position in the relationship (e.g., husband or wife, parent or adult child) is denoted by subscripts a and b, the simultaneously-determined latent constructs are:

$$\begin{aligned}
&y_{it}^{*} = \alpha_{1a} + \alpha_{2a}A_{it} + \alpha_{3a}H_{it} + \alpha_{4a}X_{it}^{S} + \beta_{1a}y_{jt} + \beta_{2a}H_{jt} + \beta_{3a}X_{jt}^{S} + \rho_a^S\xi_i + \mu_t + \epsilon_{it}^{S} \\
&y_{jt}^{*} = \alpha_{1b} + \alpha_{2b}A_{jt} + \alpha_{3b}H_{jt} + \alpha_{4b}X_{jt}^{S} + \beta_{1b}y_{it} + \beta_{2b}H_{it} + \beta_{3b}X_{it}^{S} + \rho_b^S\xi_j + \mu_t + \epsilon_{jt}^{S} \\
&y_{it} = \begin{cases} 1, \text{if } y_{it}^{*} > 0 \\ -1, \text{otherwise} \end{cases} \text{ and } y_{jt} = \begin{cases} 1, \text{if } y_{jt}^{*} > 0 \\ -1, \text{otherwise} \end{cases}
\end{aligned} \tag{3.5}$$

and $\epsilon_{it}^{S} = \epsilon_{it}^{1} - \epsilon_{it}^{-1}$, $\epsilon_{jt}^{S} = \epsilon_{jt}^{1} - \epsilon_{jt}^{-1}$.

We point out three areas where our empirical specification differs from its theory-based framework. First, we retain the individual-specific subscripts on dependent and independent variables, yet acknowledge that we consider pairs of social contacts. That is, we do not model the simultaneous actions of numerous social contacts. We also use subscripts a and b on coefficients to indicate the mean effects on each "type" of individual (e.g., husband or wife, parent or adult child) within the social pair. We offer no explicit notation to indicate that two individuals are in the same network (i.e., are a pair) in order to avoid confusion. The effects of one's different social contacts may vary by the pair relationship (e.g., spouses, friends) as well as one's type within that relationship. Thus, in this paper, we identify mean effects for each relationship pair (i.e., "the spouse effect" or "the friend effect") and we do not restrict the effects to be symmetric within the pair. That is, the effect of husbands on wives ("the husband effect") may be different from the effect of wives on husbands ("the wife effect").[10] Second, theory suggests that time-varying environmental factors such as the price of cigarettes or the sentiment toward smoking in a particular location may impact smoking decisions. In estimation, variation in such variables is often used to identify smoking behaviors. We cannot rely on this source of variation since individuals in our sample are (generally) in the same town over time. We assume that the same, time-varying unobserved environmental factors, μ_t, impact all individuals at each period t; the associated terms amount to year indicators. Third, we allow the permanent individual unobserved components (ξ_i and ξ_j) to be correlated with each other. Thus, our estimation strategy addresses biases from homophily when measuring social interaction effects by controlling for social contacts' exogenous characteristics, by controlling for (and modeling) their endogenous observables, such as their lagged smoking status, and by allowing individual unobservables to be correlated.

3.2.4 Health Transitions and Mortality

The probability of an adverse health event for individual i is:

$$\begin{aligned} &Pr(H_{it} = 1 | H^e_{it}, A_{it}, y_{jt-1}, X^H_{it}, \xi_i) \\ &= \frac{exp(\gamma_{1a} + \gamma_{2a} H^e_{it} + \gamma_{3a} A_{it} + \gamma_{4a} y_{jt-1} + \gamma_{5a} X^H_{it} + \rho^H_a \xi_i)}{1 + exp(\gamma_{1a} + \gamma_{2a} H^e_{it} + \gamma_{3a} A_{it} + \gamma_{4a} y_{jt-1} + \gamma_{5a} X^H_{it} + \rho^H_a \xi_i)} \end{aligned} \tag{3.6}$$

where one's histories of smoking (A_{it}) and of health (H^e_{it}, ever had a chronic health condition) impact health transitions. We also allow the recent smoking behavior of one's social contact to influence own health events. The vector of exogenous individual variables are superscripted by H (X^H_{it}) to indicate the inclusion of an exogenous, time-varying health-related variable.

Since the average age of individuals in our data sample is high and attrition is mostly explained by death, we explicitly model mortality (jointly with the smoking behaviors and health events) to address non-random attrition. We define an indicator for death by period $t+1$ (D_{it+1}) conditional on being alive in period t, where the probability of death depends on current adverse health events, one's health history, and past and current smoking behaviors. We also allow one's social contact's

[10]As specified, the coefficients are free to vary in these aforementioned ways, but can be restricted in estimation (by the econometrician) to be the same depending on the context or what type of mean effects we want to achieve.

current smoking behavior to influence mortality. Specifically, the probability of death is

$$\begin{aligned} &Pr(D_{it+1} = 1 | H_{it}, H^e_{it}, A_{it}, y_{it}, y_{jt}, X^H_{it}, \xi_i) \\ &= \frac{exp(\omega_{1a} + \omega_{2a} H_{it} + \omega_{3a} H^e_{it} + \omega_{4a} A_{it} + \omega_{5a} y_{it} + \omega_{6a} y_{jt} + \omega_{7a} X^H_{it} + \rho^D_a \xi_i)}{1 + exp(\omega_{1a} + \omega_{2a} H_{it} + \omega_{3a} H^e_{it} + \omega_{4a} A_{it} + \omega_{5a} y_{it} + \omega_{6a} y_{jt} + \omega_{7a} X^H_{it} + \rho^D_a \xi_i)} \end{aligned} \tag{3.7}$$

where the vector of exogenous individual variables (X^H_{it}) is potentially the same set of variables that impact adverse health events, despite the dependence of mortality on those events themselves.[11]

3.2.5 Initial Conditions

Theory suggests that smoking behaviors and health outcomes depend on lagged variables in important ways. To test this relationship empirically, we must observe the actions and health events of individuals over time. The FHS data allows us to observe individuals eight times (over approximately 38 years).[12] Accompanying this wealth of data is the common occurrence that individuals in the survey have non-zero values of the lagged variables when we initially observe them. Because we initially observe individuals at an age when they may have already engaged in smoking behavior (i.e., have a non-zero addictive stock) or have experienced a chronic health condition, we must account for the endogeneity of these initially-observed values. That is, initially-observed smoking and health may be correlated with time-invariant determinants of subsequent smoking actions and health events even conditional on one's smoking history. For example, individuals who value health less may be more likely to smoke at any age, to experience an adverse health event at any time t, or to die with higher probability than someone who values health more. To address this endogeneity, we jointly model (with the dynamic smoking, health event, and death probabilities) static probabilities for the initial conditions. Importantly, we condition the initial response

[11] Identification is achieved through variation in the history of exogenous time-varying variables facilitated through mortality's dependence on observed current adverse health events and smoking behaviors (Arellano and Bond 1991).

[12] The Original cohort of the FHS is surveyed almost every two years since 1948; in 2013, almost 100 participants were still alive. We use only those exams from the Original cohort that overlap with exams from the Offspring cohort.

probabilities on individual unobserved heterogeneity.[13]

We model the initial smoking state (i.e., ever smoked up to $t = 1$ or age of smoking initiation) for i as follows:

$$y^*_{i0} = \alpha^0_{1a} + \alpha^0_{2a} X^S_{i0} + \alpha^0_{3a} X^S_{j0} + \rho^{S0}_a \xi_i + \epsilon^S_{i0}$$
$$y_{i0} = \begin{cases} 1, \text{if } y^*_{i0} > 0 \\ -1, \text{otherwise} \end{cases} \tag{3.8}$$

and, for the other individual in the pair (j), the initial equation is analogous. Note that characteristics of one's social contact may influence initial smoking behavior.[14] Proper identification of these initial conditions requires that the vector of exogenous individual characteristics (X^S_{i0}) include a variable that affects initial smoking, but does not affect subsequent smoking conditional on one's smoking history.[15]

The probability of the initial health state of individual i (i.e., an adverse health event entering period $t = 1$) is:

$$Pr(H_{i1} = 1 | X^H_{i0}, y_{i0}, y_{j0}, \xi_i) = \frac{exp(\gamma^0_{1a} + \gamma^0_{2a} X^H_{i0} + \gamma^0_{3a} y_{i0} + \gamma^0_{4a} y_{j0} + \rho^{H0}_a \xi_i)}{1 + exp(\gamma^0_{1a} + \gamma^0_{2a} X^H_{i0} + \gamma^0_{3a} y_{i0} + \gamma^0_{4a} y_{j0} + \rho^{H0}_a \xi_i)}\ . \tag{3.9}$$

and depends on the initial smoking status of the individual as well as that of her social contact. Analogous functions are used to define health state probabilities for individual j of the relationship pair.[16] Similarly, identification requires a variable in X^H_{i0} that is excluded from the vector of exogenous individual characteristics that affect subsequent health events and mortality. Dependence of

[13] We detail assumptions necessary for estimation in Section 3.2.6. Modified methods of Wooldridge (2005) could be used for an unbalanced panel such as ours, but finite sample bias was found under some situations (e.g., Akay 2012; Rabe-Hesketh and Skrondal 2013).

[14] Because we model initial smoking prior to age 19 as the initial condition and we drop couples who were married before age 19, we do not include spouse characteristics in the initial smoking probability. We do, however, allow spouse's lagged smoking behavior to impact initial health since it represent health events in the recent past.

[15] We return to this discussion when we introduce our empirical solution for multiple equilibria.

[16] We assume one's prior history of health shocks is zero when we first observe them (i.e., $H^e_{i1} = 0$).

these initial conditions on individual unobserved heterogeneity (ξ_i) addresses potential endogeneity bias in estimated effects of lagged smoking behaviors and health events in the jointly estimated dynamic probabilities of smoking and health transitions.[17]

3.2.6 Additional Considerations for Estimation

Multiple Equilibria

The likelihood function, comprised of the probabilities defined in equations 3.5-3.9 and the observed individual outcomes over time, is an "incomplete" discrete econometric model according to the existing econometrics literature (Tamer, 2003). Directly applying a maximum likelihood method produces inconsistent estimates (Heckman 1978; Maddala 1983). In a game theory framework, the "incompleteness" is caused by the existence of multiple equilibria in the model. That is, the relationship between the covariates and error terms and the observed outcome is a correspondence, not a function (e.g., Bjorn and Vuong 1984; Bresnahan and Reiss 1991; Tamer 2003).

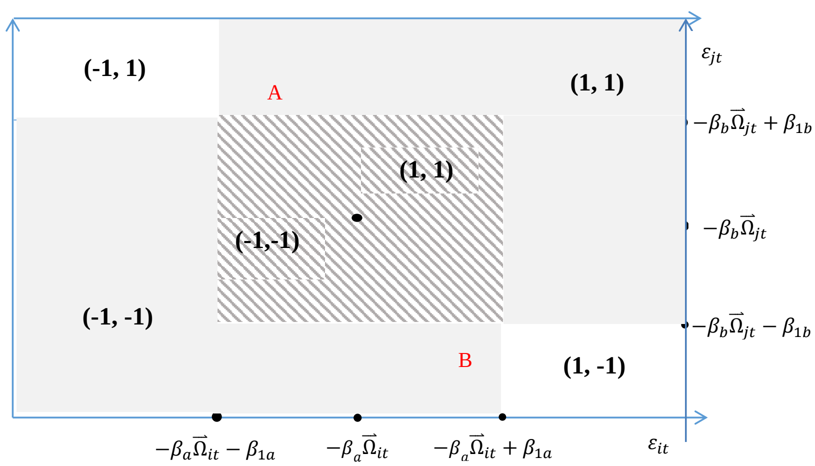
To deal with the multiple equilibria problem, several approaches have been proposed in the existing literature. The first approach is to find a common feature of all equilibria and change the model into one that predicts this feature (e.g., Bresnahan and Reiss 1991). The second approach is to specify a selection rule for the multiple equilibrium (e.g., Bjorn and Vuong 1984; Soetevent and Kooreman 2007). The third approach is to use upper and lower bounds of the choice probability to restrict the parameter estimates to a set and, with this, partially identify the parameters (e.g., Ciliberto and Tamer 2009).

Unlike the above mentioned approaches, we leave the probability of equilibrium selection as an empirical construct to be estimated. The idea is similar to Bajari et al. (2010), who estimate equilibrium selection mechanisms. To illustrate the idea, Figure 3.1 shows the equilibrium pattern for different draws of the smoking i.i.d. error terms (ignoring the evolution of health and death for this example). Suppose $\beta = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \beta_2, \beta_3, \rho^S, 1)$ and $\vec{\Omega}_{it} = (1, A_{it}, H_{it}, X_{it}^S, H_{jt}, X_{jt}^S, \xi_i, \mu_t)$

[17]The factor loading (ρ^S) on the permanent individual-level unobserved heterogeneity in the initial smoking probability (or age of smoking initiation) is normalized to one to satisfy an identification requirement for consistent estimation.

where an individual's social contact's smoking behavior (y_{jt}) is removed from the information vector and the marginal effect (β_1) is similarly removed from the coefficient vector. We see that events $\{y_{it}= 1$ and $y_{jt}= 1\}$ and $\{y_{it}= -1$ and $y_{jt}= -1\}$ may happen if $-\beta_a\vec{\Omega}_{it} - \beta_{1a} \leq \epsilon_{it} \leq -\beta_a\vec{\Omega}_{it} + \beta_{1a}$ and $-\beta_b\vec{\Omega}_{jt} - \beta_{1b} \leq \epsilon_{jt} \leq -\beta_b\vec{\Omega}_{jt} + \beta_{1b}$. Therefore, two equilibria, (1,1) and (-1,-1), exist if ϵ^S_{it} and ϵ^S_{jt} are drawn from this region (i.e., the hashed region in Figure 3.1). We assume that in the multiple equilibria region, outcome (1,1) is selected with probability $Pr(o = (1, 1))$, and thus outcome (-1,-1) is selected with probability $1-Pr(o = (1, 1))$. According to Bajari et al. (2010), β_{1a}, β_{1b} and $Pr(o = (1, 1))$ can be identified if we have exclusion restrictions that shift the smoking behavior of i (or j) but do not directly shift the smoking behavior of j (or i).[18]

Figure 3.1: Equilibrium Pattern for Error Space



Note: Blue region is the multiple equilibria region.

Identification

To identify β_{1a}, β_{1b} and $Pr(o = (1, 1))$, we use different exclusion restrictions for different types of relationship. For spouses, we use factors that shift the wife's (or the husband's) smoking status before her (or his) marriage as exclusion restrictions. Specifically, we use the proportion of individuals in the same birth year, same sex cohort of the wife (or the husband) who ever smoked

[18] $Pr(o = (1, 1))$ as an additional parameter is only identified if neither β_{1a} nor β_{1b} is zero. One could also allow the probability of the equilibrium selection to depend on individuals' observed or unobserved characteristics, but the additional parameters are hard to identify in our model.

before age 19 to capture the effect of cohort-specific factors, such as values toward smoking or the cost of smoking for the specific age-sex cohort. Since the husband and the wife belong to different age-sex groups, the values are different for the husband and the wife, and they separately shift initial smoking status (i.e., ever smoked before age 19) of the husband and the wife, which leads to different smoking stocks that separately shift their smoking behaviors after marriage. We dropped a small portion of the sample who married before 19 to guarantee that the factors shift behavior before marriage.[19] For siblings and friends, since it is more likely that they belong to the same cohort, we use smoking status of excluded social contacts as an additional exclusion restriction. The idea is similar to the literature that uses characteristics of excluded peers as instruments to deal with identification problems (Bramoullé et al. 2009; De Giorgi et al. 2010). Age of the individual also serves as an exclusion restriction, with the assumption that it does not directly shift the social contact's smoking behavior.

Identification of the effect of health status on smoking behavior requires a variable that alters health events but does not directly shift smoking behavior conditional on the observed health status. Since our measure for health status is a cardiovascular disease shock, we use an indicator for high-normal blood pressure as an exclusion restriction. The risk of high-normal blood pressure on cardiovascular disease events was not well known at the time most of our data were collected.[20] Since individuals with high-normal blood pressure are unaware of their increased risk for an adverse cardiovascular health event, we assume that it does not directly affect own smoking behavior.

Distributional Assumptions

We use a full-information maximum likelihood method to jointly estimate the probabilities (in equations 3.5-3.9) of the behaviors and outcomes we observe. We assume ϵ_{it}^S, ϵ_{jt}^S ($t= 0, 1, \dots, T$) are i.i.d. idiosyncratic errors conditional on observed exogenous and health variables as well as

[19] We assume the proportion measure is not correlated with unobserved permanent heterogeneity of the individuals since it represents a population feature (i.e., no homophily for the entire age-sex cohort).

[20] See the research milestones of FHS and Vasan, Larson, Leip, Evans, O'Donnell, Kannel, and Levy (2001).

unobserved permanent individual and time-varying environmental factors, and each follows a normal distribution $N(0, 1)$. By assumption, the health transitions and mortality error terms ϵ_{it}^{H}, ϵ_{it}^{D} ($t= 1, 2, \ldots, T$) are i.i.d. and follow standard logistic distributions. These errors are not correlated with each other and other error terms in the system, conditional on observable characteristics, permanent individual unobserved heterogeneity ξ_i, and common unobserved heterogeneity μ_t.

Following an approach proposed and used by Heckman and Singer (1984), Mroz and Guilkey (1992) and Mroz (1999), we assume the joint distribution of ξ_i and ξ_j can be approximated by a set of discrete mass points and associated weights without imposing a specific distribution. A Monte Carlo analyses demonstrates that this flexible approach performs better in simultaneous equations than assuming, for example, a joint normal distribution for error terms that may not be normally distributed (Mroz 1999). Having been specific about assumptions made for distributions of unobservables, identification, and equilibrium selection, we specify the likelihood function in Appendix C.1.

3.3 Data and Construction of Variables

3.3.1 Framingham Heart Study

The data we use to estimate social interaction effects are from the Framingham Heart Study (FHS) with its complementary social network data (FHS-Net). This ongoing longitudinal study began in 1948 under the direction of the National Heart, Lung and Blood Institute (NHLBI) and has attained iconic status in epidemiological research. Much of the now-common knowledge concerning heart disease and detrimental effects of smoking were first discovered by this study.[21] The study follows several cohorts of residents (or former residents) of the town of Framingham, Massachusetts with periodic in-person health examinations and survey measures of other health-related information. The complementary social network data link the study participants by their social ties. The network information was organized and first used by Nicholas A. Christakis, James H. Fowler, and their co-authors. In this paper, we focus on spouses, friends, siblings, and parent-child pairs. One limitation of these data is that friendship ties were collected from handwritten tracking sheets,

[21] See texttthttps://www.framinghamheartstudy.org/about-fhs/research-milestones.php.

which were designed for the purpose of facilitating follow-up. Since a majority of individuals nominate only one friend, and many of the nominated friends are not participants in the FHS study, the friend ties are relatively sparse. Yet, missing network ties is a common problem for most network datasets.

The FHS surveys individuals from several different cohorts, labeled Original, Offspring, Third Generation, New Offspring Spouse, and Omni 1 and 2. The social network data focus mostly on the Offspring cohort and some of the waves (called Exams in the FHS) of the Original cohort. The exams we use span years 1971 to 2008. Exams are not performed annually. In our data, the average gap between two adjacent exams of an individual is 5 years (for the Offspring cohort) and 2 years (for the Original cohort), and the longest gap is 11 years.[22] In this paper, based on the availability of social network data, we use Exams 1-8 of the Offspring cohort and Exams 12, 16, 19, 21, 23, 24, 26, and 28 of the Original cohort, in which exam dates most closely correspond to those of the Offspring cohort.[23] For simplicity, we re-label the chosen exams of the Original cohort as Original Cohort Exam 1 to Exam 8, respectively. Because of the addictive nature of smoking and the infrequent fluctuation in smoking behavior, we assume the exam-specific observed smoking behavior is representative of behavior over the exam period and, based on this assumption, we treat observed smoking behaviors across individuals (pairs) as concurrent.[24]

The total number of distinct individuals among the original and offspring data with overlapping exams is 8533. We drop those individuals who never report a smoking status or only report smoking status in one exam, which results in 7090 distinct individuals. We also drop individuals who enter the study below age 19, who are older than age 70, or for whom information on initial smoking

[22] Additionally, within an exam wave, individuals are not examined or surveyed at the same time.

[23] Original Cohort Exam 29 is closer to Offspring Cohort Exam 8, but we do not have access to it. Instead, we use Original Cohort Exam 28. Although the Third Generation cohort and other Omni cohorts can be linked with the network data to some extent, we do not use them because either most of the relationships we explore cannot be identified or there are only one or two exams for these cohorts.

[24] Darden et al. (2017) report that, among the Original cohort, one quarter of men never smoke, a little over a quarter always smoke, and three-quarters of the men who smoke and quit do not restart. In fact, just over a third of men who smoked and quit after age 30 begin smoking again. The average length of cessation prior to relapse is a little over three years.

status is not available. These deletions result in 6480 distinct individuals with eight exams. In addition, some individuals may not report their smoking status in some of the exams; we drop these individual-exam observations.[25] Finally, some individuals die. Our estimation sample consists of an unbalanced panel with 37906 individual-exam observations.[26] We use these observations to construct social contact pairs.[27] The social interactions that we examine are among spouse, friend, sibling, and parent and adult child pairs. The number of paired exams (i.e., exams for which we observe the behaviors and outcomes of both individuals in the pair) are $9394, 4796, 15346,$ and 10044, respectively.

3.3.2 Construction of Key Variables

We seek to measure the influence of endogenous smoking behaviors among social pairs. To that end, we require an analogous description of smoking behavior at frequent intervals. In the FHS exams, participants were asked variants of the question "Have you smoked cigarettes regularly in the last year?" We use responses to these questions to define smoking actions (i.e., yes as 1, no as -1).[28] Initial smoking status (i.e., ever regularly smoked before age 19) is constructed using retrospective questions from Exams 7, 12 and 17 for those in the Original cohort and from Exam 1 for the Offspring cohort.[29]

Our time-varying health indicator is constructed using information on cardiovascular disease events that occurred at least one year prior to each examination and since one year prior to the last

---

[25] Non-response rates of participants to smoking questions are less than one percent.

[26] We impute some other explanatory variables by a multiple imputation method. Observations with any imputed explanatory variable are less than 10 percent of the sample.

[27] If one person has multiple same type social contacts, such as multiple friends, at the same time, they are regarded as different observations for the estimation. This happens very infrequently.

[28] It is possible to use responses to other smoking-related questions in subsequent exams to reconstruct smoking behavior in a missed exam. However, we do not perform such imputations in order to avoid potential bias. We do impute the smoking action for individuals in Original Cohort Exam 16 using retrospective questions from Exam 17 because smoking status was not asked in Exam 16.

[29] Some relationships (links) were initiated in the later exams (e.g., married in Exam 2 or new friendship in Exam 3). The initial smoking statuses of those individuals are coded as the smoking status in the exam that preceded the linkages.

examination. The cardiovascular disease events include: Myocardial Infarction, Angina Pectoris, Coronary Insufficiency, Cerebrovascular Accident (e.g., Atherothrombotic Infarction of Brain, Transient Ischemic Attack, Cerebral Embolism, Intracerebral Hemorrhage, Subarachnoid Hemorrhage), and Congestive Heart Failure. Blood pressure is recorded by two numbers: systolic (SBP) and diastolic (DBP) blood pressure readings. For our purposes, we define an indicator for high blood pressure if the individual has SBP$\geqslant$140mmHg or DBP$\geqslant$90mmHg or is under hypertension treatment. High-normal blood pressure indicates the individual does not have high blood pressure but pressure readings are 130mmHg$\geqslant$SBP$\geqslant$139mmHg or 85mmHg$\geqslant$DBP$\geqslant$89mmHg.[30] Table 3.1 reports summary statistics for the endogenous and exogenous variables used in the subsequent estimation section. The table also specifies (with an asterisk) which endogenous variables define the dependent variables in equations 3.5-3.9.

3.3.3 Replication of Previous Findings

The original assembler's of the network data (Christakis and Fowler, henceforth CF) reported their findings regarding whether an individual's social contact's quit behavior influenced own smoking behavior (Christakis and Fowler 2008). They found that a person's likelihood of smoking decreased by 67 percent if their spouse did not currently smoke versus smoked currently, by 25 percent among siblings, and by 43 percent among mutual friends.[31] In an effort to compare our findings from a more comprehensive econometric model with those of CF, we first must replicate their estimation sample of network pairs. Replication is often difficult because details of researcher decisions regarding sample construction are scarce. Our attempt to replicate the CF sample revealed that those authors used the responses to a question about an individual's smoking that asked "Usual number of cigarettes smoked (now or formerly)" to construct the smoking indicator at the first exam. They coded any positive response as one; zero otherwise. Because this question contains responses about former smoking behavior, it is likely to overstate smoking

[30] The American Heart Association defines prehypertension as a SBP between 120 and 139 mm Hg or a DBP between 80 and 89 mm Hg.

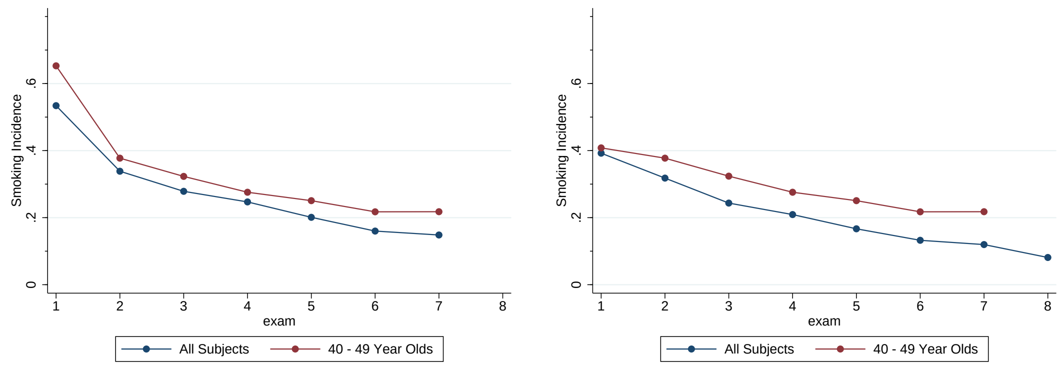
[31] CF arrive at these values by simulating smoking behavior of the egos at the mean values of the explanatory variables when the alter does and does not quit smoking.

Table 3.1: Summary Statistics for Individual-Exam Observations

Variable	Mean	SD
Endogenous Variables		
Smoke in t *	0.227	0.419
Smoke in $t-1$	0.266	0.442
Ever Regularly Smoked before $t-1$	0.417	0.493
Cardiovascular Disease Event between $t-1$ and t *	0.054	0.227
Ever Cardiovascular Disease before $t-1$	0.071	0.257
Death between t and $t+1$ *	0.064	0.244
Ever Regularly Smoked before age 19 ($t=0$) *	0.374	0.484
Cardiovascular Disease Event entering $t=1$ *	0.031	0.174
Exogenous Variables		
Demographics		
Female	0.551	0.497
Married	0.744	0.437
Age in years at t [range: 19,102]	57.353	15.134
Age in years at $t=1$ [range: 19,70]	44.005	13.667
Education		
No High School	0.055	0.228
Some High School	0.073	0.261
High School Graduate	0.336	0.472
College	0.402	0.490
Post-College	0.133	0.340
High Blood Pressure in $t-1$	0.360	0.480
High-Normal Blood Pressure in $t-1$	0.138	0.345
Proportion who Smoked before age 19 in Same Birth-year Same Sex Cohort	0.390	0.140
Individual-Exam Observations	37,906	

Note: * indicates variable is also modeled as a dependent variable. All variables in the table, except "Age in years", are indicator variables (0 or 1) .

in the current (first exam) period. Figure S1 of CF's supplementary appendix depicts the average smoking probabilities at each exam of individuals in their sample.[32] Figure 3.2a below depicts the smoking probabilities for our replicated sample when we make the same assignment mistake for first exam smoking behavior. Figure 3.2b depicts the smoking probabilities for our sample.



a.) Replicated Sample b.) Our Sample

Figure 3.2: Smoking Incidence Using Replicated Sample and Our Sample

A few additional differences between their sample and our sample exist. CF used up to seven observations on an individual while our sample includes eight. They also restricted their sample to individuals between age 21 and age 70 while we include individuals aged 19 and above. Their estimated smoking equations use only the "offspring cohort" as egos (i.e., the focal individuals). The linked individual ("alter") can be from any of the "Original Cohort," "Offspring Cohort," "Omni Cohort" and "Generation 3 Cohort." We use individuals from the "Original cohort" and "Offspring Cohort" to construct both egos and alters.

We now attempt to replicate CF's results. In estimation we use our replication of their sample, including the assignment error. We also provide results from estimation using our corrected, or preferred, sample. Their main reported findings regarding the behavioral impact of social contacts are derived from a logistic regression of ego's smoking status on alter's previous and current

[32]The supplementary appendix cited in Christakis and Fowler (2008) is available at `http://www.nejm.org/doi/suppl/10.1056/NEJMsa0706154/suppl_file/nejm_christakis_2249sa1.pdf`.

smoking status as well as ego's previous smoking status, ego's exogenous demographic characteristics, and exam indicators. They consider social pairs of the types: spouse, friend, sibling, co-worker, and neighbor. While CF account for clustering of standard errors since individuals are observed multiple times, they do not model the endogeneity of own previous smoking status or that of the social contact's current and previous smoking behavior. They claim that the inclusion of own lagged smoking behavior addresses serial correlation and that the inclusion of alter's smoking behavior addresses homophily. They also report impacts of a social contact's smoking behavior on own smoking using the estimated model's predictions of own behavior at the average values of the explanatory variables. As such, while they sometimes describe their findings as the ego's quitting response to an alter's "smoking cessation," their main findings, highlighted in their abstract, do not condition the alter's previous smoking to one and do not condition the ego's previous smoking to one. Hence, it is our opinion that the results do not reflect quit responses to social contact's quitting. Rather, from the best we can tell, their main reported results are percent changes in the ego's probability of smoking (unconditional on ego's previous smoking status) when an alter smokes and does not smoke currently (unconditional on alter's previous smoking status). We estimate their model using our replication of their miscoded sample and calculate marginal effects evaluated at the average values of all explanatory variables except for alter's current smoking behavior. We then report percent changes as CF do, rather than reporting the percentage point change. Because CF do not report the averages of ego's smoking probabilities when alter does and does not smoke, we cannot compare percentage point changes.

Table 3.2 provides the percent change in smoking behavior of the ego when the alter smokes versus does not smoke (i.e., $\frac{p(y_{it}=1|y_{jt}=1)-p(y_{it}=1|y_{jt}=0)}{p(y_{it}=1|y_{jt}=1)}$) for spouse, mutual friend, and sibling pairs. We reproduce the CF results in the first panel. Panel two details results using our replication of their sample, including its error, and CF's estimation and calculation procedure. The third panel of the table reports the percent changes using our preferred sample in estimation. Here, we use their estimation model and predict smoking probabilities as CF do. Finally, we report the percent changes using the same estimation procedure as CF on both our replicated CF sample and our preferred sample with calculations of the ego's smoking probabilities (when the alter does and does not smoke) conditional on the ego smoking in the previous period.

Table 3.2: Percentage Change in the Likelihood of Ego Smoking when Alter Smokes versus does not Smoke for Different Social Pairs

Sample	Spouses	Mutual Friends	Siblings
Christakis & Fowler (2008)	67	43	25
	[59,73]	[1,69]	[14,35]
	10,522	1,083	21,097
Replicated results			
Evaluated at sample means of all explanatory variables			
• Replicated Sample	65	39	30
	[59,72]	[7,72]	[21,39]
	10,762	1,338	18,157
• Our Sample	53	38	21
	[46,59]	[13,63]	[13,30]
	18,788	2,276	30,692
Evaluated at sample means of explanatory variables with ego smoking in $t-1$			
• Replicated Sample	26	16	11
	[23,30]	[2,29]	[8,15]
	10,762	1,338	18,157
• Our Sample	24	16	8
	[20,27]	[5,28]	[5,12]
	18,788	2,276	30,692

Note: 95% confidence intervals are reported in brackets below the percentage change figures.

Sample size is also reported.

The purpose of Table 3.2 is three-fold. First, a comparison of the results in panels one and two demonstrate that we do a pretty good job of replicating CF's results using a sample constructed like theirs and their estimation and percent change calculations. We calculate confidence intervals using 1000 bootstrapped replications, as CF does. Second, when we use our larger, preferred sample, we obtain similar, yet slightly smaller, results as CF. Third, because we wish to understand how social contacts' behaviors influence quitting behavior, we use our replicated CF sample and our preferred sample to calculate probabilities of ego's smoking conditional on the ego smoking in the previous period when the alter does and does not smoke contemporaneously. We find that the percent changes are considerably smaller, but it is difficult to say anything about a social contacts' influence on the behavior of previous smokers versus non-smokers because the baseline probabilities are so different. That is, these figures reflect percent changes and the unconditional probabilities of smoking are smaller than the probabilities of smoking if one has smoked in the previous year (i.e., persistence is significant). In the next section we present results from a different estimation procedure using our preferred sample. Our preferred estimation model addresses several concerns in the peer effects literature as well as important dimensions of dynamic models of smoking and health.

3.4 Results and Discussion

In this section, we present the results of our preferred model of the smoking behaviors of individuals and their social contacts where we specifically account for simultaneity, homophily, confounding, smoking dynamics, stochastic health, and endogenous initial conditions. We estimate the equation system (defined by equations 3.5-3.9) for spouse, friend, sibling, and parent-child pairs. We also discuss these results by comparing them with the results from specifications that do not fully address these concerns.

3.4.1 Effects of Own Observed Behavior and Characteristics

Table 3.3 provides the estimated effects of variables explaining smoking and health probabilities (equations 3.5 and 3.6) for each member of a spouse pair. We report other estimated parameters for spouse pairs and all parameters for friend, sibling, and parent-child pairs in the Appendix.[33] Models of addiction suggest that current smoking behavior depends on the history of one's own smoking behavior, which is summarized in our empirical model by indicators of smoking in the previous period and having ever regularly smoked. Any history of smoking increases the probability of smoking currently, which indicates a positive "reinforcement" effect. A recent history

[33] According to Christakis and Fowler (2013), friend pairs in their papers were defined by the question "please tell us the name of a close friend, to whom you are not related" with whom "you are close enough that they would know where you are if we can't find you." We believe that this question indicates that once a person reported another person's name, they should be mutual friends. Therefore, in our preferred estimation, unlike Christakis and Fowler (2008), we define friend pairs by nomination in either direction.

strengthens that relationship. Recall from section 3.2.1 that health shocks entering period t are observed prior to making a smoking decision. A recent health event (measured by having a cardiovascular disease event between $t-1$ and t) and one's blood pressure in the previous period capture health of the individual. Recent poor health decreases the probability of smoking. The effects of age and education are also provided in the table.

Smoking last period increases the probability of a cardiovascular disease event between period $t-1$ and t. Indicators of either high blood pressure or high-normal blood pressure in the previous period increase the probability of a health shock. Note that the significance of the high-normal blood pressure indicator is a necessary condition for identification of the health equation in this system of simultaneous equations.[34]

3.4.2 Effects of Social Contact's Observed Behavior

Important to this analysis, we examine the impacts of a social contacts' behavior, health, and exogenous characteristics on one's own smoking behavior. We find that the current smoking behavior of each spouse significantly affects the current smoking behavior of the other spouse. We find no evidence that a spouse's cardiovascular disease event affects one's own probability of smoking. However, a wife's previous smoking behavior does increase the probability of a cardiovascular disease event among husbands (but not vice versa).

The statistically significant positive estimates of the key parameter, β_1, suggest that spouse's smoking status leads an individual to choose the same smoking status. Using the estimated model, we calculate an ego's probability of smoking (i.e., smoking in the current period, given that he smoked in the previous period) when his alter smokes in the current period and when she does not smoke in the current period. Table 3.4 reports the differences in these probabilities (i.e., marginal effects) for several different versions of the model that successively include important econometric considerations: simultaneity of the ego's and the alter's smoking decision, correlated unobserved individual heterogeneity, endogeneity of the health shock, and non-random attrition due to death.

[34] We have verified the other necessary condition that high-normal blood pressure (in $t-1$) does not impact current smoking behavior conditional on the indicator for a health event between periods.

Table 3.3: Estimation Results: Smoking and Health Probabilities for Spouse Pairs

	Husband Equation		Wife Equation	
	Coeff	SE	Coeff	SE
Smoking at t (Equation 3.5)				
Smoke in $t-1$	1.004^{***}	0.031	0.768^{***}	0.030
Ever Regularly Smoked before $t-1$	0.499^{***}	0.039	0.311^{***}	0.041
Cardiovascular Disease Event between t and $t-1$	-0.224^{**}	0.102	-0.563^{***}	0.192
High Blood Pressure in $t-1$	-0.104^{**}	0.056	-0.179^{**}	0.079
(Age-19)/10 at t	-0.116	0.081	0.382^{***}	0.089
(Age-19)^2/100 at t	0.001	0.012	-0.087^{***}	0.014
No High School	-0.301^{**}	0.146	0.111	0.210
Some High School	-0.035	0.111	0.079	0.155
College	0.026	0.065	-0.113	0.080
Post-College	-0.102	0.088	-0.241	0.149
Social Contact:				
Smoke in t (β_1)	0.160^{***}	0.044	0.188^{***}	0.047
Cardiovascular Disease Event between t and $t-1$	0.156	0.155	0.125	0.109
High Blood Pressure in $t-1$	0.000	0.063	-0.037	0.067
No High School	0.085	0.158	-0.062	0.207
Some High School	0.009	0.129	0.290^{**}	0.129
College	-0.084	0.060	-0.259^{***}	0.089
Post-College	-0.327^{***}	0.114	-0.253^{*}	0.130
ρ^S	-0.572^{***}	0.167	0.267^{***}	0.052
Constant	1.296^{***}	0.179	1.690^{***}	0.174
Cardiovascular Disease Event between $t-1$ and t (Equation 3.6)				
Smoke in $t-1$	0.377^{***}	0.077	0.342^{**}	0.170
Ever Regularly Smoked before $t-1$	-0.121^{*}	0.063	-0.123	0.087
Ever Cardiovascular Disease before $t-1$	0.981^{***}	0.107	1.245^{***}	0.174
High Blood Pressure in $t-1$	0.527^{***}	0.111	0.927^{***}	0.175
High-Normal Blood Pressure in $t-1$	0.369^{***}	0.144	0.540^{**}	0.237
(Age-19)/10 at t	1.448^{***}	0.298	1.772^{***}	0.515
(Age-19)^2/100 at t	-0.106^{***}	0.032	-0.127^{**}	0.054
No High School	0.027	0.181	0.036	0.272
Some High School	0.010	0.159	0.418^{**}	0.216
Some High School	-0.093	0.108	-0.334^{**}	0.156
Post-College	-0.399^{**}	0.156	-0.555	0.377
Social Contact: Smoke in $t-1$	0.148^{**}	0.056	0.076	0.088
ρ_a^H	-0.155	0.101	0.011	0.233
Constant	-6.943^{***}	0.688	-8.909^{***}	1.256
Number of Pair-Exams:		9394		

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

Note: The table provides selected estimated parameters from the smoking and health event probabilities as part of a larger joint estimation of equations 3.5-3.9 for spouse pairs. Estimates for parameters in other equations of the jointly estimated set of equations are in the Appendix. The specification normalizes one mass point to zero ($\xi^1_{(i,j)} = 0$) and identifies another mass point, $\xi^2_{(i,j)} = -1.911$ with standard error 0.070. The joint probabilities for the mass point combinations are $P(\xi_i = 0, \xi_j = 0) = 0.16$, $P(\xi_i = -0.190, \xi_j = 0) = 0.081$, $P(\xi_i = 0, \xi_j = -0.190) = 0.479$, and $P(\xi_i = -0.190, \xi_j = -0.190) = 0.280$. Time-varying unobserved effects (μ_t for $t = 1, 2, \ldots, 8$) are controlled for by including exam indicators. The estimated probability of equilibrium (1,1) is $Pr(o = (1,1)) = 0.185$ (when the equilibrium is either (1,1) or (-1,-1)).

These specifications (in columns) allow us to evaluate aspects of the econometric modeling that address bias in estimates of a social contact's influence. The rows of Table 3.4 present the results for different relationship types. Specifically, we allow for different effects of a wife's behavior on a husband's behavior (labeled "Wife's (effect on) $\rightarrow$ Husband") and vice versa (labeled "Husband's (Effect on) $\rightarrow$ Wife"). We also examine the social interaction effects among mutual friends and siblings, as well as the effects of a parent's behavior on an adult child's behavior (labeled "Parent's (Effect on) $\rightarrow$ Child") and vice versa ("Child's (Effect on) $\rightarrow$ Parent").

The first column reports benchmark estimates from probit estimation of equation 3.5 without the multiple equilibria correction and without controlling for correlated individual unobserved heterogeneity ξ.[35] We find statistically significant marginal effects for all types of social contacts. For example, if the wife of a spousal pair does not smoke, her husband (who smoked last period) is 15.8 percentage points less likely to smoke this period.[36] To compare to the results in Table 3.2, this change reflects a 25.0 percent decrease in the husband's probability of smoking. Analogously, we measure the impact of a husband's smoking behavior on that of the wife, as the model allows for asymmetric effects within a social pair. We find that wives, who smoked in the previous period, are 14.8 percentage points (or 23.8 percent) less likely to smoke currently if the husband currently does not smoke versus smokes. These percent changes expand upon the composite spouse effect summarized in the last rows of Table 3.2.

We continue examining the impacts among spouse pairs across different estimators before discussing the impacts of other social pairs. The results in column 2 include a correction for multiple equilibria. This correction requires joint estimation of the two equations in the equation system (5), and solves the simultaneity problem. Note that the marginal effects are smaller (by a third) and remain statistically significant, which demonstrates that simultaneity creates significant upward

[35]This specification is equivalent to estimating the two equations in (5) separately. Note that the results in Table 3.2 reflect the combined estimation where social contact effects are forced to be the same.

[36]Expressed differently, if the wife of a spousal pair smokes, her husband is 15.8 percentage points more likely to smoke than if she does not smoked.

Table 3.4: Marginal Effects: Alter's Smoking Behavior on Ego's Smoking Behavior for All Social Pairs

	(1)	(2)	(3)	(4)	(5)
Wife's (Effect on) → Husband	0.158***	0.094***	0.081***	0.078***	0.078***
	(0.015)	(0.019)	(0.021)	(0.021)	(0.021)
Husband's (Effect on) → Wife	0.148***	0.102***	0.082***	0.083***	0.085***
	(0.015)	(0.019)	(0.022)	(0.022)	(0.021)
Friends	0.064***	0.040***	0.042***	0.043***	0.042***
	(0.014)	(0.012)	(0.015)	(0.016)	(0.015)
Siblings	0.068***	0.044***	0.010	0.011	0.010
	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)
Parent's (Effect on) → Child	0.089***	0.077***	0.002	0.002	0.001
	(0.014)	(0.022)	(0.277)	(0.252)	(0.322)
Child's (Effect on) → Parent	0.068***	0.023	0.032*	0.042***	0.042***
	(0.014)	(0.024)	(0.018)	(0.016)	(0.015)
Estimator addresses:					
Simultaneity	No	Yes	Yes	Yes	Yes
Unobserved Heterogeneity	No	No	Yes	Yes	Yes
Endogeneity of the Health Variable	No	No	No	Yes	Yes
Attrition Due to Death	No	No	No	No	Yes
Percent Change					
Wife's (Effect on) → Husband	25.0	17.1	17.3	16.8	16.5
Husband's (Effect on) → Wife	23.8	16.5	21.9	22.1	21.9
Friends	11.7	7.6	8.6	8.8	8.6
Siblings	11.7	7.8	2.9	2.9	2.8
Parent's (Effect on) → Child	14.0	12.0	0.4	0.4	0.3
Child's (Effect on) → Parent	14.6	5.3	12.4	16.5	15.5

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

Note: The marginal effects of alter's smoking behavior on ego's smoking behavior are simulated assuming ego smoked last period. For friends and siblings, the effects are assumed to be symmetric. Percentage point changes are reported in the top panel. Standard errors are calculated using 500 random draws from the variance-covariance matrix of estimated parameters. Percent changes are provided in the bottom panel.

bias in the marginal effects if not properly dealt with. Column 3 reports a specification that allows for individual unobserved heterogeneity ξ to address "homophily" and requires jointly estimating equations 3.5 and (8).[37] Column 4 provides results for a specification that adds the health equations 3.6 and 3.9 to the jointly estimated equation system; column 5 accounts for attrition due to death by adding equation 3.7. Note that the estimates for the marginal effects are robust with these model specifications. The results from our preferred model (column 5) suggests that a wife who does not smoke decreases her husband's probability of smoking by 7.8 percentage points (or 16.5 percent) over a wife who does smoke; a husband who does not smoke decreases his wife's probability of smoking by 8.5 percentage points (or 21.9 percent). These impacts, while still suggesting a conformative social interaction effect, are smaller than those estimated with a less-comprehensive estimator.[38]

Having discussed why we prefer the estimator in column 5 of Table 3.4, we examine the estimated marginal effects among other social pairs. We see that the decrease in smoking probabilities when a friend smokes and does not smoke falls from 6.4 percentage points (or 11.7 percent) to 4.2 percentage points (or 8.6 percent) as the estimator addresses additional econometric concerns. The decrease in smoking probabilities of siblings (or children) when the other sibling (or parents) does and does not smoke changes from being statistically significant (yet biased) at a 6.8 percentage point (8.9 percentage point) to being statistically insignificant. The largest changes happen from Column 2 to Column 3 (i.e., after addressing "homophily" due to the unobserved heterogeneity), which is consistent with the intuition that the "homophily" due to permanent unobserved factors plays a larger role for these types of relationships (siblings and parents/children). Interestingly,

[37] The best fit of our model (in terms of information criteria) has two latent factors for each person in the pair, so there are a total of four points of support in the discrete joint distribution. The results indicate that unobserved heterogeneity plays an important role as the estimated marginal effects decline even more. This finding is consistent with Clark and Etilé (2006).

[38] We also evaluated the marginal effects and percent changes at the observed values of the explanatory variables. That is, we did not condition on the ego having smoked in the previous period. Our analogous values are 4.2 percentage points (or 20.4 percent) and 4.8 percentage points (or 25.5 percent) for husband and wife egos, respectively. These results are presented in Appendix Table C8. We see that addressing simultaneity, homophily, health endogeneity, and non-random attrition is very important.

using our preferred model, we find that a non-smoking child statistically significantly decreases the smoking probability of his or her parents.

3.4.3 Effects of Social Contact's Observed Characteristics

Table 3.5 summarizes the estimated parameters for social contacts' exogenous characteristics or health status on one's smoking behavior (β_2 or β_3) for all social pairs. We find that an alter's education significantly decreases the ego's probability of smoking for several types of social pairs. We do not find evidence, for any relationship, that a social contact's cardiovascular disease shock affects one's probability of smoking. These results are consistent with Khwaja et al. (2006), who find no statistically significant effects of spousal health shocks on own smoking, and consistent with Darden and Gilleskie (2016), who find no statistically significant effects of parents' health shocks on children.

Table 3.5: Estimation Results: Alter's Characteristics on Ego's Smoking Behavior for All Social Pairs

Social Contact	Wife → Husband	Husband → Wife	Friends	Siblings	Parent → Child	Child → Parent
Female			-0.093	-0.011	0.053	0.022
			(0.087)	(0.038)	(0.064)	(0.068)
Married			-0.037	-0.008	-0.024	0.023
			(0.064)	(0.038)	(0.063)	(0.064)
No High School	0.085	-0.062	-0.083	0.233^{**}		
	(0.158)	(0.207)	(0.111)	(0.100)		
Some High School	0.009	0.290^{**}	-0.050	0.039	0.052	-0.875^{**}
	(0.129)	(0.129)	(0.092)	(0.070)	(0.082)	(0.343)
College	-0.084	-0.259^{***}	-0.098^{*}	-0.063	-0.148	-0.933^{***}
	(0.060)	(0.089)	(0.057)	(0.046)	(0.097)	(0.345)
Post-College	-0.327^{***}	-0.253^{*}	-0.161^{*}	0.044	-0.096	-0.860^{**}
	(0.114)	(0.130)	(0.091)	(0.062)	(0.164)	(0.349)
High Blood Pressure in $t-1$	0.000	-0.037	0.050	0.051	0.089^{*}	-0.030
	(0.063)	(0.067)	(0.052)	(0.038)	(0.050)	(0.096)
Cardiovascular Disease Between t and $t-1$	0.156	0.125	-0.061	0.132	0.058	0.209
	(0.155)	(0.109)	(0.109)	(0.080)	(0.067)	(0.300)
Number of Pair-Exams:	9394	9394	4796	15346	10044	10044

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

Note: This table reports effects of social contacts' characteristics on individual smoking behavior (β_2 or β_3 in equation 3.5) for each social pair from joint estimation of equations 3.5-3.9.

APPENDIX A

APPENDIX FOR CHAPTER 1

Table A1: Offense Categories and Types

Violence
1.Assault, Robbery and sexual offense
001 = Assault resulting in a serious injury
002 = Assault involving the use of a weapon
104 = Physical attack with a firearm or explosive device
003 = Assault on school personnel not resulting in a serious injury
071 = Assault on non-student w/o weapon and not resulting in serious injury
044 = Assault on student
072 = Assault on student w/o weapon and not resulting in serious injury
090 = Violent assault not resulting in serious injury
045 = Assault - other
103 = Robbery with a firearm or explosive device
010 = Robbery with a dangerous weapon
093 = Robbery without a weapon
016 = Kidnapping
023 = Extortion
015 = Taking indecent liberties with a minor
012 = Rape
013 = Sexual offense
014 = Sexual assault not involving rape or sexual offense
2.Threat and possession of a weapon
043 = Bomb threat
105 = Threat of physical attack with a firearm
106 = Threat of physical attack with a weapon
107 = Threat of physical attack without a weapon
019 = Communicating threats (G.S. 14-277.1)
008 = Possession of a firearm or powerful explosive
009 = Possession of a weapon (excluding firearms and explosives)
3.Disorderly conduct and harassment
022 = Disorderly conduct (G.S. 14-288.4(a)(6))
038 = Harassment - sexual
101 = Harassment - Racial
102 = Harassment - Disability
109 = Harassment - Sexual orientation
110 = Harassment - Religious affiliation
025 = Harassment - verbal
080 = Discrimination
4.Other violent behavior
027 = Aggressive behavior
021 = Affray (G.S. 14-33)
024 = Fighting
026 = Hazing
052 = Bullying
094 = Cyber-bullying
079 = Gang activity

Offense Categories and Types (Continue)

Drug

054 = Sale of controlled substance in violation of law - cocaine
055 = Sale of controlled substance in violation of law - marijuana
056 = Sale of controlled substance in violation of law - Ritalin
057 = Sale of controlled substance in violation of law - other
088 = Distribution of a prescription drug
049 = Use of controlled substances
048 = Use of alcoholic beverages
050 = Use of narcotics
070 = Use of tobacco
096 = Under the influence of controlled substances
095 = Under the influence of alcohol
005 = Possession of controlled substance in violation of law - cocaine
006 = Possession of controlled substance in violation of law - marijuana
007 = Possession of controlled substance in violation of law - Ritalin
017 = Possession of controlled substance in violation of law - other
020 = Alcohol Possession (G.S. 18B)
041 = Possession of tobacco
051 = Possession of chemical or drug paraphernalia
086 = Possession of student's own prescription drug
087 = Possession of another person's prescription drug

Disrespect

061 = UB: Disrespect of faculty/staff
032 = UB: Inappropriate language/disrespect
033 = UB: Insubordination

Property

053 = RO: Burning of a school building (G.S. 14-60)
036 = UB: Theft
039 = UB: Property damage
018 = UB: Unlawfully setting a fire (G.S. 14-277.1)

Truancy

075 = UB: Skipping school
030 = UB: Truancy
067 = UB: Leaving school without permission
074 = UB: Skipping class
066 = UB: Leaving class without permission

Other

1.Disruptive

042 = UB: Disruptive Behavior

2.Tardiness

064 = UB: Excessive tardiness
078 = UB: Late to class

3.Some other minor rule violations

028 = UB: Honor code violation
031 = UB: Dress code violation
060 = UB: Cell phone use
063 = UB: Excessive display of affection
068 = UB: Mutual sexual contact between two students
029 = UB: False fire alarm
035 = UB: Falsification of information
034 = UB: Gambling
059 = UB: Being in an unauthorized area
091 = UB: Misuse of school technology
040 = UB: Inappropriate items on school property
047 = UB: Use of counterfeit items
046 = UB: Possession of counterfeit items
065 = UB: No Immunization

4.Other

037 = UB: Bus misbehavior
077 = UB: Physical exam
114 = UB: Inappropriate Behavior
092 = UB: Repeat offender
058 = UB: Other School Defined Offense
069 = UB: Other

Table A2: General and Specific Deterrence Effects for Second Offense

Offense Category (or Type)	OLS	OLS &GFE	2SLS	2SLS &GFE	Sample Size	Mean of Dependent Var.
Property						
α_{21}	-0.005	-0.003	-0.016	0.005	24420	0.13
	(0.009)	(0.009)	(0.858)	(0.560)		
α_{22}	0.016***	0.014**	-1.307	-0.922		
	(0.005)	(0.005)	(2.767)	(1.684)		
Drug						
α_{21}	-0.009	-0.009	0.224	0.111	36487	0.2
	(0.010)	(0.010)	(0.584)	(0.426)		
α_{22}	-0.011*	-0.009*	-0.060	-0.123		
	(0.005)	(0.005)	(0.431)	(0.351)		
Fighting						
α_{21}	-0.023***	-0.024***	-0.282	-0.175	133341	0.16
	(0.006)	(0.006)	(3.387)	(2.858)		
α_{22}	-0.013***	-0.016***	4.770	4.435		
	(0.004)	(0.003)	(8.614)	(7.192)		
Aggressive Behavior						
α_{21}	-0.017***	-0.021***	-1.121	-1.478	117886	0.22
	(0.006)	(0.006)	(2.822)	(3.390)		
α_{22}	0.005*	0.001	1.166	1.552		
	(0.003)	(0.003)	(3.094)	(3.715)		
Disrespect to Faculty						
α_{21}	-0.030***	-0.031***	1.043	0.650	122166	0.28
	(0.006)	(0.006)	(1.279)	(0.854)		
α_{22}	0.002	-0.004	-1.265	-0.797		
	(0.003)	(0.003)	(1.507)	(1.017)		
Insubordination						
α_{21}	-0.081***	-0.084***	0.380	0.474	224651	0.37
	(0.006)	(0.006)	(0.440)	(0.451)		
α_{22}	0.002	-0.006**	-0.619*	-0.708**		
	(0.003)	(0.003)	(0.330)	(0.331)		
Inappropriate Language						
α_{21}	-0.011*	-0.015***	0.204	0.071	166167	0.26
	(0.005)	(0.005)	(0.357)	(0.286)		
α_{22}	0.007**	0.001	-0.801	-0.437		
	(0.003)	(0.003)	(0.645)	(0.514)		
Skipping Class						
α_{21}	-0.030***	-0.023***	-0.218	-0.211	124294	0.27
	(0.007)	(0.007)	(0.219)	(0.211)		
α_{22}	0.004	-0.003	0.339	0.181		
	(0.004)	(0.004)	(0.443)	(0.430)		
Disruptive Behavior						
α_{21}	-0.122***	-0.128***	-0.851**	-0.697**	358033	0.40
	(0.005)	(0.005)	(0.340)	(0.335)		
α_{22}	-0.024***	-0.030***	0.655	0.618		
	(0.002)	(0.002)	(0.502)	(0.493)		
Excessive Tardiness						
α_{21}	-0.075***	-0.069***	-14.748	43.990	111400	0.45
	(0.011)	(0.011)	(35.668)	(442.025)		
α_{22}	-0.018**	-0.022***	18.460	-56.236		
	(0.008)	(0.007)	(44.752)	(564.759)		

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

Note: This table reports estimates of the "general deterrence effect" (α_{21}) and "the specific deterrence effect" (α_{22}) in equation 1.2 for different categories (or types) of misbehaviors, estimated separately. Columns 2-5 show the results using different estimation methods. OLS: a model with all control variables and school fixed effects; OLS&GFE: adds an additional control for student unobserved heterogeneity; 2SLS: instruments the DPI variable but does not include a control for student unobserved heterogeneity; 2SLS&GFE: instruments the DPI variable and controls for student unobserved heterogeneity. The sample for each category does not include schools with less than 5 students' first offense observations of the category or without second offense observations of the category in the academic year. The mean of the dependent variable is reported in the last column. Standard errors are reported in parentheses and clustered at the student level.

Table A3: General Deterrence Effects for Second Offense by OSS Experience

Offense	Sample with OSS Experience		Sample without OSS Experience	
Category (or Type)	OLS & GFE	2SLS & GFE	OLS & GFE	2SLS & GFE
Property	0.012	0.257	-0.017	-0.056
	(0.013)	(0.353)	(0.013)	(0.234)
	12927	12927	11493	11493
Drug	0.006	0.110	-0.055***	-0.247
	(0.013)	(0.749)	(0.018)	(0.452)
	20907	20907	15580	15580
Fighting	-0.025***	1.754**	-0.056***	298.6
	(0.007)	(0.777)	(0.014)	(39616.6)
	114579	114579	18762	18762
Aggressive Behavior	0.003	-0.282	-0.044***	-0.108
	(0.009)	(0.213)	(0.009)	(0.323)
	51729	51729	66157	66157
Disrespect to Faculty	0.013	0.263	-0.062***	-0.079
	(0.011)	(0.227)	(0.008)	(0.137)
	41403	41403	80763	80763
Insubordination	-0.009	-0.113	-0.127***	-0.423
	(0.011)	(0.251)	(0.008)	(0.297)
	61334	61334	163317	163317
Inappropriate Language	0.021**	-0.024	-0.034***	-0.184*
	(0.009)	(0.195)	(0.007)	(0.107)
	59099	59099	107100	107100
Skipping Class	0.036**	-0.212	-0.024***	0.087
	(0.017)	(0.309)	(0.009)	(0.141)
	17558	17558	107881	107881
Disruptive Behavior	-0.060***	-0.556	-0.152***	-0.498***
	(0.008)	(0.473)	(0.007)	(0.177)
	83497	83497	278704	278704
Excessive Tardiness	-0.115***	2.775	-0.062***	0.659*
	(0.033)	(11.33)	(0.015)	(0.344)
	9028	9028	102372	102372

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

Note: This table reports estimates of α_{21} (general deterrence effects on students' second offense) in equation 1.2 by separately using the student-observations that were out-of-school suspended for their first offenses ($p_{1ist} = 1$) and the student-observations that were not out-of-school suspended for their first offense ($p_{1ist} = 0$). Standard errors are reported in parentheses and clustered at the student level. The third number in each cell is number of observations.

Table A4: First Stage Results for 2SLS&GFE Specification in Table 1.5

Offense	All Offenses	Violence	Disrespect	Truancy	Drug	Property
Coefficient on IV (γ_{11})	0.058***	0.056***	0.074***	0.051***	0.064***	0.017***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
F-Statistic	5261.360	4139.040	9069.750	2588.320	3567.880	349.142

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

Note: Standard errors are reported in parentheses.

Table A5: First Stage Results for 2SLS&GFE Specification in Table 1.8

Offense	Violence	Disrespect	Truancy
Dependent Variable: $\tilde{P}_{2st}$ (equation 1.4)			
Coefficient on the first instrumental variable (γ_{21})	-0.014^{***}	0.021^{***}	0.077^{***}
	(0.005)	(0.005)	(0.009)
Coefficient on the second instrumental variable (γ_{22})	0.092^{***}	0.034^{***}	-0.043^{***}
	(0.005)	(0.004)	(0.008)
Dependent Variable: P_{1ist} (equation 1.5)			
Coefficient on the first instrumental variable (γ_{31})	-0.019	0.017^{*}	0.050^{***}
	(0.012)	(0.010)	(0.001)
Coefficient on the second instrumental variable (γ_{31})	0.089^{***}	0.047^{***}	-0.025^{*}
	(0.011)	(0.009)	(0.015)

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

Note: Standard errors are reported in parentheses.

Table A6: First Stage Results for 2SLS&GFE Specifications in Table 1.10

	End-of-Grade Math		Dropout		ACT Composite Score	
	Ever OSS	First OSS	Ever OSS	First OSS	Ever OSS	First OSS
Coefficient on IV	0.051***	0.031***	0.091***	0.055***	0.082***	0.030*
	(0.006)	(0.005)	(0.018)	(0.016)	(0.020)	(0.017)
F-Statistic	74.106	34.464	26.333	11.683	17.036	2.881

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

Note: Standard errors are reported in parentheses.

Table A7: First Stage Results for 2SLS&GFE Specification in Table 1.11

	End-of-Grade Math				Dropout		ACT Composite Score	
	Grade 3-8		Grade 6-8					
	All	No Record	All	No Record	All	No Record	All	No Record
Coefficient on IV	0.069***	0.078***	0.038***	0.051***	0.087***	0.100***	0.063***	0.097***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)	(0.000)	(0.004)
F-Statistic	3974.170	2108.690	924.528	750.942	764.660	559.193	722.402	689.602

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

Note: Standard errors are reported in parentheses.

Table A8: Comparison of Sample Means of Student Characteristics for Each Sample in Table 1.8

	All Offenses	Violence	Disrespect	Truancy	Drug	Property
Race						
White	0.529	0.525	0.526	0.533	0.549	0.529
	(0.499)	(0.499)	(0.499)	(0.499)	(0.498)	(0.499)
Black	0.282	0.286	0.289	0.289	0.278	0.286
	(0.450)	(0.452)	(0.453)	(0.453)	(0.448)	(0.452)
Hispanic	0.114	0.114	0.112	0.105	0.102	0.111
	(0.318)	(0.318)	(0.315)	(0.307)	(0.302)	(0.314)
Asian	0.025	0.025	0.025	0.024	0.025	0.026
	(0.156)	(0.157)	(0.156)	(0.154)	(0.155)	(0.160)
Multi-Racial	0.036	0.035	0.035	0.033	0.032	0.034
	(0.185)	(0.185)	(0.183)	(0.179)	(0.176)	(0.182)
American Indian	0.014	0.014	0.014	0.014	0.014	0.012
	(0.117)	(0.117)	(0.117)	(0.118)	(0.119)	(0.108)
Other Race	0.001	0.001	0.001	0.001	0.001	0.001
	(0.0245)	(0.0247)	(0.0244)	(0.0249)	(0.0249)	(0.0248)
Disability						
No Disability	0.861	0.861	0.864	0.870	0.872	0.866
	(0.346)	(0.345)	(0.343)	(0.337)	(0.334)	(0.340)
Physical Disability	0.060	0.059	0.057	0.050	0.048	0.054
	(0.237)	(0.236)	(0.232)	(0.218)	(0.213)	(0.226)
Intellectual Disability	0.079	0.079	0.079	0.080	0.080	0.079
	(0.269)	(0.270)	(0.270)	(0.272)	(0.272)	(0.270)
Other Dichotomous Characteristics (omitted: alternative group)						
Female	0.489	0.489	0.489	0.489	0.489	0.489
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Economically Disadvantaged	0.503	0.504	0.502	0.487	0.465	0.487
	(0.500)	(0.500)	(0.500)	(0.500)	(0.499)	(0.500)
Limited English Proficiency	0.058	0.058	0.055	0.047	0.044	0.053
	(0.233)	(0.233)	(0.227)	(0.211)	(0.206)	(0.224)
Academically and Intellectually Gifted - Reading	0.123	0.124	0.126	0.132	0.133	0.133
	(0.329)	(0.329)	(0.332)	(0.338)	(0.339)	(0.340)
Academically and Intellectually Gifted - Math	0.134	0.135	0.137	0.142	0.142	0.145
	(0.341)	(0.341)	(0.344)	(0.349)	(0.349)	(0.352)
Old in the Grade	0.147	0.151	0.154	0.164	0.172	0.158
	(0.354)	(0.358)	(0.361)	(0.371)	(0.377)	(0.365)
Repeating Grade in the Academic Year	0.030	0.031	0.032	0.038	0.044	0.034
	(0.171)	(0.173)	(0.177)	(0.191)	(0.205)	(0.181)
Mean of Lagged Scores						
Lagged Normalized Math Score	0.006	0.004	0.006	0.007	0.012	0.008
	(0.896)	(0.901)	(0.905)	(0.921)	(0.910)	(0.918)
Lagged Normalized Reading Score	0.025	0.023	0.027	0.035	0.048	0.035
	(0.883)	(0.886)	(0.887)	(0.898)	(0.883)	(0.896)
Lagged Score Missing Indicator	0.190	0.185	0.169	0.136	0.154	0.146
	(0.392)	(0.388)	(0.375)	(0.343)	(0.361)	(0.354)
Grade level						
Grade 3	0.085	0.078	0.057	0.004	0.000	0.025
	(0.280)	(0.268)	(0.232)	(0.0621)	(0.0212)	(0.156)
Grade 4	0.087	0.079	0.059	0.005	0.001	0.027
	(0.281)	(0.271)	(0.236)	(0.0706)	(0.0254)	(0.161)
Grade 5	0.086	0.080	0.061	0.007	0.002	0.029
	(0.281)	(0.271)	(0.239)	(0.0861)	(0.0406)	(0.168)
Grade 6	0.108	0.111	0.118	0.124	0.081	0.134
	(0.310)	(0.315)	(0.323)	(0.330)	(0.272)	(0.340)
Grade 7	0.108	0.112	0.119	0.130	0.086	0.136
	(0.310)	(0.315)	(0.324)	(0.336)	(0.281)	(0.343)
Grade 8	0.108	0.111	0.119	0.130	0.087	0.136
	(0.310)	(0.315)	(0.323)	(0.336)	(0.282)	(0.343)
Grade 9	0.122	0.125	0.137	0.175	0.217	0.150
	(0.327)	(0.331)	(0.343)	(0.380)	(0.412)	(0.357)
Grade 10	0.107	0.110	0.120	0.154	0.191	0.132
	(0.309)	(0.313)	(0.325)	(0.361)	(0.393)	(0.339)
Grade 11	0.097	0.100	0.109	0.140	0.173	0.119
	(0.297)	(0.300)	(0.311)	(0.347)	(0.378)	(0.324)
Grade 12	0.091	0.094	0.102	0.131	0.162	0.112
	(0.288)	(0.291)	(0.303)	(0.338)	(0.369)	(0.315)
Sample Size	4545364	4372421	4003540	2984461	2321135	2870740

mean coefficients; sd in parentheses

Table A9: Other Estimation Results for Specification "All Offenses" in Table 1.5 (Part 1: $\hat{\beta}_{12}$, $\hat{\beta}_{13}$, $\hat{\beta}_{10}$)

	OLS	OLS &GFE	2SLS	2SLS &GFE
Race (ref. white)				
Black	0.092^{***}	0.107^{***}	0.092^{***}	0.107^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.013^{***}	-0.012^{***}	-0.013^{***}	-0.012^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Asian	-0.071^{***}	-0.070^{***}	-0.071^{***}	-0.070^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Multi-Racial	0.042^{***}	0.049^{***}	0.042^{***}	0.049^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
American Indian	0.048^{***}	0.057^{***}	0.048^{***}	0.057^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Other Race	0.100^{***}	0.091^{***}	0.102^{***}	0.094^{***}
	(0.009)	(0.008)	(0.009)	(0.008)
Disability (ref. no disability)				
Physical Disability	0.041^{***}	0.049^{***}	0.041^{***}	0.049^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Intellectual Disability	0.001	0.005^{***}	0.001	0.005^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Other Dichotomous Characteristics				
Female	-0.095^{***}	-0.108^{***}	-0.095^{***}	-0.108^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Economically Disadvantaged	0.081^{***}	0.083^{***}	0.081^{***}	0.084^{***}
	(0.001)	(0.000)	(0.001)	(0.000)
Limited English Proficiency	-0.014^{***}	-0.012^{***}	-0.014^{***}	-0.012^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Academically and Intellectually Gifted - Reading	-0.029^{***}	-0.031^{***}	-0.028^{***}	-0.031^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Academically and Intellectually Gifted - Math	-0.045^{***}	-0.049^{***}	-0.046^{***}	-0.049^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Old in the Grade	0.028^{***}	0.035^{***}	0.028^{***}	0.035^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Repeating Grade in the Academic Year	0.095^{***}	0.070^{***}	0.095^{***}	0.070^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Grade level (ref. grade 3)				
Grade 4	0.048^{***}	0.040^{***}	0.051^{***}	0.042^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Grade 5	0.066^{***}	0.057^{***}	0.069^{***}	0.061^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Grade 6	0.110^{***}	0.105^{***}	0.113^{***}	0.108^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Grade 7	0.112^{***}	0.116^{***}	0.114^{***}	0.120^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Grade 8	0.115^{***}	0.124^{***}	0.117^{***}	0.129^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Grade 9	0.019^{***}	0.031^{***}	0.015^{***}	0.030^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Grade 10	0.019^{***}	0.034^{***}	0.015^{***}	0.034^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Grade 11	0.019^{***}	0.029^{***}	0.014^{***}	0.028^{***}
	(0.004)	(0.004)	(0.004)	(0.004)
Grade 12	-0.007^{*}	0.000	-0.013^{***}	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)
Lagged Scores ($\hat{\beta}_{12}$)				
Lagged Normalized Math Score	-0.005^{***}	-0.004^{***}	-0.006^{***}	-0.004^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged Normalized Reading Score	0.002^{***}	0.000	0.003^{***}	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged Score Missing Indicator	0.026^{***}	0.019^{***}	0.026^{***}	0.019^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Constant ($\hat{\beta}_{10}$)	0.904^{***}	0.868^{***}	1.129^{***}	1.078^{***}
	(0.045)	(0.041)	(0.076)	(0.057)
Sample Size	4545364	4545364	4545364	4545364

Standard errors in parentheses

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

Table A10: Other Estimation Results for Specification "All Offenses" in Table 1.5 (Part 2: $\hat{\beta}_{11}$, $\hat{\beta}_{15}$)

	OLS	OLS &GFE	2SLS	2SLS &GFE
Offenses in Previous Year ($\hat{\beta}_{11}$)				
Lagged Frequency of Assault and Robbery	0.104^{***}	0.055^{***}	0.104^{***}	0.055^{***}
	(0.003)	(0.002)	(0.003)	(0.002)
Lagged Frequency of Threat and Weapon	0.108^{***}	0.030^{***}	0.108^{***}	0.031^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
Lagged Frequency of Harassment	0.070^{***}	0.042^{***}	0.070^{***}	0.042^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Lagged Frequency of Other Violence	0.112^{***}	0.076^{***}	0.111^{***}	0.075^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Frequency of Drug	0.111^{***}	0.073^{***}	0.111^{***}	0.073^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Lagged Frequency of Disrespect	0.054^{***}	0.035^{***}	0.054^{***}	0.035^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged Frequency of Disruption	0.051^{***}	0.031^{***}	0.051^{***}	0.031^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Lagged Frequency of Truancy	0.061^{***}	0.042^{***}	0.061^{***}	0.042^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Frequency of Tardiness	0.034^{***}	0.021^{***}	0.034^{***}	0.021^{***}
	(0.001)	(0.000)	(0.001)	(0.000)
Lagged Frequency of Property	0.098^{***}	0.033^{***}	0.098^{***}	0.033^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
Lagged Frequency of Minor Rule Violation	0.057^{***}	0.028^{***}	0.057^{***}	0.028^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Frequency of Other Offense	0.063^{***}	0.040^{***}	0.062^{***}	0.040^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Same Grade Peers' Observed Characteristics ($\hat{\beta}_{15}$)				
Ratio of Black Students	0.047^{***}	0.048^{***}	0.053^{***}	0.042^{***}
	(0.006)	(0.006)	(0.006)	(0.006)
Ratio of Other Minority Students	-0.030^{***}	-0.007	-0.038^{***}	-0.011
	(0.007)	(0.007)	(0.008)	(0.007)
Ratio of Female Students	-0.076^{***}	-0.078^{***}	-0.073^{***}	-0.077^{***}
	(0.005)	(0.004)	(0.005)	(0.004)
Ratio of Exceptional Students	-0.054^{***}	-0.040^{***}	-0.054^{***}	-0.035^{***}
	(0.006)	(0.006)	(0.006)	(0.006)
Ratio of AIG Students	0.031^{***}	0.037^{***}	0.020^{***}	0.030^{***}
	(0.003)	(0.003)	(0.004)	(0.003)
Ratio of Students with Limited English Proficiency	-0.004	-0.022^{***}	0.016^{*}	-0.013^{*}
	(0.008)	(0.007)	(0.009)	(0.007)
Ratio of Students Who Repeated Grades This Tear	0.185^{***}	0.162^{***}	0.219^{***}	0.206^{***}
	(0.007)	(0.007)	(0.012)	(0.011)
Ratio of students Who Are Above the Typical Age In the Grade	0.150^{***}	0.135^{***}	0.125^{***}	0.088^{***}
	(0.004)	(0.004)	(0.008)	(0.010)
Ratio of Economically Disadvantaged Students	0.042^{***}	0.072^{***}	0.014	0.068^{***}
	(0.004)	(0.004)	(0.009)	(0.004)
Mean of Last Year Math Standard Scores	0.003^{***}	0.014^{***}	0.004^{***}	0.015^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Last Year Reading Standard Scores	0.002^{***}	0.002^{***}	0.001	0.002^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Sample Size	4545364	4545364	4545364	4545364

Standard errors in parentheses

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

Table A11: Other Estimation Results for Specification “All Offenses” in Table 1.5 (Part 3: $\hat{\beta}_{14}$)

	OLS	OLS &GFE	2SLS	2SLS &GFE
Time Varying School Observables				
Violent crime cases last year (N)	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Students involved in misbehavior last year (N)	0.000***	0.000***	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Students assigned OSS or expulsion last year (N)	−0.000	−0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Assault, robbery or sexual offense cases (N	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Threat or possession of a weapon cases (N)	0.000*	0.000**	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Disorderly conduct or harassment cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Other violent cases (N)	−0.000*	0.000*	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Drug related cases (N)	0.000	0.000***	−0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Disrespect cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Disruptive behavior cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Truancy cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Tardiness cases (N)	−0.000***	0.000	−0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Property cases (N)	−0.000**	−0.000	−0.000*	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Other rule violation cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Other disciplinary infraction cases (N)	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Minor (average OSS days < 0.55) misbehavior cases (N)	0.000***	0.000***	0.000	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Moderate ($0.55 \leq$ average OSS days ≤ 1) misbehavior cases (N)	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Major (average OSS days > 1) misbehavior cases (N)	−0.000***	−0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Ratio of black students	0.087***	0.026***	0.090***	0.027***
	(0.008)	(0.007)	(0.008)	(0.007)
Ratio of Hispanic students	−0.027***	−0.017*	−0.079***	−0.048***
	(0.011)	(0.010)	(0.018)	(0.011)
Ratio of other minority students	0.096***	0.121***	0.098***	0.128***
	(0.016)	(0.014)	(0.016)	(0.014)
School mean of normalized math score last year	−0.008***	−0.004*	−0.004	−0.001
	(0.003)	(0.002)	(0.003)	(0.002)
School mean of normalized reading score last year	−0.026***	−0.016***	−0.031***	−0.016***
	(0.003)	(0.002)	(0.003)	(0.002)
Proportion of students – math scores 2 sd below state average last year	−0.048***	−0.034**	−0.050***	−0.052***
	(0.017)	(0.016)	(0.017)	(0.016)
Proportion of students – reading scores 2 sd below state average last year	−0.044***	0.024	−0.077***	0.038**
	(0.017)	(0.015)	(0.019)	(0.015)
Title I eligible school	−0.001	0.005***	−0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
School-wide title I	−0.004***	−0.000	−0.006***	−0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Magnet School Indicator	−0.030***	−0.030***	−0.034***	−0.033***
	(0.002)	(0.002)	(0.003)	(0.002)
PBIS - Green Ribbon School	−0.007***	−0.005***	−0.010***	−0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
PBIS - Model School	−0.006***	−0.004***	−0.009***	−0.007***
	(0.001)	(0.001)	(0.002)	(0.001)
PBIS - Exemplar School	−0.015***	−0.014***	−0.019***	−0.017***
	(0.002)	(0.002)	(0.002)	(0.002)
Ratio of teachers licensed in the school for more than 5 years	−0.017***	0.001	−0.032***	−0.007***
	(0.002)	(0.002)	(0.004)	(0.003)
Ratio of female personnels	0.011*	0.020***	0.003	0.012*
	(0.006)	(0.006)	(0.007)	(0.006)
Ratio of black personnels	−0.065***	−0.062***	−0.062***	−0.059***
	(0.007)	(0.006)	(0.007)	(0.006)
Ratio of non-white non-black personnels	0.091***	0.118***	0.090***	0.132***
	(0.013)	(0.012)	(0.013)	(0.012)
Total student number	−0.000***	−0.000***	−0.000***	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Students who are economically disadvantaged %	−0.029***	−0.010**	−0.025***	0.009*
	(0.004)	(0.004)	(0.005)	(0.006)
Total full-time equivalent classroom teachers	0.000***	−0.000	0.001***	−0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Fully licensed teachers %	0.028***	0.015**	0.019**	0.020***
	(0.007)	(0.007)	(0.008)	(0.007)
Teachers with experience 4-10 years %	0.025***	0.011**	0.029***	0.018***
	(0.005)	(0.004)	(0.005)	(0.004)
Teachers with experience more than 11 years %	0.019***	0.004	0.018***	0.004
	(0.005)	(0.005)	(0.005)	(0.005)
Teachers with Advanced Degrees %	−0.008	−0.011**	−0.008	−0.006
	(0.005)	(0.005)	(0.005)	(0.005)
Teacher Turnover Rate %	0.015***	0.016***	0.017***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)
Average daily school attendance %	−1.049***	−1.160***	−1.262***	−1.374***
	(0.035)	(0.034)	(0.068)	(0.052)
Students per Instructional Computer (N)	0.001***	0.000*	0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Books per Student (N)	−0.000	−0.000	−0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Average age of Books in library or media center	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Classes taught by highly qualified teachers	−0.109***	−0.105***	−0.101***	−0.099***
	(0.008)	(0.008)	(0.009)	(0.008)
Adequate yearly progress target met %	−0.009***	−0.001	−0.003	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Classrooms connected to the Internet %	0.035***	0.035***	0.040***	0.043***
	(0.004)	(0.004)	(0.005)	(0.004)
Total number of classroom teachers	0.000	0.000	−0.000	−0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
School Variable Missing Indicator	0.081***	0.064***	0.082***	0.064***
	(0.002)	(0.002)	(0.002)	(0.002)
Sample Size	4545364	4545364	4545364	4545364

Standard errors in parentheses

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

Table A12: Selected Estimation Results ($\hat{\beta}_{33}$, $\hat{\beta}_{32}$, $\hat{\beta}_{30}$) for Specification "OLS&GFE" for Equation 1.7

	End of Grade Math (Grade 3 -8)	Dropout (Grade 9)	ACT Composite Score (Grade 9 -10)
Race (ref. white)			
Black	-0.027^{***}	-0.087^{***}	-1.179^{***}
	(0.005)	(0.005)	(0.048)
Hispanic	0.001	-0.050^{***}	-0.476^{***}
	(0.003)	(0.006)	(0.054)
Asian	0.036^{***}	-0.077^{***}	0.846^{***}
	(0.011)	(0.013)	(0.154)
Multi-Racial	-0.007	-0.039^{***}	-0.334^{***}
	(0.005)	(0.007)	(0.070)
American Indian	-0.010	-0.014	-0.655^{***}
	(0.008)	(0.011)	(0.105)
Other Race	-0.047	-0.037	-0.832
	(0.040)	(0.056)	(0.575)
Disability (ref. no disability)			
Physical Disability	-0.258^{***}	-0.001	-0.334^{***}
	(0.005)	(0.005)	(0.053)
Intellectual Disability	-0.360^{***}	0.018^{***}	-0.286^{***}
	(0.006)	(0.005)	(0.046)
Other Dichotomous Characteristics			
Female	-0.007	-0.008^{***}	0.231^{***}
	(0.006)	(0.003)	(0.024)
Economically Disadvantaged	-0.026^{***}	0.069^{***}	-0.385^{***}
	(0.007)	(0.008)	(0.068)
Limited English Proficiency	0.019^{***}	0.029^{***}	-1.220^{***}
	(0.005)	(0.007)	(0.070)
Academically and Intellectually Gifted - Reading	0.000	-0.024^{***}	2.417^{***}
	(0.004)	(0.008)	(0.076)
Academically and Intellectually Gifted - Math	0.028^{***}	-0.036^{***}	2.580^{***}
	(0.004)	(0.008)	(0.078)
Old in the Grade	-0.056^{***}	0.129^{***}	-0.894^{***}
	(0.005)	(0.005)	(0.053)
Repeating Grade in the Academic Year	0.079^{***}		
	(0.008)		
Grade level (ref. grade 3 for End of Grade Math, grade 9 for ACT Composite Score)			
Grade 4	-0.021		
	(0.014)		
Grade 5	-0.017		
	(0.020)		
Grade 6	0.014		
	(0.028)		
Grade 7	0.017		
	(0.035)		
Grade 8	0.028		
	(0.038)		
Grade 10			0.225^{***}
			(0.038)
Lagged Scores ($\hat{\beta}_{32}$)			
Lagged Normalized Math Score	0.691^{***}	-0.001	0.504^{***}
	(0.005)	(0.003)	(0.024)
Lagged Normalized Reading Score	0.110^{***}	-0.002	1.440^{***}
	(0.007)	(0.004)	(0.032)
Lagged Score Missing Indicator	-0.149^{***}	-0.111^{***}	-0.570^{***}
	(0.008)	(0.006)	(0.037)
Constant ($\hat{\beta}_{30}$)	-0.586	1.260^{***}	19.819^{***}
	(0.703)	(0.342)	(3.116)
Sample Size	548893	114642	97419

Standard errors in parentheses

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

APPENDIX B

APPENDIX FOR CHAPTER 2

Table B1: Estimation Results: Black White Differences in OOS by Offense Type

Offense Type	Coeff.	SE	Sample
Fighting	0.012^{***}	(0.002)	167164
Inappropriate language/disrespect	-0.009^{***}	(0.003)	242416
Insubordination	0.003	(0.002)	448476
Disruptive behavior	0.005^{***}	(0.001)	753039
Disrespect of faculty/staff	-0.018^{***}	(0.003)	203359
Skipping Class	0.003	(0.002)	242792

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression results (with school-year fixed effects) for being a black student (relative to a white student), separately estimated for different types of offenses. Only schools that reported all the listed types of offenses in the academic year are used for the estimation. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

Table B2: Estimation Results: Black White Differences in OOS by Offense Type

Offense Type	Coeff.	SE	Sample
Fighting	0.011^{***}	(0.003)	96803
Dress code violation	0.012^{***}	(0.004)	61508
Inappropriate language/disrespect	-0.009^{**}	(0.003)	163262
Insubordination	0.002	(0.002)	326865
Theft	0.031^{***}	(0.009)	21874
Disruptive behavior	0.007^{***}	(0.002)	514009
Disrespect of faculty/staff	-0.020^{***}	(0.004)	136429
Leaving Class without permission	-0.007	(0.005)	43145
Skipping Class	0.005	(0.003)	186840

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports OLS regression (with school-year fixed effects) results for being a black student (relative to a white student), separately estimated for different types of offenses. Only schools that reported all the listed types of offenses in the academic year are used for the estimation. The dependent variable is the out-of-school suspension indicator. Standard errors are reported in parentheses and clustered at the school-year level.

APPENDIX C

APPENDIX FOR CHAPTER 3

C.1. Likelihood Function

Under the equilibia selection assumption desscribed in Section 3.2.5, we have the following conditional probability:

$$
\begin{aligned}
Pr(y_{it} = 1 \text{ and } y_{jt} = 1 | \Omega_{it}, \Omega_{jt}, \mu_t) &= Pr(\epsilon_{it} > -\beta_a \vec{\Omega}_{it} - \beta_{1a} \text{ and } \epsilon_{jt} > -\beta_b \vec{\Omega}_{jt} - \beta_{1b}) \\
&\quad - (1 - Pr(o = (1,1))) \\
&\quad Pr(-\beta_a \vec{\Omega}_{it} - \beta_{1a} < \epsilon_{it} \leq -\beta_a \vec{\Omega}_{it} + \beta_{1a} \text{ and } -\beta_b \vec{\Omega}_{jt} - \beta_{1b} < \epsilon_{jt} \leq -\beta_b \vec{\Omega}_{jt} + \beta_{1b}); \\
Pr(y_{it} = -1 \text{ and } y_{jt} = -1 | \Omega_{it}, \Omega_{jt}, \mu_t) &= Pr(\epsilon_{it} < -\beta_a \vec{\Omega}_{it} + \beta_{1a} \text{ and } \epsilon_{jt} < -\beta_b \vec{\Omega}_{jt} + \beta_{1b}) \\
&\quad - Pr(o = (1,1)) \\
&\quad Pr(-\beta_a \vec{\Omega}_{it} - \beta_{1a} < \epsilon_{it} \leq -\beta_a \vec{\Omega}_{it} + \beta_{1a} \text{ and } -\beta_b \vec{\Omega}_{jt} - \beta_{1b} < \epsilon_{jt} \leq -\beta_b \vec{\Omega}_{jt} + \beta_{1b}); \\
Pr(y_{it} = 1 \text{ and } y_{jt} = -1 | \Omega_{it}, \Omega_{jt}, \mu_t) &= Pr(\epsilon_{it} \geq -\beta_a \vec{\Omega}_{it} + \beta_{1a} \text{ and } \epsilon_{jt} \leq -\beta_b \vec{\Omega}_{jt} - \beta_{1b}); \\
Pr(y_{it} = -1 \text{ and } y_{jt} = 1 | \Omega_{it}, \Omega_{jt}, \mu_t) &= Pr(\epsilon_{it} \leq -\beta_a \vec{\Omega}_{it} - \beta_{1a} \text{ and } \epsilon_{jt} \geq -\beta_b \vec{\Omega}_{jt} + \beta_{1b});
\end{aligned}
$$

Suppose we treat μ_t ($t = 1, 2 \ldots, T$) as additional parameters to estimate without imposing a distributional assumption (i.e., year indicators). Combined with equations 3.7 and 3.8, we have the following conditional probability of observing smoking status and health states for individuals i and j (here, we have ignored the mortality probability for simplicity of explanation):

$$
\begin{aligned}
&Pr(y_{i0}, y_{i1}, \ldots, y_{iT}, y_{j0}, y_{j1}, \ldots, y_{jT}, H_{i1}, \ldots, H_{iT}, H_{j1}, \ldots, H_{jT} | X_i, X_j, \xi_i, \xi_j) = \\
&\qquad \prod_{t=0}^{T} Pr(y_{it}, y_{jt} | \Omega_{it}, \Omega_{jt}) \\
&\qquad \times \prod_{t=1}^{T} Pr(H_{it} | H_{it}^l, X_{it}^H, A_{it}, y_{jt-1}, \xi_i) Pr(H_{jt} | H_{jt}^l, X_{jt}^H, A_{jt}, y_{it-1}, \xi_j) \quad \text{(C.1)}
\end{aligned}
$$

where we define $\Omega_{it} = \{A_{it}, H_{it}, X_{it}^S, \xi_i\}$ $\Omega_{jt} = \{A_{jt}, H_{jt}, X_{jt}^S, \xi_j\}$ for t=1, …,T, $\Omega_{i0} = \{X_{i0}^S, \xi_i\}$, $\Omega_{j0} = \{X_{j0}^S, \xi_j\}$, and X_i, X_j as all of the exogenous covariates for all the periods, and assume $H_{i0} =$

$H_{j0} = 0$ (i.e., we ignore the initial health state probability for simplicity of explanation).

Therefore, the likelihood function unconditional on the permanent individual unobserved heterogeneity, ξ, is:

$$\begin{aligned}
&Pr(y_{i0}, y_{i1}, \ldots, y_{iT}, y_{j0}, y_{j1}, \ldots, y_{jT}, H_{i1}, \ldots, H_{iT}, H_{j1}, \ldots, H_{jT} | X_i, X_j) \\
&= \int_{supp(\xi_i, \xi_j)} Pr(y_{i0}, y_{i1}, \ldots, y_{iT}, y_{j0}, y_{j1}, \ldots, y_{jT}, H_{i1}, \ldots, H_{iT}, H_{j1}, \ldots, H_{jT} | X_i, X_j, \xi_i, \xi_j) \\
&\qquad f(\xi_i, \xi_j | X_i, X_j) d\xi_i \xi_j \qquad \text{(C.2)}
\end{aligned}$$

where we apply the distributional assumptions introduced in Section 3.2.5 for $\xi_i, \xi_j, \epsilon_i, \epsilon_j, \epsilon_i^h$, and ϵ_j^h.

Table C1: Estimation Results: Smoking and Health Probabilities for Friend Pairs

	Coeff	SE
Smoking at t (Equation 3.5)		
Smoke in $t-1$	1.046	0.032
Ever Regularly Smoked before $t-1$	0.430	0.042
Cardiovascular Disease Between t and $t-1$	-0.103	0.133
High Blood Pressure in $t-1$	-0.062	0.054
Female	0.208	0.089
Married	-0.157	0.063
(Age-19)/10 at t	0.204	0.084
(Age-19)^2/100 at t	-0.049	0.012
No High School	-0.094	0.112
Some High School	0.087	0.095
College	-0.054	0.058
Post-College	-0.036	0.089
Mode of Excluded Peers' Smoking Status	0.265	0.058
Excluded Peers are Missing	0.017	0.094
Social Contact:		
Smoke in t	0.082	0.030
Cardiovascular Disease Between t and $t-1$	-0.061	0.109
High Blood Pressure in $t-1$	0.050	0.052
Female	-0.093	0.087
Married	-0.037	0.064
No High School	-0.083	0.111
Some High School	-0.050	0.092
College	-0.098	0.057
Post-College	-0.161	0.091
ρ^S	-0.187	0.095
Constant	0.940	0.177
Cardiovascular Disease Event between $t-1$ and t (Equation 3.6)		
Smoke in $t-1$	0.482	0.078
Ever Regularly Smoked before $t-1$	-0.009	0.061
Ever Cardiovascular Disease before $t-1$	1.043	0.133
High Blood Pressure in $t-1$	0.631	0.130
High-Normal Blood Pressure in $t-1$	0.591	0.169
Female	-0.819	0.118
Married	0.143	0.140
(Age-19)/10 at t	2.172	0.391
(Age-19)^2/100 at t	-0.166	0.041
No High School	0.380	0.183
Some High School	0.356	0.164
College	-0.049	0.129
Post-College	-0.335	0.206
Social Contact: Smoke in $t-1$	0.041	0.064
ρ^H_a	-0.188	0.113
Constant	-9.259	0.938
Number of Pair-Exams:	4796	

Note: The table provides selected estimated parameters from the smoking and health event probabilities as part of a larger joint estimation of equations 3.5-3.9 for friend pairs. The specification normalizes one mass point to zero ($\xi^1_{(i,j)} = 0$) and identifies another mass point, $\xi^2_{(i,j)} = -2.225$ with standard error 0.198. The joint probabilities for the mass point combinations are $P(\xi_i = 0, \xi_j = 0) = 0.547$, $P(\xi_i = -2.36, \xi_j = 0) = 0.164$, $P(\xi_i = 0, \xi_j = -2.36) = 0.229$, and $P(\xi_i = -2.36, \xi_j = -2.36) = 0.06$. Time-varying unobserved effects (μ_t for t= $1, 2, 3, \ldots, 8$) are controlled for by including exam indicators. The estimated probability of equilibrium (1,1) is $Pr(o = (1, 1)) = 0.488$.

Table C2: Estimation Results: Smoking and Health Probabilities for Sibling Pairs

	Coeff	SE
Smoking at t (Equation 3.5)		
Smoke in $t-1$	0.708	0.017
Ever Regularly Smoked before $t-1$	0.314	0.026
Cardiovascular Disease Between t and $t-1$	-0.451	0.093
High Blood Pressure in $t-1$	-0.145	0.039
Female	0.055	0.039
Married	-0.230	0.037
(Age-19)/10 at t	0.012	0.045
(Age-19)^2/100 at t	-0.041	0.007
No High School	-0.071	0.109
Some High School	0.366	0.066
College	-0.144	0.046
Post-College	-0.600	0.067
Mode of Excluded Peers' Smoking Status	0.296	0.038
Excluded Peers are Missing	0.085	0.040
Social Contact:		
Smoke in t	0.022	0.016
Cardiovascular Disease Between t and $t-1$	0.132	0.080
High Blood Pressure in $t-1$	0.051	0.038
Female	-0.011	0.038
Married	-0.008	0.038
No High School	0.233	0.100
Some High School	0.039	0.070
College	-0.063	0.046
Post-College	0.044	0.062
ρ^S	0.282	0.029
Constant	1.967	0.096
Cardiovascular Disease Event between $t-1$ and t (Equation 3.6)		
Smoke in $t-1$	0.523	0.073
Ever Regularly Smoked before $t-1$	-0.091	0.039
Ever Cardiovascular Disease before $t-1$	1.379	0.079
High Blood Pressure in $t-1$	0.823	0.081
High-Normal Blood Pressure in $t-1$	0.578	0.103
Female	-0.791	0.072
Married	-0.108	0.076
(Age-19)/10 at t	1.659	0.200
(Age-19)^2/100 at t	-0.135	0.022
No High School	0.134	0.130
Some High School	-0.106	0.105
College	-0.191	0.078
Post-College	-0.761	0.143
Social Contact: Smoke in $t-1$	0.052	0.040
ρ^H_a	-0.138	0.105
Constant	-7.501	0.477
Number of Pair-Exams:	15346	

Note: The table provides selected estimated parameters from the smoking and health event probabilities as part of a larger joint estimation of equations 3.5-3.9 for sibling pairs. The specification normalizes one mass point to zero ($\xi^1_{(i,j)} = 0$) and identifies another mass point, $\xi^2_{(i,j)} = -1.843$ with standard error 0.048. The joint probabilities for the mass point combinations are $P(\xi_i = 0, \xi_j = 0) = 0.106$, $P(\xi_i = -1.847, \xi_j = 0) = 0.164$, $P(\xi_i = 0, \xi_j = -1.847) = 0.161$, and $P(\xi_i = -1.847, \xi_j = -1.847) = 0.569$. Time-varying unobserved effects (μ_t for $t = 1, 2, 3, \ldots, 8$) are controlled for by including exam indicators. The estimated probability of equilibrium (1,1) is $Pr(o = (1,1)) = 1$.

Table C3: Estimation Results: Smoking and Health Probabilities for Parent-Child Pairs

	Child Equation	SE	Parent Equation	SE
Smoking at t (Equation 3.5)				
Smoke in $t-1$	0.690	0.025	0.694	0.032
Ever Regularly Smoked before $t-1$	0.238	0.035	0.259	0.048
Cardiovascular Disease Between t and $t-1$	−0.194	0.467	−0.043	0.096
High Blood Pressure in $t-1$	−0.119	0.079	−0.101	0.063
Female	0.007	0.060	0.258	0.073
Married	−0.435	0.057	−0.252	0.081
(Age-19)/10 at t	0.093	0.085	−0.945	0.304
(Age-19)^2/100 at t	−0.073	0.020	0.041	0.031
No High School or Some High School	−0.392	0.416	−0.085	0.091
College	−0.526	0.415	−0.033	0.102
Post-College	−0.999	0.426	−0.537	0.213
Mode of Excluded Peers' Smoking Status	0.029	0.065	0.416	0.075
Excluded Peers are Missing	−0.044	0.056	0.141	0.068
Social Contact:				
Smoke in t	0.003	0.014	0.107	0.037
Cardiovascular Disease Between t and $t-1$	0.058	0.067	0.209	0.300
High Blood Pressure in $t-1$	0.089	0.050	−0.030	0.096
Female	0.053	0.064	0.022	0.068
Married	−0.024	0.063	0.023	0.064
No High School or Some High School	0.052	0.082	−0.875	0.343
College	−0.148	0.097	−0.933	0.345
Post-College	−0.096	0.164	−0.860	0.349
ρ^S	0.403	0.044	0.147	0.052
Constant	2.654	0.456	5.625	0.858
Cardiovascular Disease Event between $t-1$ and t (Equation 3.6)				
Smoke in $t-1$	0.452	0.320	0.441	0.083
Ever Regularly Smoked before $t-1$	−0.262	0.139	−0.088	0.041
Ever Cardiovascular Disease before $t-1$	1.995	0.326	0.719	0.079
High Blood Pressure in $t-1$	0.736	0.242	0.935	0.104
High Normal Blood Pressure in $t-1$	0.487	0.289	0.712	0.132
Female	−0.807	0.227	−0.403	0.080
Married	−0.494	0.240	−0.249	0.083
(Age-19)/10 at t	2.220	0.752	−0.773	0.506
(Age-19)^2/100 at t	−0.197	0.108	0.062	0.043
No High School or Some High School	0.117	0.795	0.072	0.088
College	−0.164	0.805	−0.199	0.104
Post-College	−0.558	0.867	0.243	0.172
Social Contact: Smoke in $t-1$	0.257	0.144	0.051	0.037
ρ_a^H	0.059	0.459	−0.368	0.120
Constant	−8.261	1.861	−0.343	1.490
Number of Pair-Exams:		10044		

Note: The table provides selected estimated parameters from the smoking and health event probabilities as part of a larger joint estimation of equations 3.5-3.9 for parent-child pairs. The specification normalizes one mass point to zero ($\xi^1_{(i,j)} = 0$) and identifies another mass point, $\xi^2_{(i,j)} = -1.999$ with standard error 0.066. The joint probabilities for the mass point combinations are $P(\xi_i = 0, \xi_j = 0) = 0.069$, $P(\xi_i = -2.008, \xi_j = 0) = 0.138$, $P(\xi_i = 0, \xi_j = -2.008) = 0.174$, and $P(\xi_i = -2.008, \xi_j = -2.008) = 0.619$. Time-varying unobserved effects (μ_t for $t = 1, 2, \ldots, 8$) are controlled for by including exam indicators. The estimated probability of equilibrium (1,1) is $Pr(o = (1,1)) = 0.563$.

Table C4: Additional Estimation Results: Initial Conditions and Attrition for Spouse Pairs

	Husband Equation		Wife Equation	
	Coeff	SE	Coeff	SE
Initial Smoking				
Constant	-1.894	0.278	-1.215	0.128
No High School	0.041	0.145	-0.019	0.183
Some High School	0.165	0.127	0.240	0.142
College	-0.424	0.081	-0.254	0.071
Post-College	-0.785	0.105	-0.176	0.125
Proportion of Smokers Before 19 in Same Birth-year Same Sex Cohort	3.538	0.515	3.530	0.297
Initial Health				
Constant	-11.786	1.967	-19.434	6.339
(Age-19)/10 at t	4.343	1.122	7.526	3.388
(Age-19)^2/100 at t	-0.531	0.160	-0.865	0.456
No High School	0.113	0.373	-0.240	0.689
Some High School	-0.431	0.447	0.366	0.614
College	0.315	0.294	0.033	0.489
Post College	0.036	0.484	0.714	1.109
High Blood Pressure in $t-1$	1.220	0.271	0.280	0.443
High-Normal Blood Pressure in $t-1$	0.275	0.512	0.528	0.795
Smoke in $t-1$	0.336	0.158	-0.020	0.285
Social Contact: Smoke in $t-1$	-0.084	0.155	0.151	0.206
ρ_a^{H0}	-0.093	0.220	-0.086	0.335
Attrition (Death)				
Constant	-8.867	0.913	-8.563	1.225
(Age-19)/10 at t	1.522	0.373	1.387	0.502
(Age-19)^2/100 at t	-0.037	0.038	-0.037	0.052
No High School	-0.032	0.175	-0.293	0.286
Some High School	0.021	0.170	0.400	0.239
College	0.011	0.123	-0.169	0.170
Post-College	-0.198	0.180	-0.114	0.359
High Blood Pressure in $t-1$	0.263	0.106	0.383	0.157
Cardiovascular Disease Between t and $t-1$	0.863	0.130	1.011	0.221
Ever Cardiovascular Disease before $t-1$	0.616	0.119	0.402	0.217
Ever Regularly Smoked before $t-1$	0.110	0.059	0.079	0.094
Smoke in t	0.359	0.081	0.006	0.162
Social Contact: Smoke in t	0.076	0.071	-0.064	0.114
ρ_a^D	0.098	0.096	0.459	0.201

Note: The table provides estimated parameters from the initial condition and attrition equations (3.7, 3.8, and 3.9) as part of a larger joint estimation of equations 3.5-3.9 for spouse pairs.

Table C5: Additional Estimation Results: Initial Conditions and Attrition for Friend Pairs

	Coeff	SE
Initial Smoking		
Constant	-1.599	0.138
Female	0.228	0.082
No High School	-0.298	0.104
Some High School	0.125	0.085
College	-0.123	0.053
Post-College	-0.524	0.086
Social Contact: Female	-0.097	0.069
Social Contact: No High School	0.075	0.097
Social Contact: Some High School	0.133	0.085
Social Contact: College	-0.024	0.053
Social Contact: Post-College	-0.044	0.083
Proportion of Smokers Before 19 in Same Birth-year Same Sex Cohort	2.761	0.228
Initial Health		
Constant	-13.430	3.241
Female	-0.719	0.289
Married	-0.120	0.350
(Age-19)/10 at t	4.472	1.644
(Age-19)^2/100 at t	-0.468	0.211
No High School	0.844	0.351
Some High School	0.391	0.412
College	0.257	0.331
Post College	0.309	0.585
High Blood Pressure in $t-1$	0.758	0.274
High-Normal Blood Pressure in $t-1$	0.565	0.670
Smoke in $t-1$	0.392	0.161
Social Contact: Smoke in $t-1$	-0.065	0.140
ρ_a^{H0}	-0.356	0.283
Attrition (Death)		
Constant	-8.189	0.959
Female	-0.643	0.124
Married	-0.221	0.137
(Age-19)/10 at t	1.369	0.388
(Age-19)^2/100 at t	-0.035	0.039
No High School	0.183	0.180
Some High School	0.445	0.174
College	0.152	0.138
Post-College	0.176	0.206
High Blood Pressure in $t-1$	0.232	0.117
Cardiovascular Disease Between t and $t-1$	0.878	0.150
Ever Cardiovascular Disease before $t-1$	0.750	0.149
Ever Regularly Smoked before $t-1$	0.113	0.060
Smoke in t	0.189	0.086
Social Contact: Smoke in t	0.076	0.075
ρ_a^{D}	0.274	0.140

Note: The table provides estimated parameters from the initial condition and attrition equations (3.7, 3.8, and 3.9) as part of a larger joint estimation of equations 3.5-3.9 for friend pairs.

Table C6: Additional Estimation Results: Initial Conditions and Attrition for Sibling Pairs

	Coeff	SE
Initial Smoking		
Constant	-1.133	0.118
Female	-0.015	0.047
No High School	0.044	0.093
Some High School	0.110	0.066
College	-0.282	0.043
Post-College	-0.556	0.060
Social Contact: Female	0.006	0.035
Social Contact: No High School	-0.025	0.091
Social Contact: Some High School	0.027	0.066
Social Contact: College	0.084	0.043
Social Contact: Post-College	-0.140	0.061
Proportion of Smokers Before 19 in Same Birth-year Same Sex Cohort	3.395	0.204
Initial Health		
Constant	-13.595	1.976
Female	-1.275	0.255
Married	0.167	0.296
(Age-19)/10 at t	4.615	1.080
(Age-19)^2/100 at t	-0.466	0.150
No High School	0.189	0.284
Some High School	-0.033	0.294
College	-0.365	0.318
Post College	-1.008	0.746
High Blood Pressure in $t-1$	0.787	0.254
High-Normal Blood Pressure in $t-1$	1.731	0.310
Smoke in $t-1$	0.182	0.114
Social Contact: Smoke in $t-1$	-0.107	0.117
ρ_a^{H0}	-0.134	0.212
Attrition (Death)		
Constant	-7.145	0.497
Female	-0.304	0.080
Married	-0.102	0.085
(Age-19)/10 at t	1.109	0.213
(Age-19)^2/100 at t	-0.018	0.023
No High School	-0.042	0.141
Some High School	0.497	0.111
College	0.108	0.093
Post-College	-0.185	0.159
High Blood Pressure in $t-1$	0.179	0.080
Cardiovascular Disease Between t and $t-1$	1.146	0.105
Ever Cardiovascular Disease before $t-1$	0.532	0.100
Ever Regularly Smoked before $t-1$	0.096	0.042
Smoke in t	-0.060	0.075
Social Contact: Smoke in t	0.055	0.050
ρ_a^{D}	0.506	0.093

Note: The table provides estimated parameters from the initial condition and attrition equations (3.7, 3.8, and 3.9) as part of a larger joint estimation of equations 3.5-3.9 for sibling pairs.

Table C7: Additional Estimation Results: Initial Conditions and Attrition for Parent-Child Pairs

	Child Equation		Parent Equation	
	Coeff	SE	Coeff	SE
Initial Smoking				
Constant	0.064	0.452	−1.467	0.469
Female	−0.040	0.068	0.219	0.140
No High School or Some High School	−0.884	0.394	−0.066	0.078
College	−1.250	0.394	0.048	0.089
Post-College	−1.574	0.398	−0.233	0.153
Social Contact: Female	0.016	0.057	0.075	0.060
Social Contact: No High School or Some High School	0.105	0.071	−0.324	0.420
Social Contact: College	0.171	0.082	−0.330	0.419
Social Contact: Post-College	0.235	0.149	−0.448	0.423
Proportion of Smokers Before 19 in Same Birth-year Same Sex Cohort	3.143	0.449	4.476	0.487
Initial Health				
Constant	−8.327	4.338	−4.665	5.097
Female	−1.275	1.060	−0.643	0.174
Married	0.15	1.252	−0.543	0.213
(Age-19)/10 at t	−1.506	3.636	0.042	2.424
(Age-19)^2/100 at t	1.254	0.838	0.079	0.287
No High School or Some High School	−0.245	1.758	0.347	0.192
College	0.013	1.948	0.079	0.225
Post-College	0.599	2.100	0.554	0.375
High Blood Pressure in $t-1$	0.473	1.270	0.956	0.174
Smoke in $t-1$	−0.134	0.531	0.322	0.087
ρ_a^{H0}	0.074	0.931	−0.486	0.178
Attrition (Death)				
Constant	−14.23	2.727	−9.602	1.296
Female	−0.287	0.271	−0.104	0.065
Married	−0.649	0.279	0.054	0.077
(Age-19)/10 at t	4.684	1.305	1.774	0.448
(Age-19)^2/100 at t	−0.493	0.187	−0.060	0.039
No High School or Some High School	0.461	1.075	−0.117	0.080
College	0.223	1.092	−0.107	0.091
Post-College	0.673	1.140	−0.082	0.167
High Blood Pressure in $t-1$	0.067	0.291	0.004	0.070
Cardiovascular Disease Between t and $t-1$	0.523	0.591	0.927	0.082
Ever Cardiovascular Disease before $t-1$	1.474	0.457	0.793	0.075
Ever Regularly Smoked before $t-1$	0.275	0.169	0.093	0.039
Smoke in t	0.512	0.405	0.071	0.089
Social Contact: Smoke in t	0.023	0.223	0.042	0.038
ρ_a^{D}	−0.156	0.525	0.248	0.100

Note: The table provides estimated parameters from the initial condition and attrition equations (3.7, 3.8, and 3.9) as part of a larger joint estimation of equations 3.5-3.9 for parent-child pairs.

Table C8: Marginal Effects: Alter's Smoking Behavior on Ego's Smoking Behavior Unconditional on Ego's Previous Smoking

	(1)	(2)	(3)	(4)	(5)
Wife's (Effect on) → Husband	0.076	0.044	0.042	0.040	0.042
	(0.007)	(0.009)	(0.012)	(0.012)	(0.013)
Husband's (Effect on) → Wife	0.065	0.044	0.045	0.047	0.048
	(0.007)	(0.008)	(0.012)	(0.012)	(0.012)
Friends	0.030	0.019	0.020	0.021	0.020
	(0.007)	(0.005)	(0.007)	(0.008)	(0.008)
Siblings	0.033	0.021	0.006	0.006	0.006
	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)
Parent's (Effect on) → Child	0.050	0.042	0.002	0.001	0.001
	(0.009)	(0.013)	(0.263)	(0.245)	(0.294)
Child's (Effect on) → Parent	0.025	0.008	0.016	0.021	0.021
	(0.005)	(0.009)	(0.010)	(0.008)	(0.008)
Estimator addresses:					
Simultaneity	No	Yes	Yes	Yes	Yes
Unobserved Heterogeneity	No	No	Yes	Yes	Yes
Endogeneity of the Health Variable	No	No	No	Yes	Yes
Attrition Due to Death	No	No	No	No	Yes
Percent Change					
Wife's (Effect on) → Husband	30.3	20.4	21.1	20.4	20.4
Husband's (Effect on) → Wife	28.6	20.3	25.3	25.9	25.5
Friends	13.8	9.0	10.4	10.6	10.2
Siblings	13.7	9.1	3.4	3.4	3.3
Parent's (Effect on) → Child	16.6	14.2	0.6	0.4	0.3
Child's (Effect on) → Parent	17.6	6.4	14.5	18.9	18.2

Note: The marginal effects of social contacts' smoking behavior on individual smoking behavior are simulated at the observed explanatory variables (i.e., unconditional on ego's previous smoking behavior). For friends and siblings, the effects are assumed to be symmetric. Percentage point changes are reported in the top panel. Standard errors are calculated using 500 random draws from the variance-covariance matrix of estimated parameters. Percent changes are provided in the bottom panel.

REFERENCES

Aizer, A. and J. J. Doyle (2015). Juvenile incarceration, human capital and future crime: evidence from randomly-assigned judges. *130*(2), 759–803.

Akay, A. (2012). Finite-sample comparison of alternative methods for estimating dynamic panel data models. *Journal of Applied Econometrics 27*(7), 1189–1204.

Anyon, Y., J. M. Jenson, I. Altschul, J. Farrar, J. McQueen, E. Greer, B. Downing, and J. Simmons (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review 44*, 379–386.

Apel, R. (2013). Sanctions, perceptions, and crime: Implications for criminal deterrence. *Journal of quantitative criminology 29*(1), 67–101.

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.

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies 58*, 277–297.

Badev, A. (2013). Discrete games in endogenous networks: Theory and policy.

Bajari, P., H. Hong, and S. P. Ryan (2010). Identification and estimation of a discrete game of complete information. *Econometrica 78*(5), 1529–1568.

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Becker, G. S. and K. Murphy (1988). A theory of rational addiction. *Journal of Political Economy 96*(4), 675–700.

Bjorn, P. A. and Q. H. Vuong (1984). Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation.

Bonhomme, S., T. Lamadon, and E. Manresa (2016a). Discretizing unobserved heterogeneity: Approximate clustering methods for dimension reduction.

Bonhomme, S., T. Lamadon, and E. Manresa (2016b). A distributional framework for matched employer employeee data.

Bonhomme, S. and E. Manresa (2015). Grouped patterns of heterogeneity in panel data. *Econometrica 83*(3), 1147–1184.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of econometrics 150*(1), 41–55.

Bramoullé, Y., R. Kranton, and M. D'amours (2014). Strategic interaction and networks. *The*

American Economic Review 104(3), 898–930.

Bresnahan, T. F. and P. C. Reiss (1991). Empirical models of discrete games. *Journal of Econometrics 48*(1), 57–81.

Brock, W. A. and S. N. Durlauf (2001). Discrete choice with social interactions. *The Review of Economic Studies 68*(2), 235–260.

Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer (1985). Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political economy 93*(3), 488–511.

Chalfin, A. and J. McCrary (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature 55*(1), 5–48.

Christakis, N. A. and J. H. Fowler (2008). The collective dynamics of smoking in a large social network. *New England journal of medicine 358*(21), 2249–2258.

Christakis, N. A. and J. H. Fowler (2013). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in medicine 32*(4), 556–577.

Ciliberto, F. and E. Tamer (2009). Market structure and multiple equilibria in airline markets. *Econometrica 77*(6), 1791–1828.

Clark, A. E. and F. Etilé (2006). Don't give up on me baby: Spousal correlation in smoking behaviour. *Journal of Health Economics 25*(5), 958–978.

Cutler, D. M. and E. L. Glaeser (2010). Social interactions and smoking. *Research Findings in the Economics of Aging*.

Darden, M. and D. Gilleskie (2016). The effects of parental health shocks on adult offspring smoking behavior and self-assessed health. *25*, 939–954.

Darden, M., D. Gilleskie, and K. Strumpf (2017). Smoking and mortality: New evidence from a long panel.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics 2*(2), 241–275.

Doyle, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *The American Economic Review 97*(5), 1583–1610.

Durlauf, S. N. and Y. M. Ioannides (2010). Social interactions. *Annual Review of Economics 2*(1), 451–478.

Durlauf, S. N. and D. S. Nagin (2011). Imprisonment and crime. *Criminology & Public Policy 10*(1), 13–54.

Ehrlich, I. (1981). On the usefulness of controlling individuals: an economic analysis of rehabilitation, incapacitation and deterrence. *The American Economic Review 71*(3), 307–322.

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics 21*, 768–769.

Gilleskie, D. and Y. S. Zhang (2010). Friendship formation and smoking initiation among teens.

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Heckman (1978). Dummy endogenous variables in a simultaneous equation system. *46*, 931–959.

Heckman, J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, 271–320.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics 24*(3), 411–482.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation1. *Econometrica 73*(3), 669–738.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Khwaja, A., F. Sloan, and S. Chung (2006). Learning about individual risk and the decision to smoke. (24), 683–699.

Kinsler, J. (2011). Understanding the black–white school discipline gap. *Economics of Education Review 30*(6), 1370–1383.

Kinsler, J. (2013). School discipline: a source or salve for the racial achievement gap? *International Economic Review 54*(1), 355–383.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *The American economic review 96*(3), 863–876.

Lee, T., D. Cornell, A. Gregory, and X. Fan (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children 34*(2), 167–192.

Levitt, S. D. (1998). Juvenile crime and punishment. *Journal of political Economy 106*(6), 1156–1185.

Levitt, S. D. and T. J. Miles (2006). Economic contributions to the understanding of crime. *Annu. Rev. Law Soc. Sci. 2*, 147–164.

Li, C. (2017). The effectiveness and consequences of exclusionary school discipline. *Working Paper*.

Lin, C.-C. and S. Ng (2012). Estimation of panel data models with parameter heterogeneity when group membership is unknown. *Journal of Econometric Methods 1*(1), 42–55.

Losen, D. J. and R. J. Skiba (2010). Suspended education: Urban middle schools in crisis.

Maddala (1983). *Limited-dependent and qualitative variables in econometrics*. New York: Cambridge University Press.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Marvell, T. B. and C. E. Moody (1994). Prison population growth and crime reduction. *Journal of Quantitative criminology 10*(2), 109–140.

Mendez, L. M. R. and H. M. Knoff (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*, 30–51.

Morris, E. W. and B. L. Perry (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems 63*(1), 68–86.

Mroz, T. A. (1999). Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *journal of Econometrics 92*(2), 233–274.

Mroz, T. A. and D. K. Guilkey (1992). Discrete factor approximation for use in simultaneous equation models with both continuous and discrete endogenous variables.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy 104*(5), 869–895.

Perry, B. L. and E. W. Morris (2014). Suspending progress collateral consequences of exclusionary punishment in public schools. *American Sociological Review 79*(6), 1067–1087.

Rabe-Hesketh, S. and A. Skrondal (2013). Avoiding biased versions of wooldridge's simple solution to the initial conditions problem. *Economics Letters 120*(2), 346–349.

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

Skiba, R. and M. Rauch (2015). Reconsidering exclusionary discipline: The efficacy and equity of out-of-school suspension and expulsion. *Handbook of Classroom Management*, 116–138.

Skiba, R. J., C.-G. Chung, M. Trachok, T. Baker, A. Sheya, and R. Hughes (2014). Where should we intervene? *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion 132*.

Skiba, R. J. and K. Knesting (2001). Zero tolerance, zero evidence: An analysis of school disciplinary practice. *New Directions for Youth Development 2001*(92), 17–43.

Skiba, R. J., R. S. Michael, A. C. Nardo, and R. L. Peterson (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The urban review 34*(4), 317–342.

Smith, E. J. and S. R. Harper (2015). Disproportionate impact of k-12 school suspension and expulsion on black students in southern states. *Centre for the Study of Race and Equity in Education, university of Pennsylvania (Philadelphia, PA: 2015). https://www. gse. upenn. edu/equity/sites/gse. upenn. edu. equity/files/publications/Smith_Harper_ Report. pdf*.

Soetevent, A. R. and P. Kooreman (2007). A discrete-choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics 22*(3), 599–624.

Steinley, D. (2006). K-means clustering: a half-century synthesis. *British Journal of Mathematical and Statistical Psychology 59*(1), 1–34.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80–108.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies 70*(1), 147–165.

Tauchen, H., A. D. Witte, and H. Griesinger (1993). Criminal deterrence: Revisiting the issue with a birth cohort. Technical report, National Bureau of Economic Research.

Vasan, R. S., M. G. Larson, E. P. Leip, J. C. Evans, C. J. O'Donnell, W. B. Kannel, and D. Levy (2001). Impact of high-normal blood pressure on the risk of cardiovascular disease. *New England Journal of Medicine 345*(18), 1291–1297.

Wald, J. and D. J. Losen (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development 2003*(99), 9–15.

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

Wettach, J., J. Owen, and C. K. Hoffman (2015). Instead of suspension: Alternative strategies for effective school discipline.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics 20*(1), 39–54.