

# The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia

By Matthew P. Steinberg and Johanna Lacoë

Foreword by David Griffith and Amber M. Northern





The Thomas B. Fordham Institute promotes educational excellence for every child in America via quality research, analysis, and commentary, as well as advocacy and exemplary charter school authorizing in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at [www.edexcellence.net](http://www.edexcellence.net).

The Institute is neither connected with nor sponsored by Fordham University.

# Contents

<b>Foreword</b> .....	<b>4</b>
<b>Introduction</b> .....	<b>9</b>
<b>Background</b> .....	<b>11</b>
Prior Research .....	<b>11</b>
Changes to the School District of Philadelphia’s Code of Conduct .....	<b>12</b>
<b>Data and Methods</b> .....	<b>15</b>
Data .....	<b>15</b>
Methods and Evidence Claims .....	<b>15</b>
<b>Findings</b> .....	<b>18</b>
Question 1: <i>Did Philadelphia’s discipline policy reform reduce the use of out-of-school suspensions?</i> .....	<b>18</b>
Question 2: <i>Was the policy reform associated with changes in suspensions, achievement, and school attendance for students who were suspended prior to the reform?</i> .....	<b>24</b>
Question 3: <i>Was the policy reform associated with changes in achievement and school attendance for peers who were not suspended prior to the reform?</i> .....	<b>26</b>
Question 4: <i>Was the policy reform associated with a change in racial disproportionality?</i> .....	<b>29</b>
<b>Takeaways</b> .....	<b>31</b>
<b>Endnotes</b> .....	<b>33</b>

# Foreword

*By David Griffith and Amber M. Northern*

Over the past thirty years, the school discipline pendulum—much like the criminal justice pendulum—has swung wildly from one extreme to the other, as policymakers have struggled to solve an inherently difficult problem. Today, the “zero tolerance” policies that were all the rage at the end of the last century are generally viewed as heavy-handed and blunt, removing administrator discretion and treating many different kinds of offenses as equally injurious. Yet as the tide of elite—and education reform—opinion has turned against over-suspension, the instinctive response of policymakers has once again been to tie the hands of teachers, principals, and local officials, this time with the explicit goal of reducing the use of suspensions, especially for traditionally disadvantaged groups.

This near-total reversal on school discipline policy was promoted by actions of the federal Office for Civil Rights under the Obama Administration, where critics of “exclusionary discipline” found a sympathetic ear. Meanwhile, in the past five years, at least twenty-two states and the District of Columbia have revised their laws to require or encourage schools to limit the use of suspensions or expulsions. A 2013 study by the American Association of School Administrators found that more than half of the 464 districts surveyed had revised their student codes of conduct to include changes in the use of non-punitive responses to misbehavior, out-of-school suspensions, and expulsions.

We support evidence-based efforts to create a positive school culture that reduces the need for suspensions and expulsions. But in too many places, the push for “alternatives to suspension” has led to empirically unproven strategies such as restorative justice, which seeks to “rehabilitate offenders through reconciliation with victims,” but which can also radically underestimate the severity of the challenges a school faces. And then there’s the problem of implementation: almost nine in ten teachers in California—where supporters of Senate Bill 607 are trying to permanently ban suspensions for defiant behavior through fifth grade and impose a five-year temporary ban in middle and high schools—say they need more training and support if they are to deploy alternative discipline techniques successfully.<sup>1</sup>



Complicating matters further, the question of whether or not suspensions are effective has been conflated with concerns about discrimination and racial bias. Yes, there are still many schools where large numbers of African American and Latino students are suspended or expelled, and we do not doubt that some of America’s 100,000-plus schools discriminate against minority children. Yet studies also show that a disproportionate number of these youngsters face challenges that put them at risk of antisocial behavior.<sup>2</sup> Tragically, they are much more likely to be poor, grow up in a single-parent family, have a parent in prison, and live in neighborhoods where poverty is concentrated. So it should not shock us to discover that, in some circumstances and communities, minority students misbehave at “disproportionate” rates. No responsible social scientist would attribute America’s stubborn racial and socioeconomic achievement gaps solely to educator bias. Yet the unspoken assumption behind then Secretary Arne Duncan’s decision to apply “disparate impact theory” to school discipline enforcement policy was that differences in suspension rates must be attributable to such bias. According to a 2015 EdNext poll, half of the public and 59 percent of teachers opposed this approach.<sup>3</sup>

**Complicating matters further, the question of whether or not suspensions are effective has been conflated with concerns about discrimination and racial bias.**

Overall, we agree that suspensions are unlikely to benefit suspended students. But the problem with the current debate about discipline policy is that those aren’t the only students whose futures are at stake. For example, a 2008 study found that children from troubled families “significantly decrease their peers’ reading and math test scores and significantly increase misbehavior of others in the classroom.”<sup>4</sup> And a 2009 study found that when disruptive students from New Orleans relocated to Houston schools after Hurricane Katrina they “increased native absenteeism and disciplinary problems.”<sup>5</sup>

**Overall, we agree that suspensions are unlikely to benefit suspended students.**

Yes, one important question about school discipline is whether it helps or harms those being disciplined. But a second, equally important, question is whether the push to reduce the number of suspensions is harmful to the rule-abiding majority. According to a 2004 study, 85 percent of teachers and 73 percent of parents felt the “school experience of most students suffers at the expense of a few chronic offenders.”<sup>6</sup> And that was *before* the push to reduce suspensions. A more recent study by the Manhattan Institute’s Max Eden showed that the percentages of students and teachers in New York City reporting drug use, gang activity, and physical fights rose dramatically in the years following discipline reforms initiated by Mayor

Bill de Blasio, which require that principals obtain written approval from the city before suspending a student for “uncooperative/noncompliant” or “disorderly” behavior.

To examine these matters further, we enlisted Matthew Steinberg, assistant professor of education at the University of Pennsylvania, and Johanna Lacoë, an experienced researcher at Mathematica Policy Research, both of whom have previous experience studying school discipline. For this study, they examined outcomes in the School District of Philadelphia (SDP), which was gracious enough to share its data despite the sensitive nature of the subject. The SDP made dramatic changes to its code of conduct in the 2012–13 school year. Most notably, it instituted a new district-wide ban on out-of-school suspensions (OSS) for low-level “conduct” offenses—such as profanity or failure to follow classroom rules—and reduced the length of OSS for more serious infractions. To gauge the impacts of these changes, the authors examined data before and after they were implemented, and together, penned two scholarly papers: one that focuses on the district-level effects of the change in discipline policy and a second that utilizes student-level data to explore patterns of attendance and achievement at the school, grade, and individual levels in the wake of the policy change. For the benefit of our readers, this report combines those papers and synthesizes their key findings.

Before turning to that synthesis, we’re obliged to point out that school discipline is an extraordinarily difficult subject to study, in part because we do not observe student behavior directly, but only documented responses to it. So it is difficult to distinguish between the effects of a suspension and the effects of the behavior or conditions that triggered it. Drs. Steinberg and Lacoë deal with these challenges transparently by explaining the study’s limitations and the level of evidentiary rigor that readers can expect given their best efforts to improve upon prior work in this field. Below is a summary of what they found:

- Changes in district policy had no long-term impact on the number of low-level “conduct” suspensions, and most schools did not comply with the ban on such suspensions.
- Changes in district policy were associated with improved attendance—but not improved achievement—for previously suspended students.
- “Never-suspended” peers (i.e., students who didn’t receive a suspension in any of the years considered by the study) experienced worse outcomes in the most economically and academically disadvantaged schools, which were also the schools that did not (or could not) comply with the ban on conduct suspensions.
- Revising the district’s code of conduct was associated with an increase in racial disproportionality at the district level.

Based on these findings, we draw three conclusions:

**1 Schools may respond very differently to district mandates, depending on their demographics, achievement levels, and prior suspension rates, as well as other factors bearing on policy implementation and compliance.**

In Philadelphia, the “toughest” schools tended to ignore the district’s ban on conduct suspensions, while the highest-achieving schools appeared unaffected—because they didn’t have any conduct suspensions to begin with—and the schools in the middle stayed lukewarm (meaning they reduced OSS but didn’t eliminate them). In other words, schools responded (or didn’t) much as you might expect them to, given their pre-existing challenges.

**2 Top-down mandates can have unintended consequences—even when they emanate from local decision makers rather than distant state or federal governments.**

In Philadelphia, never-suspended students in many schools, including most schools that reduced their suspension rates, experienced a decline in academic performance, relative to the most plausible comparison group. Furthermore, the district-wide decline in conduct suspensions coincided with a suspicious *increase* in the number of minority students suspended for more serious infractions (though it is impossible to know whether some less serious offenses were reclassified as a result of the policy change). Clearly, these are not the types of responses that district leaders intended.

**3 Policymakers should respect the wisdom of practitioners when it comes to school discipline.**

For us, the biggest lesson from Philadelphia’s experience is that “discipline reform”—however defined or conceptualized—is best initiated at the school level rather than the district level, where the law of unintended consequences is more apt to prevail. Suspensions may have costs for suspended students, but these must be balanced against the necessity of maintaining an orderly learning environment. And the individuals best positioned to make those judgment calls, and to gauge how effective future approaches to discipline may be, are those on the front lines. Teachers and administrators who are struggling to manage disorder cannot be expected to comply willingly or well with a directive that eliminates one of their most important tools.

Overall, the report’s findings speak to the stubborn realities that educators must contend with, which brings us to our final point: a plea for the removal of the rose-colored glasses that so many observers and critics seem to don when viewing school discipline. Everyone knows that changing a district’s policy on suspensions is unlikely to alter the underlying issues in tough

schools—or in peaceful ones. So viewing this as a civil rights issue and trying to fix it with top-down decrees is impractical and potentially harmful, whether those decrees emanate from the district, the state, or the banks of the Potomac. If the goal is finding more effective ways to build a safe and strong school culture, it is far better to work with staff in high-poverty schools than to imply that they have racist tendencies and may be deliberately violating students' civil rights.

We harbor no illusions that this study will put an end to the discipline debate. But we hope it will inject a measure of nuance into the conversation—and perhaps help the discipline policy pendulum find a more stable resting place somewhere in the vast and sensible middle ground.

## Acknowledgments

This report was made possible through the support of our sister organization, the Thomas B. Fordham Foundation. We are particularly grateful to the authors, Matthew Steinberg and Johanna Lacoë, for their willingness to tackle such a difficult subject and their patience as we worked to distill their comprehensive results. We also thank external reviewers Scott Carrell, professor of economics at the University of California, Davis, and Dan Goldhaber, director of the Center on Education Data and Research at the University of Washington, for their insightful comments on draft papers.

On the Fordham side, we thank President Michael Petrilli and President Emeritus Chester E. Finn, Jr. for reviewing drafts and providing valuable feedback. We also thank Alyssa Schwenk for managing report dissemination; Caryn Morgan for handling funder relations; and Jonathan Lutton for overseeing report production and layout. Kudos also to Shannon Last for copyediting the report and Edward Alton of Alton Creative, Inc. for designing the report's layout. Shutterstock provided the image for the report's cover.

The report authors wish to thank the School District of Philadelphia and the Pennsylvania Department of Education for providing access to the data used in this report. They also thank Peter Schochet, John MacDonald, and Matthew Chingos for feedback on earlier versions of the working papers upon which these findings are based; Cameron Anglum for excellent research assistance; conference participants at the Association for Public Policy Analysis and Management and Association for Education Finance and Policy; and seminar participants at the University of Arkansas, Department of Education Reform Lecture Series for helpful comments and discussions. The authors contributed equally to the writing and analyses contained in this report. The complete working papers, on which this report is based, are available at <https://edexcellence.net/publications/discipline-in-philadelphia>.



# Introduction

By May 2015, at least twenty-two states and the District of Columbia had revised their discipline policies. These reforms have required or encouraged schools to limit the use of exclusionary discipline practices (such as out-of-school suspension and expulsion), to implement supportive (i.e., non-punitive) discipline strategies that rely on behavioral interventions, and to provide support services such as counseling, dropout prevention, and guidance services for at-risk students. A recent study by the American Association of School Administrators found that more than half of the 464 districts surveyed had revised their student code of conduct to include changes in the use of non-punitive responses to misbehavior, out-of-school suspensions and expulsions, and the length of suspensions.<sup>7</sup> By the 2015–16 school year, twenty-three of the nation’s one hundred largest school districts had implemented policy reforms that limited the use of suspensions or required less punitive discipline strategies.<sup>8</sup>

Among these was the School District of Philadelphia (SDP), which made dramatic changes to its code of conduct in 2012–13. Most notably, the SDP prohibited the use of out-of-school suspensions (OSS) for low-level conduct offenses—such as profanity and failure to follow classroom rules—and reduced the length of OSS for more serious infractions.

Although numerous studies have examined the relationship between OSS and student outcomes, little is known about the effects of district-level policy reforms—and in particular, their implications for non-offending peers in schools where the use of OSS is most prevalent. Consequently, we address four questions:

## **Question 1**

*Did Philadelphia’s discipline policy reform reduce the use of out-of-school suspensions?*

## **Question 2**

*Was the policy reform associated with changes in suspensions, achievement, and school attendance for students who were suspended prior to the reform?*

### Question 3

*Was the policy reform associated with changes in achievement and school attendance for peers who were not suspended prior to the reform?*

### Question 4

*Was the policy reform associated with a change in racial disproportionality?*

To answer Question 1, we analyzed a decade of district-level data from every public school district in Pennsylvania, including seven years of data from before the policy change (to gauge pre-policy trends) and three years of data from after the policy change (to observe these same trends post-policy). To answer Questions 2–4, we analyzed three years of student-level data provided by SDP, which include the years immediately before, during, and after district-wide implementation of the policy change. These data sources allowed us to examine the outcomes of Philadelphia’s revised discipline policy at multiple levels: district, school, and student.

The SDP is the fifth-largest district in the United States, serving approximately 200,000 students across 220 traditional public schools and eighty-eight charter schools.<sup>9</sup> Given the breadth of the movement to reform school discipline—similar reforms have been enacted in other large urban districts, including Chicago and New York City—we believe our results have important lessons for other school districts considering changes to their discipline policies.

This report is a summary of two more technical scholarly articles that can be found in “Report Materials” here: <https://edexcellence.net/publications/discipline-in-philadelphia>. The report is organized as follows: Section II provides a brief review of the literature on school discipline as well as background on the SDP and the changes to its Code of Conduct. Section III describes the data and methods used, Section IV summarizes the findings of the analysis, and Section V presents the key takeaways.

# Background

## Prior Research

A vast body of descriptive research has shown that students in traditionally disadvantaged subgroups are more likely to be suspended than other students. For example, African American students are more than four times as likely to be suspended as white students.<sup>10</sup> Research suggests that underlying social factors such as poverty and exposure to neighborhood violence,<sup>11</sup> or principals' perspectives on discipline (e.g., a more preventative approach to discipline versus a more exclusionary approach), may contribute to differences in suspensions by student race and ethnicity. But they do not fully explain these differences.<sup>12</sup>

Similarly, a vast body of correlational research has demonstrated that suspended students have lower grades and test scores,<sup>13</sup> are less likely to be promoted to the next grade level<sup>14</sup> and to graduate from high school,<sup>15</sup> and are far more likely to wind up in the criminal justice system.<sup>16</sup> Yet there is essentially no causal evidence on the effect that out-of-school-suspensions (OSS) have on the achievement and attendance of suspended students, due to the methodological challenges that are also present in the (far smaller) literature dealing with the effects of suspensions on peers (see *Methodological Challenges*).

Finally, little is known about the efficacy of changes in discipline policy that are initiated at the district level, the degree to which policy changes are adopted by schools, or how such changes might affect school climate, student behavior, and student achievement. One study of Chicago Public Schools found that, among ninth grade students, reducing the length of OSS for more serious misconduct had little impact on peer achievement or school climate.<sup>17</sup> However, Chicago's effort represents a very different approach to discipline reform than the policy enacted in Philadelphia, which prohibits suspensions for specific, lower-level infractions. And a recent descriptive analysis of a school climate survey given to students throughout the New York City public schools found that the proportion of students and teachers reporting increases in drug use, gang activity, and "physical fights" increased dramatically in the years following discipline reforms initiated by Mayor Bill de Blasio, which required principals to obtain written approval from the city before suspending students for "uncooperative" or "disorderly" behavior.<sup>18</sup>

## METHODOLOGICAL CHALLENGES

Nearly all school discipline studies must contend with at least two methodological challenges: First, there is usually a “selection problem,” meaning that the decision to suspend students is not random. Second, there is almost always “omitted variable bias,” meaning it is impossible to account for all the factors that make a student both more likely to be suspended *and* less likely to succeed academically. In particular, because we do not observe student behavior, it is difficult to distinguish between the effects of suspensions and the effects of behavior or conditions that triggered an infraction (such as unobserved changes in a student’s home life).

As a result of these challenges, it is likely that the costs associated with suspensions are biased, though it is impossible to empirically assess whether these costs are understated or overstated without the benefit that an experimental study would offer. (It is of course impractical and unethical to randomly assign suspensions to students in order to isolate the direct impact of suspensions on student outcomes.) Similarly, studies that seek to estimate the impact of suspensions on peers do not fully account for changes in behavior at the school or classroom level that may also be related to changes in student outcomes. As a result, estimates of the effect on peers are plagued by biases similar to those present in estimates of suspensions on suspended students.

## Changes to the School District of Philadelphia’s Code of Conduct

Compared to other school districts in Pennsylvania (and the United States), the School District of Philadelphia serves an unusually disadvantaged student population. For example, approximately three-quarters of SDP pupils are minority (black or Hispanic), and approximately 80 percent are eligible for free or reduced-price lunch. In the 2011–12 school year—the year prior to the district’s policy change—approximately 15 percent of Philadelphia students in grades 3–12 received at least one out-of-school suspension. (See Table 1, page 14, for infractions subject to suspension in 2011-12.) This is roughly 2.5 times the national average of 6.4 percent. Moreover, a disproportionate share of students who were suspended in the 2011–12 year in Philadelphia were black (73 percent, compared to 57 percent district-wide) or economically disadvantaged (73 percent, compared to 65 percent district-wide).

In August 2012, the SDP held a summit of school principals to discuss preventive strategies to improve school safety. Meanwhile, a private Philadelphia foundation funded a two-year fellowship to develop a “school safety and climate strategy” for the district, based on

collaboration among the SDP; the Philadelphia Departments of Behavioral Health, Police, and Human Services; and student and parent representatives.<sup>19</sup> These reform efforts culminated in changes to the SDP’s Code of Student Conduct.

Beginning in September 2012, SDP adopted a revised code of conduct that emphasized intervention rather than suspension or disciplinary transfers (i.e., transferring students across schools for disciplinary reasons) and prevented students from receiving OSS for less severe conduct infractions. Under the revised code of conduct, students were no longer to be removed from school for two specific infractions: first, failing to follow classroom rules, and second, using profane or obscene language or gestures. Instead, the maximum allowable punishment for students changed from one to three days of OSS to in-school intervention. Similarly, for other infractions—such as public displays of affection, inappropriate use of electronic devices, and forgery of an adult’s signature—the policy change required in-school intervention as a first response, with OSS to be used only as a last resort. Finally, for more serious offenses—such as theft, harassment and bullying, consensual sex, breaking and entering, robbery, extortion, or simple assault—maximum punishments were changed from expulsion to suspension. These punishments potentially could be paired with assignment to a disciplinary school or another type of transfer. Note that although we have detailed information on OSS, we lack data on “in-school interventions,” which include everything from in-school suspensions to any number of activities instituted to replace OSS (such as restorative justice practices).



● **TABLE 1:** Prior to discipline reform, Philadelphia’s code of student conduct deemed the following infractions subject for suspension.

Infraction Type	
<b>Conduct</b>	<ul style="list-style-type: none"> <li>• Failure to follow classroom rules/disruption*</li> <li>• Profane/obscene language or gestures*</li> <li>• Alteration of grade reporting/excuses/school documents</li> <li>• Forgery of administrator, teacher, or parent’s/guardian’s signature</li> <li>• Inappropriate use of electronic devices</li> <li>• Public display of affection/inappropriate touching</li> </ul>
<b>Non-Conduct</b>	<ul style="list-style-type: none"> <li>• Aggravated assault (documented serious bodily injury)</li> <li>• Assault of school personnel</li> <li>• Breaking and entering school property</li> <li>• Destruction and/or theft of property (less than \$1,000)</li> <li>• Destruction and/or theft of property (totaling \$1,000 or more)</li> <li>• Extortion; fighting (two students engaged in mutual combat)</li> <li>• Harassment/bullying/cyber-bullying/intimidation</li> <li>• Instigation or participation in group assaults</li> <li>• Mutual fight (with documented serious bodily injury)</li> <li>• Possession of a weapon</li> <li>• Possession of alcohol or drugs with intent to distribute</li> <li>• Possession of alcohol or drugs with intent to use</li> <li>• Possession or use of fireworks/incendiary devices/explosives</li> <li>• Robbery</li> <li>• Sexual acts (consensual); sexual acts (non-consensual)</li> <li>• Simple assault (documented unprovoked attack by one student on another)</li> <li>• Threatening students/staff with aggravated assault</li> </ul>

**Note:** The district-level suspension data from the Pennsylvania Department of Education are available for the broad categories of “conduct” and “non-conduct.” The student-level suspension data from the School District of Philadelphia are available by infraction type. To show how the data from the two sources correspond, the table displays the specific infraction types by the broader “conduct” and “non-conduct” categories. The two infractions noted by an asterisk were subject to the policy reform in SPD in the 2012–13 school year.

# Data and Methods

## Data

We utilize a variety of data sources to address the research questions. To address Question 1, we utilize district-level data from the 2005–06 through 2014–15 school years. These data come from two sources: the Common Core of Data collected by the National Center for Education Statistics, which includes demographic data on every district in Pennsylvania; and the Pennsylvania Department of Education, which collects information on enrollment, achievement, attendance, suspensions, and serious incidents. With these data, we are able to compare outcomes in Philadelphia to those in other districts across the state that did not implement discipline policy reform.

To address Questions 2–4, we utilize student-level data from three academic years (2011–12 through 2013–14) provided by the School District of Philadelphia. These data contain detailed information on student demographics, enrollment, attendance, achievement, and discipline records. Discipline data are reported at the infraction level, allowing us to observe each behavioral infraction a student commits that corresponds to an out-of-school suspension (OSS), the length of the associated suspension, and the specific type of student misconduct (i.e., the behavioral reason for the suspension). This detailed, infraction-level information allows us to distinguish between suspensions for low-level “conduct” infractions, which were the target of the district’s policy change, and suspensions for other, more serious “non-conduct” infractions (see Table 1 on the previous page for a list of infractions corresponding to either a *conduct* or *non-conduct* suspension).

## Methods and Evidence Claims

Many of the limitations that have plagued previous school discipline studies also apply to this one (see *Methodological Challenges*, in Section II). In particular, because we do not observe student behavior (which may or may not result in an OSS), it is difficult for us to distinguish between the effects of suspensions and the effects of the behaviors or conditions that triggered them (such as changes in a student’s home life that lead to an increase in disruptive behavior).

Consequently, it is possible that our estimates overstate the impacts of suspensions. However, this study improves upon prior studies in three critical ways: first, we provide the first empirical estimate of the effect of a discipline policy reform (designed to reduce suspensions for low-level infractions) on the prevalence of suspensions; second, we employ a more rigorous strategy (i.e., difference-in-differences) to estimate the relationship between the policy change and outcomes for suspended students and their non-offending peers; and third, we explore variation in the school-level implementation of the district’s policy change.

The research questions are addressed via a variety of quasi-experimental research designs, which lend themselves to varying degrees of evidentiary rigor. Below is a short description of each question, design, and evidence level:

### **Question 1**

*Did Philadelphia’s discipline policy reform reduce the use of out-of-school suspensions?*

Estimates of the effect of the district’s policy reform on the prevalence of suspensions were generated using a difference-in-differences approach, which compares changes in outcomes for Philadelphia before and after the policy change to changes in outcomes in all other Pennsylvania districts during the same time period. This approach controls for observable characteristics of districts—including enrollment, student poverty levels, and pre-reform achievement—as well as unobserved differences about districts that do not vary over time. This approach also accounts for the fact that OSS in Philadelphia—both overall and for conduct infractions—was declining in the pre-policy period (i.e., 2005–06 through 2011–12), while average OSS rates in all other Pennsylvania districts remained relatively stable. We categorize evidence generated using this approach as “strong.”

### **Question 2**

*Was the policy reform associated with changes in suspensions, achievement, and school attendance for students who were suspended prior to the reform?*

Estimates of the effect of the policy change on previously suspended students were generated using a difference-in-differences approach, which compares changes in the achievement and attendance of students who received suspensions for conduct infractions prior to the policy change to students who did not. This approach controls for observable student characteristics such as race and socioeconomic status, as well as unobservable characteristics that usually do not change over time, such as parental education. However, because it cannot account for unobserved time-varying student characteristics that may also affect outcomes—such as a death in the family or exposure to neighborhood violence—these factors may also contribute

to the estimated changes in student outcomes. We categorize estimates generated by this approach as “associational.”

### **Question 3**

*Was the policy reform associated with changes in achievement and school attendance for peers who were not suspended prior to the reform?*

Like estimates of the effect of the policy change on previously suspended students, estimates of changes in outcomes for peers—students who were not suspended in the years immediately before or after the policy change—were generated using a difference-in-differences approach. In this case, however, the comparison group consists of peers in schools that did not suspend students for conduct-related infractions in either the pre- or post-policy years (2011–12 and 2012–13, respectively). These schools did not change suspension levels in response to the policy; therefore we label them as “comparison schools.” Because schools exhibited varying levels of compliance with the policy reform, we categorize the remaining schools as full compliers, partial compliers, or non-compliers, and compare the change in outcomes for non-suspended peers in each of these groups to non-suspended peers in the comparison schools. Given the differences in the student characteristics of schools in the different implementation groups and the non-random selection into compliance status, we categorize estimates generated by this approach as “associational.”

### **Question 4**

*Was the policy reform associated with a change in racial disproportionality?*

Access to student-level data allowed us to characterize the level of racial disproportionality both before and after the policy change. To do so, we again employed a difference-in-differences approach to estimate whether changes in the use of suspensions, by racial and ethnic group, varied in the post-reform period. The models control for time-invariant characteristics of schools and observable characteristics of students. Because this approach is unable to control for differences in student behavior that do not result in OSS, we categorize the evidence as “associational.”

For a full explanation of the methods, see the working papers found in “Report Materials” here: <https://edexcellence.net/publications/discipline-in-philadelphia>.

# Findings

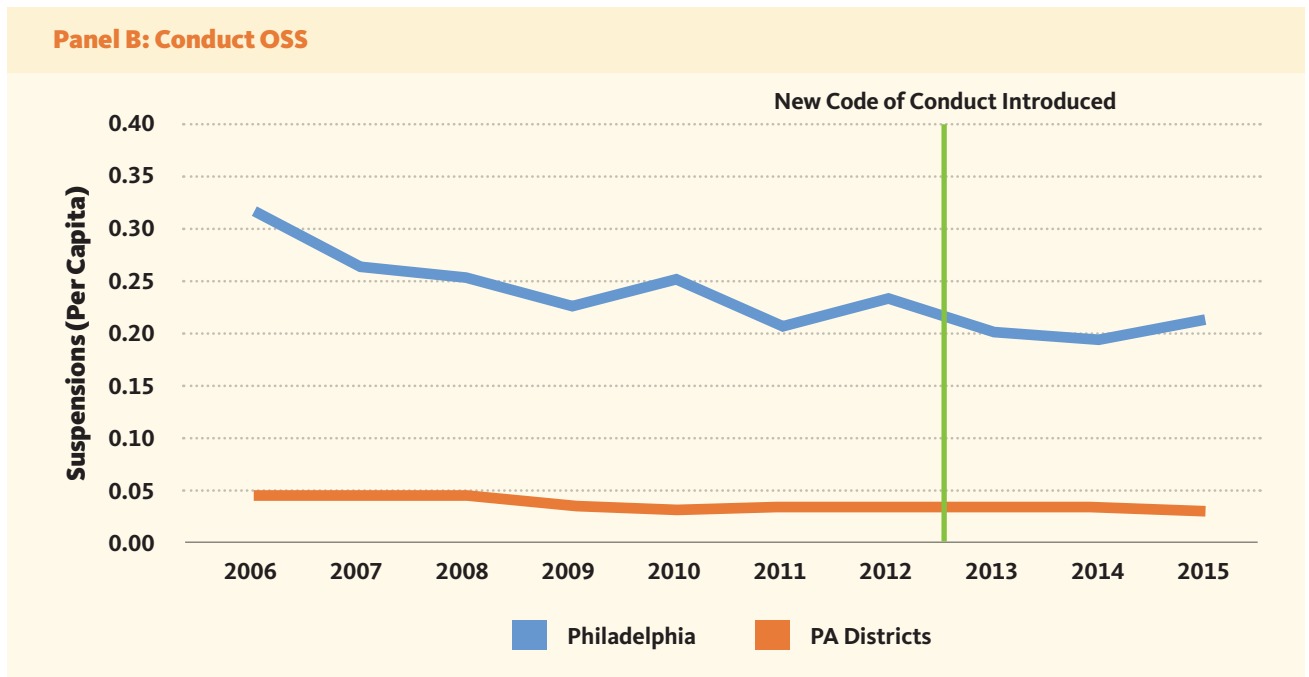
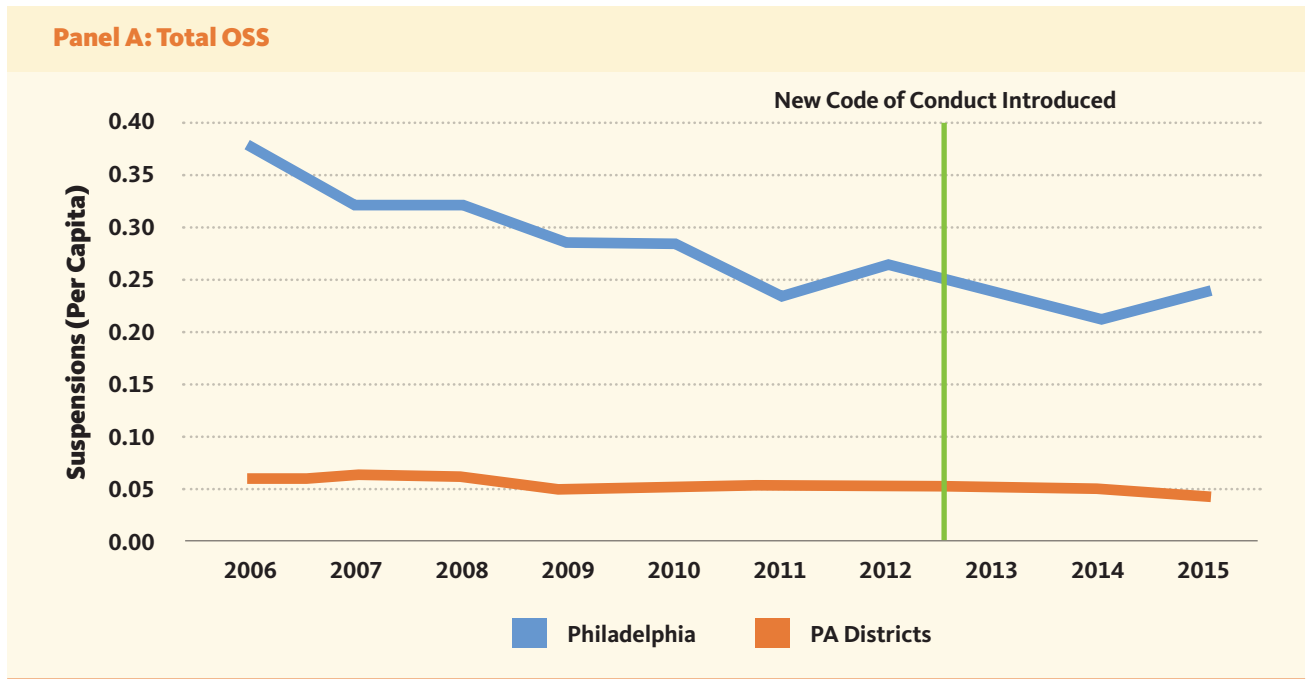
## Question 1: Did Philadelphia's discipline policy reform reduce the use of out-of-school suspensions?

**Changes in district policy resulted in an initial reduction in the number of low-level conduct suspensions, but the decrease did not persist. Notably, most schools did not comply with the policy change prohibiting such suspensions (strong evidence).**

As shown in Figure 1 (Panel A), total per capita OSS in Philadelphia was falling steadily in the pre-reform period (i.e., 2005–06 through 2011–12), while total out-of-school suspensions (OSS) among all other Pennsylvania districts remained fairly stable (on average). Specifically, during the pre-reform period, total OSS fell by 37 percent in Philadelphia—from 0.38 (or 38 suspensions per 100 students) in 2005–06 to 0.26 (or 26 suspensions per 100 students) in 2011–12. Among all other Pennsylvania districts, total per capita OSS declined by 11 percent—from 0.06 in 2005–06 to 0.05 in 2011–12. For conduct OSS, the target of Philadelphia's policy reform, a similar trend is apparent. Specifically, conduct OSS in Philadelphia declined from 0.32 to 0.23 during the pre-reform period, while average conduct OSS among all other Pennsylvania districts declined from 0.045 to 0.034 (see Figure 1, Panel B). The pre-policy trend for non-conduct OSS also shows that they declined, but less steadily over time than total and conduct OSS (see Figure 1, Panel C).

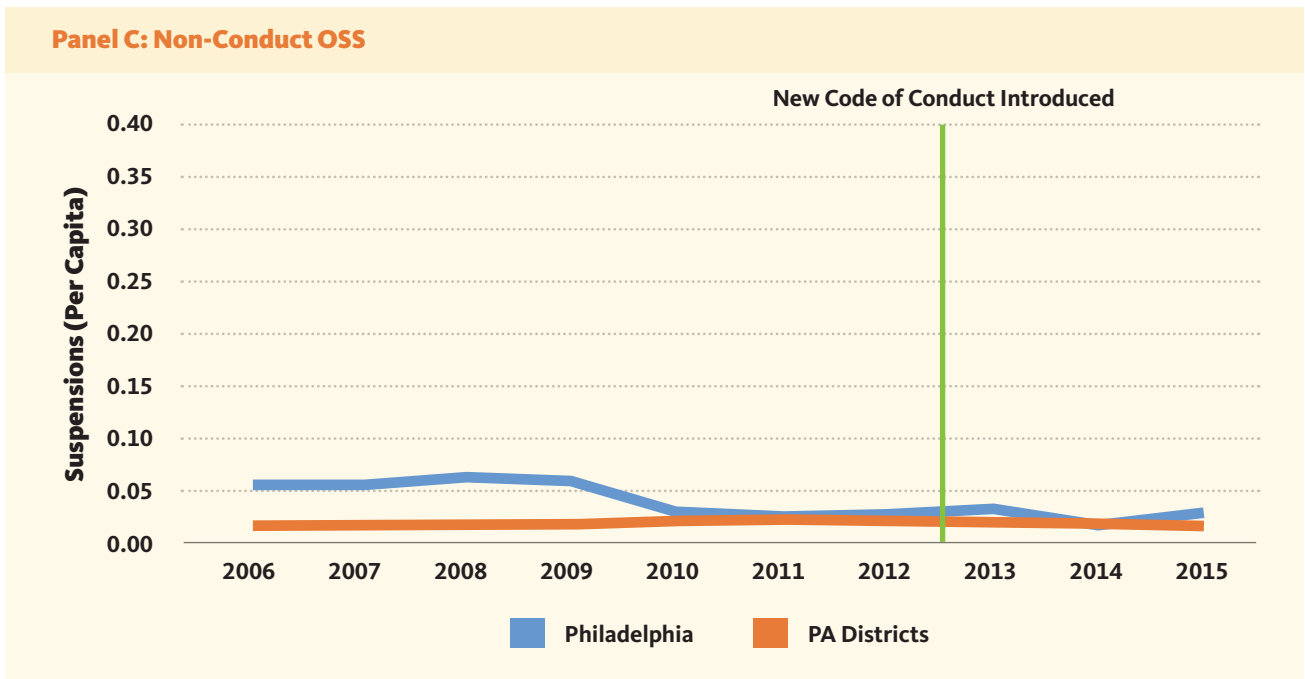


● **FIGURE 1:** Total, conduct, and non-conduct out-of-school suspensions (OSS) were all declining prior to the policy change.



Note: PA Districts include n=482 districts.

- **FIGURE 1 (continued):** Total, conduct, and non-conduct out-of-school suspensions (OSS) were all declining prior to the policy change.



**Note:** PA Districts include n=482 districts.

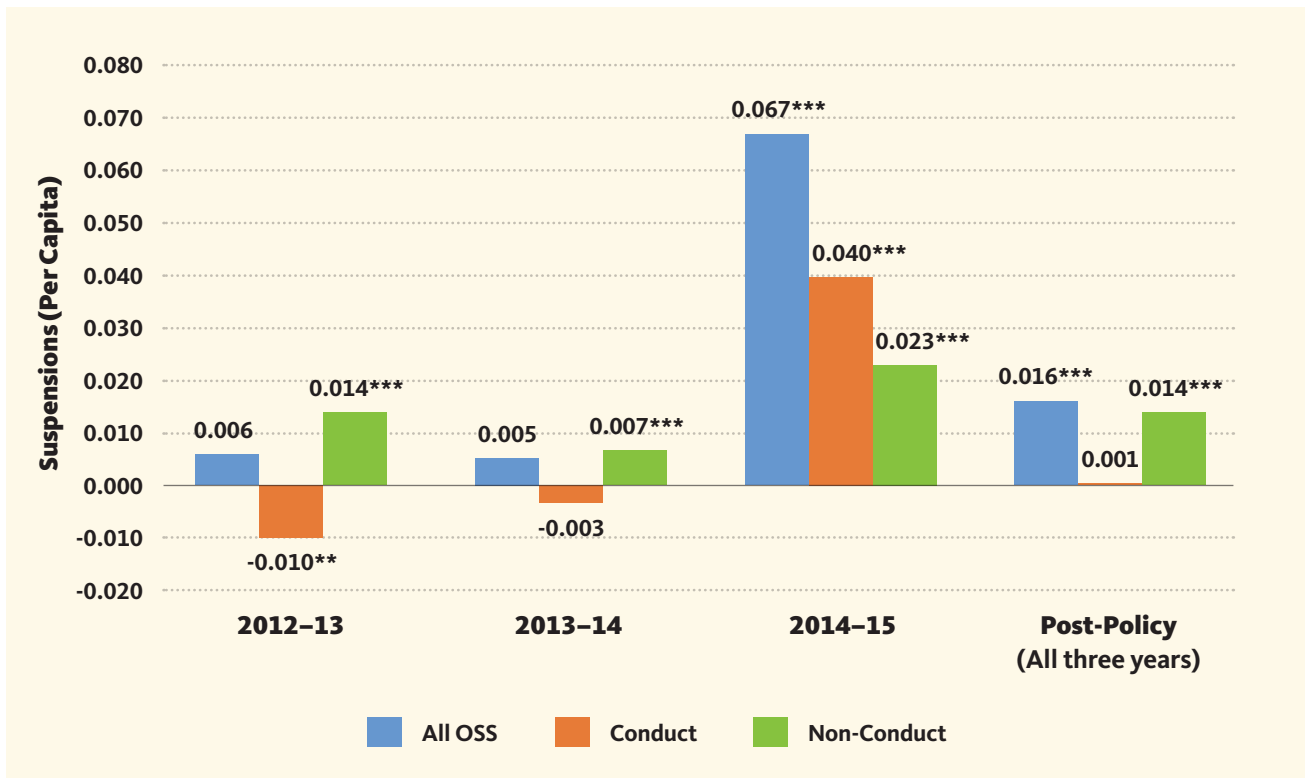
Those are the raw trends. But what was the effect of Philadelphia’s discipline policy reform?

To answer this question, we turn to our difference-in-differences approach, which compares changes in suspensions in Philadelphia to changes in suspensions among all Pennsylvania school districts (see Figure 2). First, we estimate that in the three years following Philadelphia’s policy change total OSS increased on average by 1.6 suspensions per 100 students per year. This average masks important year-specific variation. Notably, in the first two post-policy years—2012–13 and 2013–14—we do not find any significant change (positive or negative) in total OSS. However, by the third post-policy year (2014–15), total OSS increased by approximately 7 suspensions per 100 students.

Second, we find that there was no overall change in the rate of conduct suspensions in Philadelphia as a result of changes in district policy. However, this estimate also varies by post-policy year, with an initial decrease in conduct suspensions in 2012–13 of 1 fewer conduct OSS per 100 students, no statistically significant change in conduct suspensions in 2013–14, and a larger, offsetting increase in conduct suspensions in 2014–15.

Third, we find consistent evidence that non-conduct OSS in Philadelphia increased by 1.4 suspensions per 100 students in the post-policy period, relative to other Pennsylvania districts. In fact, non-conduct OSS increased in each of the three post-policy years.

● **FIGURE 2:** Conduct suspensions declined in the year following the policy reform. However, because the number of non-conduct suspensions increased, overall suspensions increased.



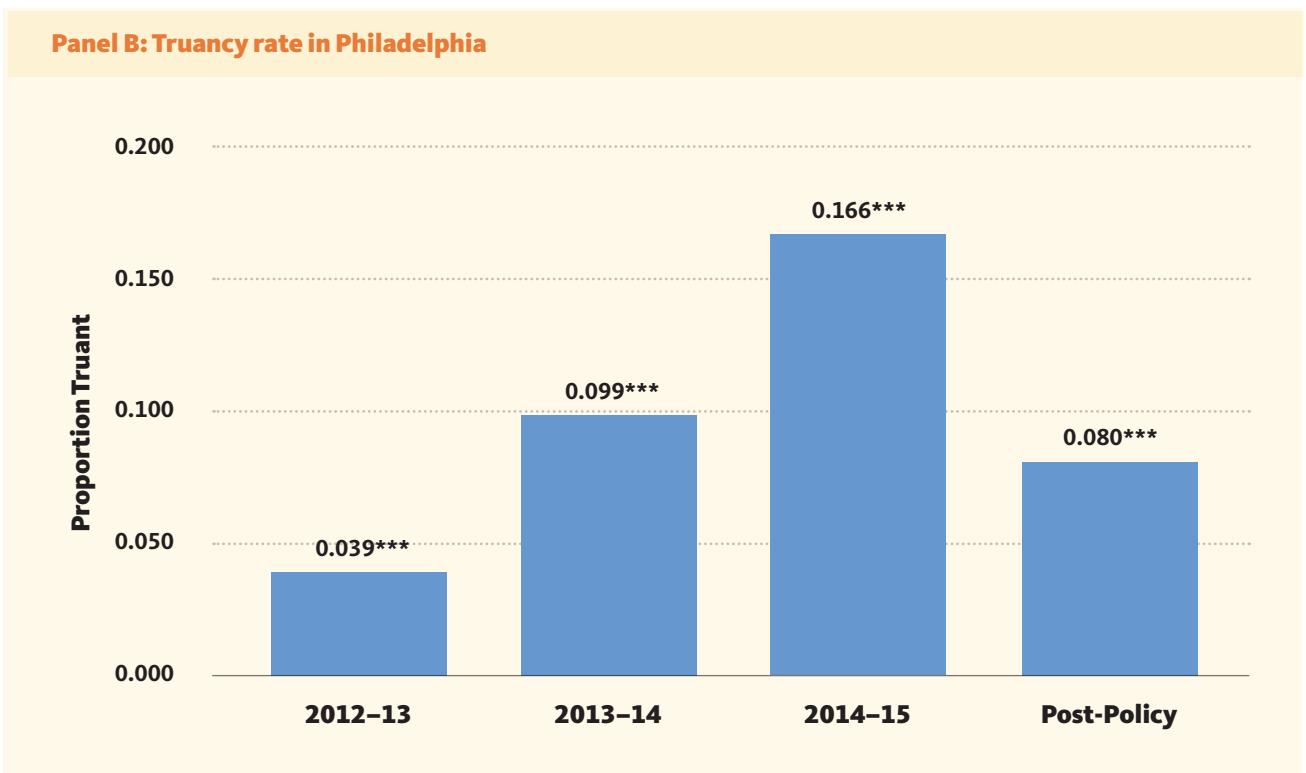
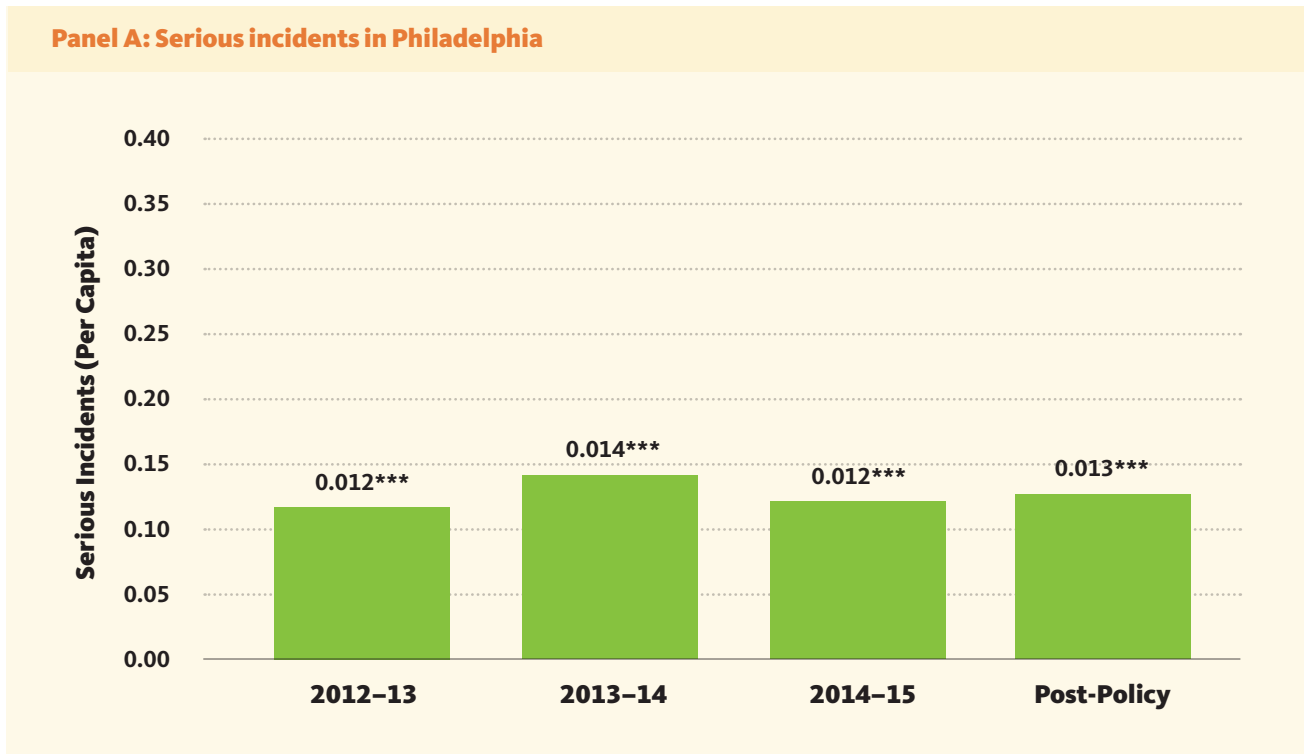
**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.

**How to read this figure:** In the first year following the discipline policy reform (2012–13), there was 1 fewer conduct suspension in Philadelphia (per one hundred students) due to changes in district policy, as compared to other Pennsylvania districts (see orange bar). In the second post-policy year, there was a smaller decrease in conduct suspensions (see second orange bar). In year 3 (2014–15), there was an increase in conduct suspensions of 4 suspensions (per one hundred students). In all three years, there were increases in non-conduct suspensions (green bars). *Post-Policy* indicates the average annual change across the three post-policy years: 2012–13, 2013–14 and 2014–15.

Finally, consistent with the finding for non-conduct OSS, we find that the policy change led to sustained increases in “serious incidents” (1.3 per one hundred students) in each post-policy year (see Figure 3, Panel A), as well as increases in the truancy rate (see Figure 3, Panel B).

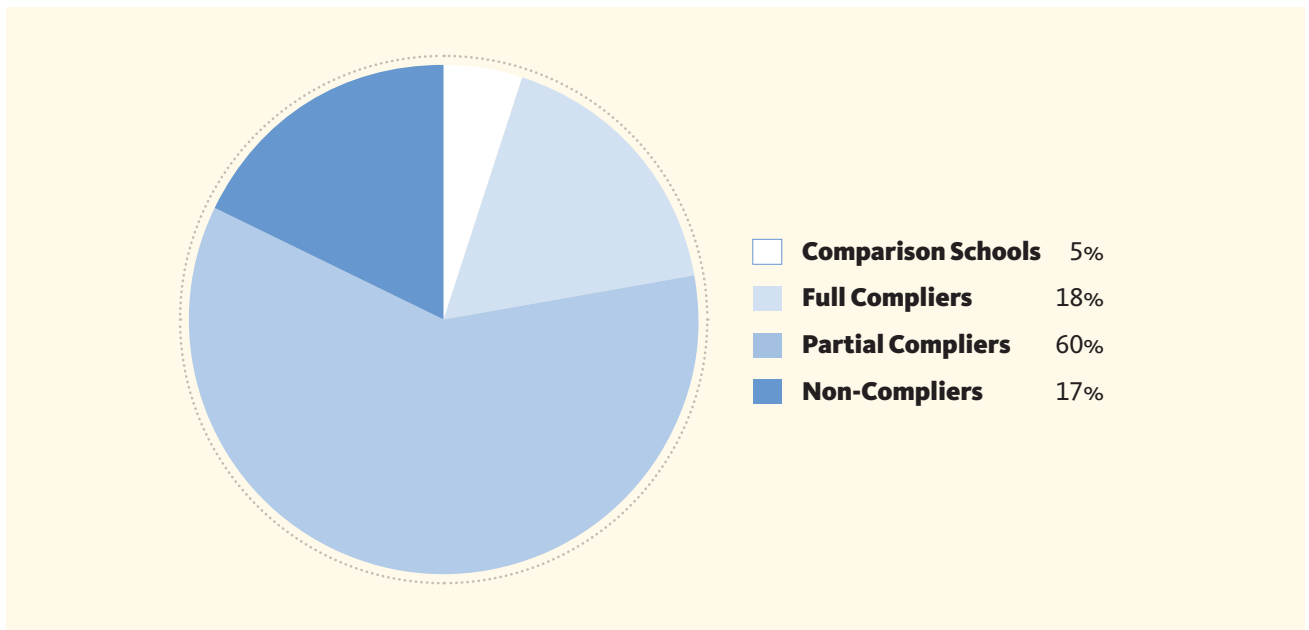
While the number of suspensions declined immediately following the policy change, some schools in Philadelphia continued to suspend students for low-level conduct violations, even though the revised code of conduct no longer included OSS as a consequence for these offenses. This raises the question of which schools continued to suspend students for conduct violations and which fully complied with the new policy. Fortunately, access to student-level data allows us to observe the extent to which each of the 238 traditional (i.e., non-charter) schools in Philadelphia that were open in both the 2011–12 and 2012–13 school years complied with the policy change.

● **FIGURE 3:** The number of serious incidents and the truancy rate increased in Philadelphia following the policy change.<sup>20, 21</sup>



**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.

● **FIGURE 4:** Less than a quarter of Philadelphia schools fully complied with the ban on low-level conduct suspensions.



Based on their level of compliance, we divide schools into four categories (see Figure 4):

*Twelve schools (or 5 percent) were “comparison schools,” meaning they had no conduct suspensions before or after the policy change and therefore did not change their practice as a result of the policy. Of the students in these schools, 51 percent were eligible for free and reduced-price lunch, 53 percent were black or Hispanic, 77 percent were proficient in math, and 68 percent were proficient in English language arts.<sup>22</sup> Prior to the policy change, only 2 percent of students in these schools were suspended for any infraction in a given school year.*

*Forty-three schools (or 18 percent) were “full compliers,” meaning they eliminated conduct suspensions as required by the policy. Of the students in these schools, 65 percent were eligible for free and reduced-price lunch, 70 percent were black or Hispanic, 58 percent were proficient in math, and 51 percent were proficient in English language arts. Prior to the policy change, 11 percent of students in these schools were suspended for any infraction in a given school year. (Three percent were suspended for either profanity or failure to follow classroom rules.)*

*One hundred and forty-two schools (or 60 percent) were “partial compliers,” meaning they reduced—but did not eliminate—conduct suspensions. Of the students in these schools, 65 percent were eligible for free and reduced-price lunch, 76 percent were black or Hispanic, 50 percent were proficient in math, and 43 percent were proficient in English language arts. Prior to the policy change, 16 percent of students in these schools were suspended for any infraction*



in a given school year. (Six percent were suspended for either profanity or failure to follow classroom rules.)

*Forty-one schools (or 17 percent) were “non-compliers,” meaning (in most cases) that they increased conduct suspensions.*<sup>23</sup> Of the students in these schools, 65 percent were eligible for free and reduced-price lunch, 85 percent were black or Hispanic, 47 percent were proficient in math, and 42 percent were proficient in English language arts. Prior to the policy change, 13 percent of students in these schools were suspended for any infraction in a given school year. (Three percent were suspended for either profanity or failure to follow classroom rules.)

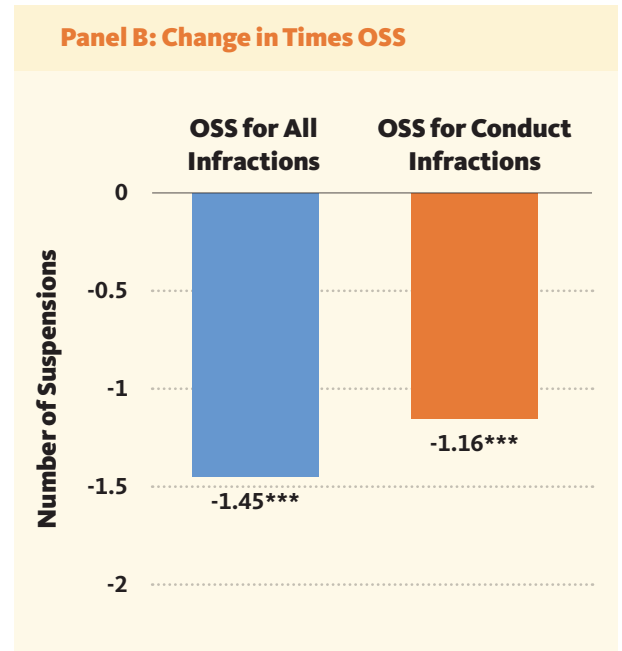
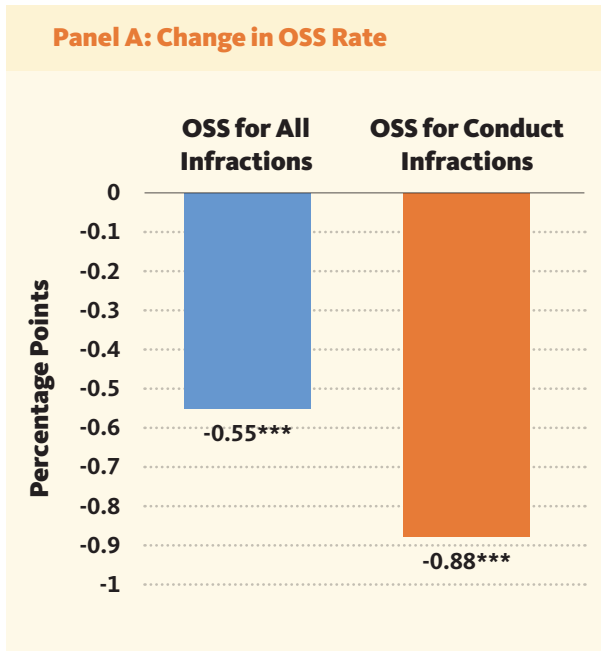
As these numbers suggest, schools’ responses to the policy change were remarkably varied. For example, among full compliers, the mean conduct suspension rate declined from an average of 2.5 percent in 2011–12 to zero percent in 2012–13. Similarly, among partial compliers, it declined from 6 percent in 2011–12 to 3 percent in 2012–13. However, contrary to the explicit intent of the policy change, the conduct suspension rate for non-complier schools *doubled* from 3 percent in 2011–12 to 6 percent in 2012–13.

## **Question 2: Was the policy reform associated with changes in suspensions, achievement, and school attendance for students who were suspended prior to the reform?**

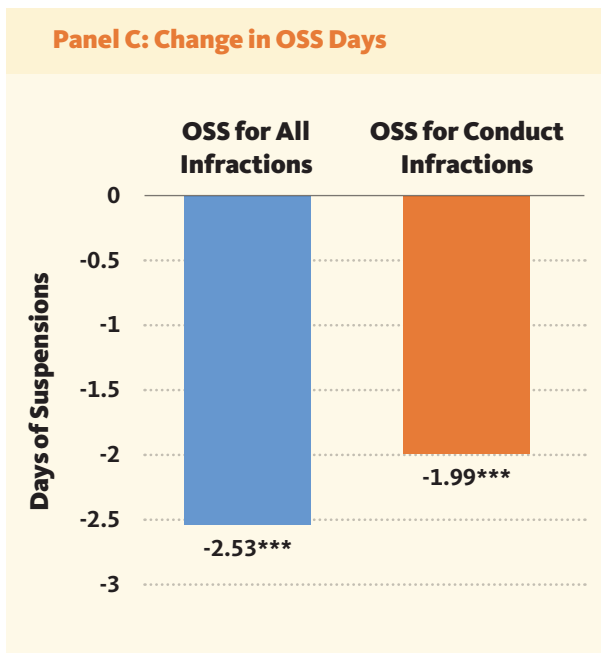
**Previously suspended students were less likely to be suspended after the policy change (associational evidence).**

Students who were suspended for conduct infractions in the year prior to the policy change were less likely to be suspended in the wake of the policy change, and they were far less likely to receive a conduct suspension (see Figure 5, Panel A). Students suspended for conduct infractions in the pre-reform year received, on average, 1.45 fewer suspensions for any reason and 1.20 fewer conduct suspensions (see Figure 5, Panel B). Furthermore, students suspended for conduct infractions in the pre-reform year received 2.50 fewer days of suspension for any reason and 1.99 fewer days of conduct suspension (see Figure 5, Panel C).

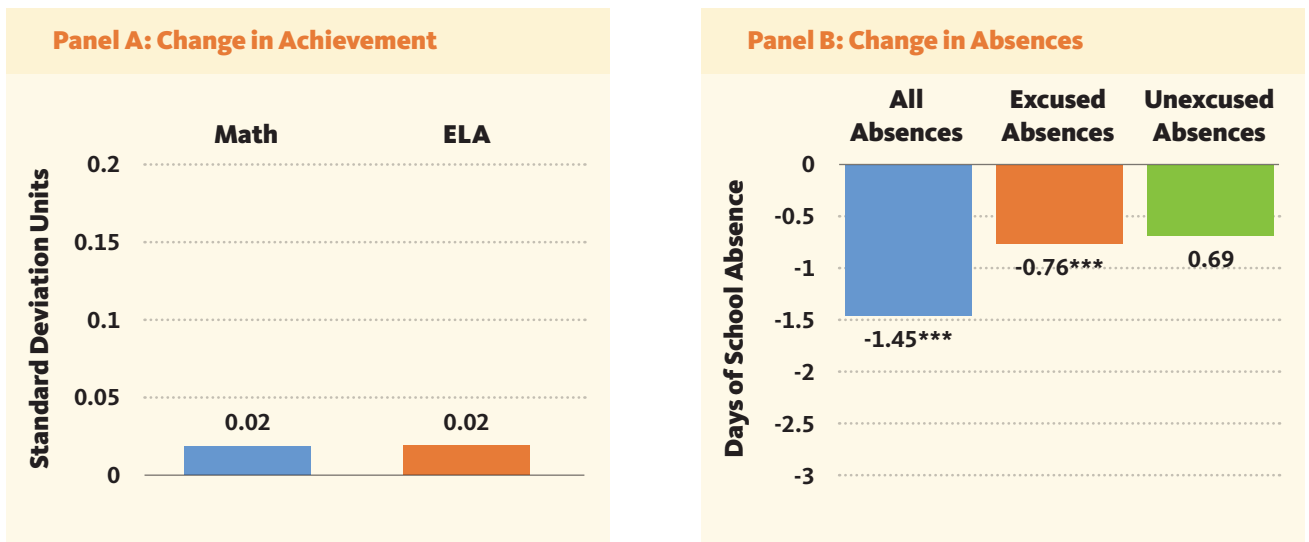
● **FIGURE 5:** Students who were previously suspended for a conduct infraction were less likely to receive a conduct suspension following the policy change.



**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.



● **FIGURE 6:** Changes in district policy were associated with improved attendance—but not improved achievement—for previously suspended students.



**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.

To estimate the association between the policy change and student outcomes—meaning achievement and school absences—for suspended students, we first identified students by their suspension status prior to the policy change. Students who received a conduct suspension in the pre-reform year (2011–12) were considered the group “treated” by the policy change; students who did not receive a conduct suspension in the pre-reform year were considered the comparison group. **Our estimates indicate that the policy change was not associated with changes in achievement among previously suspended students** (see Figure 6, Panel A).

**However, there were improvements in attendance following the policy change.**

Students suspended prior to the policy change had 1.45 fewer school absences in the year after the change, compared to students who were not suspended pre-reform (see Figure 6, Panel B).<sup>24</sup>

### Question 3: Was the policy reform associated with changes in achievement and school attendance for peers who were not suspended prior to the reform?

Peers who did not receive a conduct suspension prior to the change experienced worse outcomes in schools that didn’t (or couldn’t) comply with the policy change prohibiting conduct suspensions (associational evidence).

Prior research suggests that “peer effects” are strongest at the classroom level. However, because we don’t have data on course taking, for the purposes of this report, non-offending peers include any student in the same school and grade as a suspended student, which is the next best thing to defining “peer” at the classroom level. This means that our estimates of discipline policy change on peers are necessarily based on comparisons between students in different schools (but the same grade) rather than between students in different classrooms within the same school and grade.

Of the four groups of schools, the comparison schools are the nearest thing to a control group because they did not have any conduct suspensions in either the pre- or post-policy years, meaning they were unaffected by the policy change. Consequently, we can get some sense of how the policy change influenced peer outcomes by comparing peers in the three other school types to their comparison school counterparts. Here is what this comparison revealed for each of those three school types:

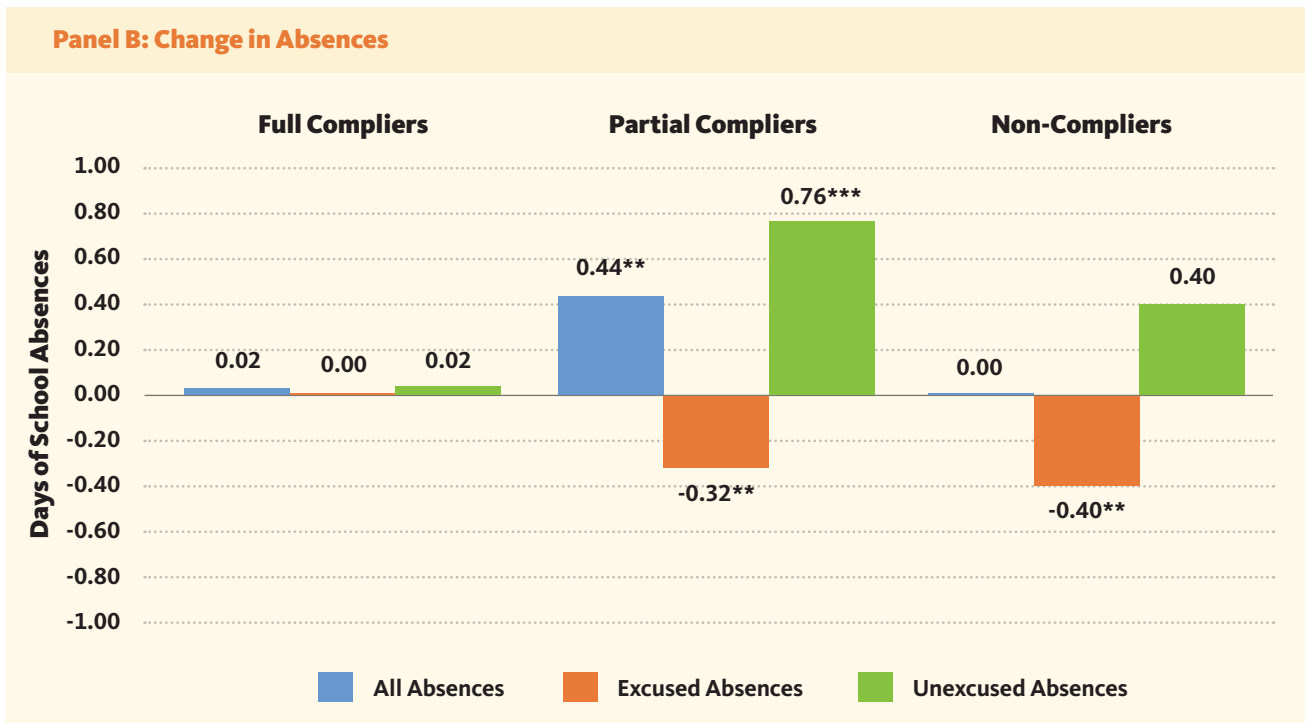
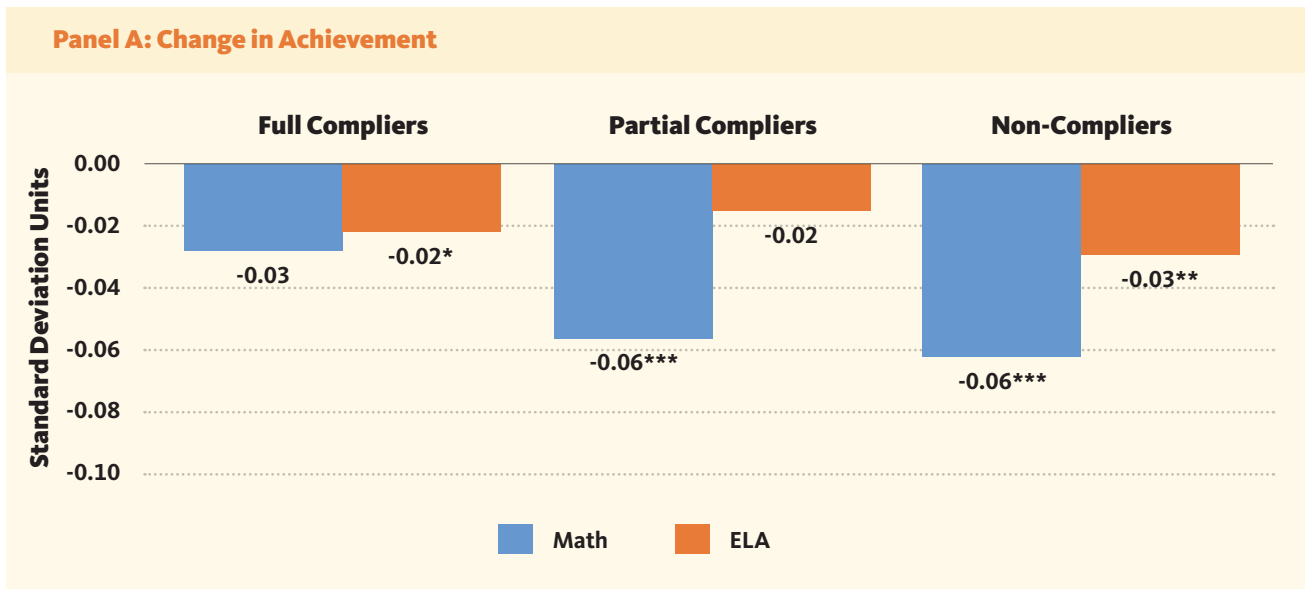
**1 Full compliers:** Recall that these are the 18 percent of schools that eliminated conduct suspensions as required by the policy. We estimate that peers in full complier schools experienced *no significant changes* in achievement (see Figure 7, Panel A) or attendance (see Figure 7, Panel B) relative to their non-suspended peers in comparison schools.

**2 Partial compliers:** These are the 60 percent of schools that reduced, but did not eliminate, conduct suspensions in the post-reform year. In contrast to non-suspended peers in full complier schools, peers in partial complier schools experienced a 0.06 standard deviation *decline in math achievement*, relative to their comparison school counterparts. Total absences increased by 0.44 days per student (or forty-four days per one hundred students), representing a 3 percent increase over 2011–12 levels. The increase in total absences was driven by an increase in unexcused absences, on the order of 0.76 days per student and representing an 8 percent increase over 2011–12 levels.

**3 Non-compliers:** These are the 17 percent of schools that (in all but one case) increased conduct suspensions in the post-reform year. Peers in non-complier schools experienced a 0.06 standard deviation *decline in math achievement* and a 0.03 standard deviation *decline in ELA achievement*, relative to their comparison school counterparts. We do not, however, find any change in total absences following the district’s policy change.

In short, average peer math achievement declined in the 77 percent of Philadelphia schools that did not fully comply with the policy change, relative to the most plausible comparison group. As noted in the first finding, partial complier schools had higher average suspension rates than full complier schools in the year prior to the policy change. Thus, one reasonable

**FIGURE 7:** Peers in Full Compliers experienced *no significant changes* in achievement or attendance relative to peers in comparison schools. In contrast, peers in Partial Compliers and Non-Compliers experienced a decline in achievement and an increase in unexcused absences.



**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.



interpretation of these results is that a policy change prohibiting the use of conduct suspensions has more negative consequences for peers in schools that serve more disruptive students—perhaps because the marginal student who returns to the classroom is more disruptive. Conversely, the results for peers in full complier schools suggest that schools with fewer or less disruptive students may be able to reduce or potentially even eliminate conduct suspensions at little to no cost to peer achievement and attendance.

## **Question 4: Was the policy reform associated with a change in racial disproportionality?**

### **Revising the district’s code of conduct was associated with an increase in racial disproportionality at the district level (associational evidence).**

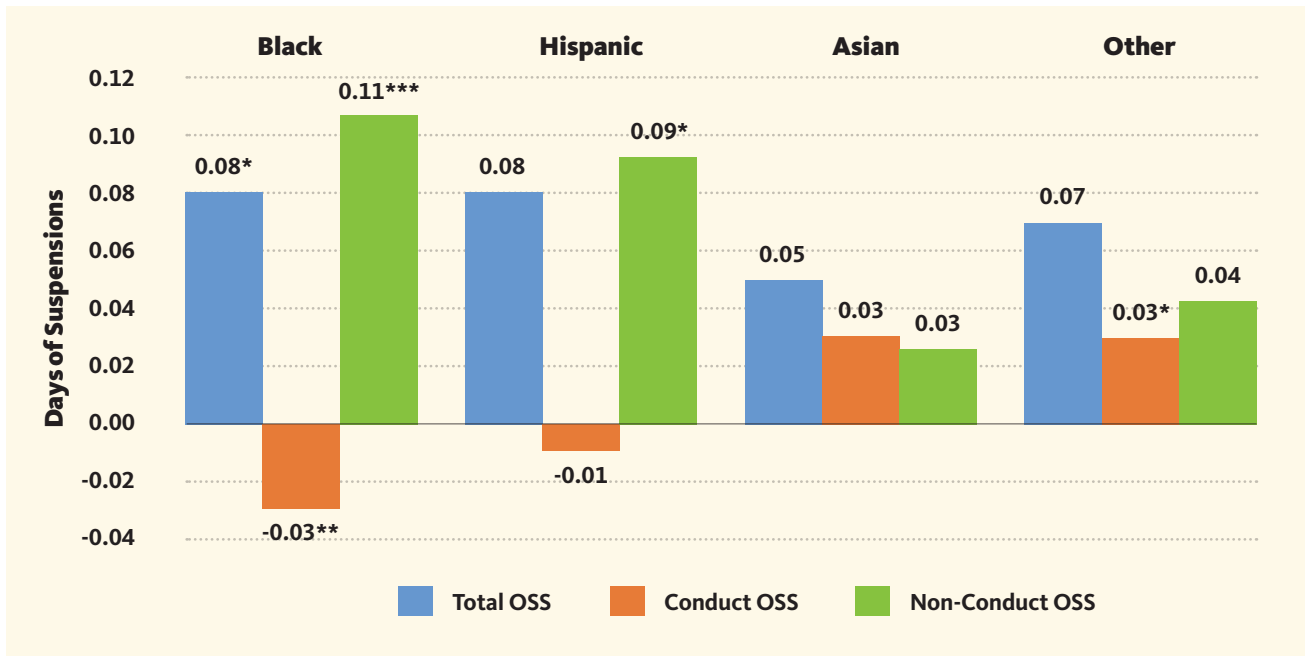
Philadelphia’s discipline policy change aimed to reduce the overall use of suspensions. However, because racial gaps in suspension use have been widely documented elsewhere,<sup>25</sup> we explored whether racial disproportionality in Philadelphia that existed prior to the policy reform changed in the post-reform years.

First, we confirm that pre-reform levels of suspension varied by racial groups. In 2011–12, black students in grades 3–12 received, on average, 0.77 days of suspension for any infraction (and 0.15 days of suspension for conduct infractions). Average days of suspensions for black students were approximately 2.5 times greater than the average days of suspension for white students, who received 0.29 days of suspension for any infraction (and 0.06 days of suspension for conduct infractions). Hispanic students in grades 3–12 received, on average, 0.48 days of suspension for any infraction (and 0.09 days for conduct infractions).

Second, we find that the black-white gap in suspensions for conduct infractions decreased in the wake of the policy change, though modestly. Compared to the 2011–12 school year, the black-white gap in conduct suspensions declined by 0.03 days per student in 2012–13 (see Figure 8). However, any goal to reduce racial disproportionality in the overall use of suspensions was not achieved as a consequence of the policy change because this decline was more than offset by an increase in suspensions for more serious incidents among black students, on the order of 0.11 days of suspension (relative to white students). As a consequence of the increase in suspensions for more serious infractions in the first post-policy year, black students experienced an average increase in suspension days of eight days per one hundred students relative to white students. We also find nearly identical patterns of suspensions for Hispanic students.

At least two factors seem to have contributed to this unintended result. First, schools with lower achievement, higher levels of unexcused absences, and more minority pupils

● **FIGURE 8:** Days of out-of-school suspension increased for black and Hispanic students, relative to white students, following the policy change.



**Note:** Asterisks denote statistical significance at the \*10%, \*\*5%, and \*\*\*1% levels.

experienced smaller declines in conduct suspensions than other schools. For example, as noted in the first finding, 85 percent of students in non-complier schools were black or Hispanic, compared to 76 percent of students in partial compliers, 70 percent of students in full compliers, and 53 percent in comparison schools.

Second, the suspension rate for more serious non-conduct offenses actually *increased* in the non-complier and partial complier schools between the 2011–12 and 2012–13 school years. Specifically, in 2011–12, 10 percent of students attending either a non-complier or partial complier school were suspended for non-conduct offenses. Similarly, in 2012–13, 12 percent and 11 percent were suspended for these more serious offenses in non-complier and partial complier schools, respectively. Meanwhile, full complier schools experienced a slight decline in non-conduct suspensions—from 8 percent to 7 percent—between 2011–12 and 2012–13, while in comparison schools only 2 percent of students received non-conduct suspensions in both 2011–12 and 2012–13. One possible explanation for this pattern is that at least some conduct offenses committed by non-white students may have been reclassified as non-conduct offenses as a result of the policy change. Alternatively, the increase in racial disproportionality could be the result of changes in student behavior (consistent with the estimated increase in serious incidents following the policy reform). Compliance with the district’s revised policy on expulsions, however, explains at most a fraction of this increase due to the relative infrequency of expulsions relative to suspensions.

# Takeaways

## **1 The discipline policy change in Philadelphia initially reduced suspensions for conduct infractions targeted by the reform, but this reduction did not persist.**

Our results show that suspensions for conduct infractions declined in the first year of implementation of the district-level policy change, relative to other districts in Pennsylvania, but this decline did not persist into the second and third years. Furthermore, a deeper look within the School District of Philadelphia demonstrates that the degree of adherence to the policy varied across schools. Some schools fully complied by eliminating conduct suspensions, while others partially complied by reducing but not eliminating these suspensions. And some schools actually increased conduct suspensions. This last group of schools also served a lower-achieving and more minority student population than schools that at least partially complied with the district's policy change.

## **2 Changes in peer outcomes following the policy change varied at the school level and likely depend on the same contextual factors that drive a school's suspension rate, such as the severity of its disciplinary challenges and its capacity to implement and sustain alternative disciplinary approaches.**

If there is any clear takeaway from our results for peers, it is that context matters. Peer math achievement declined and school absences increased in schools that did not fully comply with the district's policy change. In contrast, we found no adverse academic or attendance consequences for peers in schools that reduced conduct suspensions to zero. Notably, full complier schools served a more academically advantaged student population, and had lower suspension levels—both overall and for conduct infractions—in the pre-reform year than did partial complier schools.

Our results suggest that a policy change prohibiting suspensions for conduct infractions may have more negative consequences in schools that serve more students who are struggling academically and that have higher suspension rates. Since some schools are better positioned to implement district-level policy changes, these results suggest that the intended effects of policies designed to reduce suspensions may not be realized uniformly across schools with varying academic and behavioral climates.

### **3 Changing a district’s code of conduct to reduce or eliminate low-level suspensions can have unintended consequences.**

Our results suggest that changing a district’s code of conduct to reduce or eliminate the incidence of low-level suspensions may have unintended consequences. Compared to other districts in Pennsylvania, Philadelphia schools experienced increases in suspensions for more serious infractions in the years following the policy change. Furthermore, the policy change may have had a negative impact on peers in schools facing disciplinary challenges, and it may have inadvertently increased (rather than reduced) racial disproportionality because suspensions declined more quickly for some groups than others.

The evidence that schools in Philadelphia encountered implementation challenges is considerable: a supermajority of schools did not fully comply with the policy change, and the poorest, most racially homogenous, and most academically challenged schools were the least compliant. Peers in these schools (including those that reduced their conduct suspension rates) experienced a decline in performance, relative to the most plausible comparison group. Finally, the district-wide decline in conduct suspensions for black students coincided with an *increase* in the number of non-conduct suspensions for minority students. As a result, racial disproportionality actually increased in the wake of the district’s policy change.

### **4 To minimize the potential for unintended consequences, any changes to a district’s code of conduct that are designed to reduce or eliminate suspensions should be coupled with additional supports for schools that face disciplinary challenges.**

Districts should provide additional resources and disciplinary options to support schools so that they can implement policy changes fully and sustainably. If policymakers are focused on reducing the number of low-level suspensions by changing the statutory penalties for non-violent misconduct, they should couple district-level reforms with additional resources that support teacher training in alternative discipline strategies—preferably ones that are backed by promising research, such as Positive Behavioral Interventions and Supports (PBIS), which has been shown to successfully decrease suspensions and improve student perceptions of school safety.<sup>26</sup>

Ultimately, schools that are struggling the most with student misconduct—those that are the lowest achieving and the most racially segregated—likely require such supports if they are to successfully implement discipline reforms without adversely affecting the majority of students who are not subject to behavioral consequences such as out-of-school suspensions.

# Endnotes

- 1 J. Adams, “Most Teachers in California Say They Need More Training in Alternatives to Suspensions, Survey Finds” (Oakland, CA: EdSource, May 7, 2017), <https://edsource.org/2017/most-teachers-in-california-say-they-need-more-training-in-alternatives-to-suspensions-survey-finds/581195>.
- 2 S. Kakar, “Understanding the Causes of Disproportionate Minority Contact: Results of Focus Group Discussions,” *Journal of Criminal Justice* 34 (July 2006), <https://www.researchgate.net/publication/222962491>, 369–381; C. Bradshaw et al., “Examining the Effects of Schoolwide Positive Behavioral Interventions and Supports on Student Outcomes: Results from a Randomized Controlled Effectiveness Trial in Elementary Schools,” *Journal of Positive Behavior Interventions* 12, no. 3 (2010), <http://journals.sagepub.com/doi/abs/10.1177/1098300709334798>, 133–148.
- 3 M. Henderson et al., “The 2015 EdNext Poll on School Reform,” *Education Next* 16, no. 1 (Winter 2016), <http://educationnext.org/2015-ednext-poll-school-reform-opt-out-common-core-unions/>.
- 4 S. Carrell and M. Hoekstra, “Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone’s Kids,” *American Economic Journal: Applied Economics* 2, no. 1 (January 2010), [https://www.jstor.org/stable/25760199?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/25760199?seq=1#page_scan_tab_contents), 211–228.
- 5 S. Imberman et al., *Katrina’s Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees*, Working Paper 15291 (Washington, D.C.: National Bureau of Economic Research, 2009), <http://www.nber.org/papers/w15291.pdf>.
- 6 Public Agenda, *Teaching Interrupted: Do Discipline Policies in Today’s Public Schools Foster the Common Good?* (New York, NY: Public Agenda, May 2004), <http://files.eric.ed.gov/fulltext/ED485312.pdf>.
- 7 American Association of School Superintendents, “School Discipline Consensus Project Survey of School System Leaders” (Alexandria, VA: AASA, n/d), <http://blogs.edweek.org/edweek/rulesforengagement/AASA%20School%20Climate%20Survey.pdf>.
- 8 M. Steinberg and J. Lacoë, “What Do We Know about School Discipline Reform? Assessing the Alternatives to Suspensions and Expulsions,” *Education Next* 17, no. 1 (Winter 2017), <http://educationnext.org/what-do-we-know-about-school-discipline-reform-suspensions-expulsions/>.
- 9 As of the 2016–17 school year, approximately 135,000 students attended one of 220 traditional public schools, and approximately 65,000 students attended one of eighty-eight charter schools. Source: School District of Philadelphia (<https://www.philasd.org/performance/school-information/>).

- 10 D. Losen and J. Gillespie, *Opportunities Suspended: The Disparate Impact of Disciplinary Exclusion from School* (Los Angeles, CA: Los Angeles, CA: University of California, Los Angeles, Civil Rights Project, August 2012), <http://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-ccrr-research>.
- 11 A. Gregory et al., “The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin?” *Educational Researcher* 39, no. 59 (2010), <https://pdfs.semanticscholar.org/5e29/3a4598fece049e2e33b1254dd46d73106780.pdf>.
- 12 R. Skiba et al., “Parsing Disciplinary Disproportionality: Contributions of Infraction, Student, and School Characteristics to Out-of-School Suspension and Expulsion,” *American Educational Research Journal* 51, no. 4 (August 2014), <http://journals.sagepub.com/doi/abs/10.3102/0002831214541670>, 640–670.
- 13 E. Arcia, “Achievement and Enrollment Status of Suspended Students: Outcomes in a Large, Multicultural School District,” *Education and Urban Society* 38, no. 3 (2006), <http://journals.sagepub.com/doi/abs/10.1177/0013124506286947>, 359–369; J. Davis and W. Jordan, “The Effects of School Context, Structure, and Experiences on African American Males in Middle and High School,” *Journal of Negro Education* 63, no. 4 (Fall 1994), <https://eric.ed.gov/?id=EJ511149>, 570–587; M. Rausch and R. Skiba, “The Academic Cost of Discipline: The Relationship Between Suspension/Expulsion and School Achievement” (Bloomington, IN: Center for Evaluation and Education Policy, School of Education, Indiana University, April 2005), <http://www.indiana.edu/~atlantic/wp-content/uploads/2014/03/Academic-Cost-of-School-Discipline.pdf>; A. Noltemeyer et al., “Relationship between School Suspension and Student Outcomes: A Meta-Analysis,” *School Psychology Review* 44, no. 2 (June 2015), <http://naspjournals.org/doi/10.17105/spr-14-0008.1>, 224–240.
- 14 M. Marchbanks III, et al., “The Economic Effects of Exclusionary Discipline on Grade Retention and High School Dropout,” in D. Losen (Ed.), *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion* (New York, NY: Teachers College Press, 2015), <https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/state-reports/the-economic-effects-of-exclusionary-discipline-on-grade-retention-and-high-school-dropout/marchbanks-exclusionary-discipline-ccrr-conf.pdf>, 59–74.
- 15 R. Balfanz et al., “Sent Home and Put Off Track: The Antecedents, Disproportionalities, and Consequences of Being Suspended in the Ninth Grade,” in D. Losen (Ed.), *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion* (New York, NY: Teachers College Press, 2015), <https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/state-reports/sent-home-and-put-off-track-the-antecedents-disproportionalities-and-consequences-of-being-suspended-in-the-ninth-grade/balfanz-sent-home-ccrr-conf-2013.pdf>, 17–30; R. Rumberger and D. Losen, *The High Cost of Harsh Discipline and its Disparate Impact* (Los Angeles, CA: Los Angeles, CA: University of California, Los Angeles, Civil Rights Project, June 2016), [https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/the-high-cost-of-harsh-discipline-and-its-disparate-impact/UCLA\\_HighCost\\_6-2\\_948.pdf](https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/the-high-cost-of-harsh-discipline-and-its-disparate-impact/UCLA_HighCost_6-2_948.pdf); A. Noltemeyer et al., “Relationship between School Suspension and Student Outcomes: A Meta-Analysis.”
- 16 K. Wolf and A. Kupchik, “School Suspensions and Adverse Experiences in Adulthood,” *Justice Quarterly* 34, no. 3 (2017), <http://www.tandfonline.com/doi/full/10.1080/07418825.2016.1168475?scroll=top&needAccess=true>, 407–420.



- 17 N. Mader et al, “When Suspensions are Shorter: The Effects on School Climate and Student Learning,” Working Paper, 2016.
- 18 M. Eden, *School Discipline Reform and Disorder: Evidence from New York City Public Schools, 2012–2016* (Washington, D.C.: Manhattan Institute, March 2017), <https://www.manhattan-institute.org/sites/default/files/R-ME-0217v2.pdf>.
- 19 S. Snyder, “School Summit Focuses on Violence Prevention,” *Philadelphia Inquirer*, August 14, 2012, <http://www.philanthropynetwork.org/news/100002/School-summit-focuses-on-violence-prevention.htm>.
- 20 Coefficients are from difference-in-differences models (standard errors clustered at the district level). There are 483 districts in the sample across the 2005–06 through 2014–15 school years. Models that produced the regression estimates include district fixed effects, year fixed effects, district-specific linear time trends, and controls for the following district characteristics: the percent of students receiving free or reduced-price lunch, the percent of minority (i.e., African American or Hispanic) students, the percent of English language learners, the percent of special education/disabled students, serious incidents (per-capita), and district math achievement (the percent of district students academically proficient on the PSSA math exam).
- 21 Coefficients are from difference-in-differences models (standard errors clustered at the district level). There are 483 districts in the sample across the 2005–06 through 2014–15 school years. Models that produced the regression estimates include district fixed effects, year fixed effects, district-specific linear time trends, and controls for the following district characteristics: the percent of students receiving free or reduced-price lunch, the percent of minority (i.e., African American or Hispanic) students, the percent of English language learners, the percent of special education/disabled students, Total OSS (per-capita); and math achievement.
- 22 Descriptive statistics reflect school characteristics in the 2011–12 school year. For all school types, demographic changes pre- and post-policy implementation were minimal.
- 23 Of these forty-one schools, only one school’s use of conduct OSS remained unchanged between the 2011–12 and 2012–13 school years.
- 24 Note that days of school missed due to absences (for any reason, excused or unexcused) are independent of days of school missed due to suspension.
- 25 Steinberg and Lacoë, “What Do We Know about School Discipline Reform? Assessing the Alternatives to Suspensions and Expulsions.”
- 26 C. Bradshaw et al., “Examining the Effects of Schoolwide Positive Behavioral Interventions and Supports on Student Outcomes.”