Banning Progress: Suspension Bans and Schoolwide Academic Growth

The past decade has seen a burgeoning practical and theoretical interest in reforming how schools manage discipline. A growing number of states and school districts have banned or limited suspensions for all but the most serious offenses. Part of the behavior management literature makes a strong case for suspension bans: financial savings, racial fairness, and academic growth. In terms of academic impact, however, a second perspective – shared by some educators – is that suspension bans harm student learning by removing an essential tool teachers use to manage their classrooms. To settle this debate, this paper uses a difference-in-difference research design to discover that the 2013 suspension ban in Los Angeles Unified School District led to a substantial, 0.2 standard deviation decrease in academic growth among schools that had previously issued the banned suspensions. Additional analyses alleviate concerns that the results are driven by selection bias, omitted variable bias, or a specific measurement of academic growth. The results suggest that states and districts should not ban suspensions; instead, they should carefully test more gradual efforts to decrease suspension rates.

Of the approximately 50 million students in U.S. public elementary and secondary schools, close to seven percent – 3.5 million – are suspended each year. Descriptive statistics suggest that at least some suspensions are racially biased and unnecessary: African-American, Latino, and American Indian students are at least twice as likely as other students to be suspended (Wallace et al. 2008), and suspension rates have doubled over the past several decades (Losen 2011).[1] Simultaneously, many schools and districts have successfully managed student behavior while issuing few to no out-of-school suspensions (e.g. Christle et al. 2005; Luiselli et al. 2005; Skiba & Sprague 2008).

A variety of actors want to reduce suspensions. Notably, the U.S. Department of Education promotes alternatives to suspension: the Behavior Education Program (Crone et al. 2010), function-based interventions (e.g. Liaupsin et al. 2006), and school-wide positive behavior support (Luiselli et al. 2005; Putnam et al. 2006; Skiba and Sprague 2008). The National Association of State Boards of Education recommends that states include suspension rates in their accountability systems (Charis and Losen 2017). Likewise, advocacy organizations emphasize the public costs of suspensions (e.g. Rumberger and Losen 2017) and highlight schools that report high overall or subgroup suspension rates (Losen et al. 2015). Given this environment, it is unsurprising that many schools and districts are eager to reduce their suspension rates. The fastest way to do that is to ban suspensions for all but the most serious offenses.

Some suspension critics say that such suspension bans will produce schoolwide academic benefits. They point to research that suspensions reduce academic performance for both suspended students and, in some cases, even their non-suspended classmates (Arcia 2006; Perry and Morris 2014). However, the literature provides a second perspective: that suspensions can counter disruptive behavior that would otherwise reduce learning opportunities for all students (McFarland 2001). Given the competing hypotheses, how do suspension bans impact schoolwide academic performance?

---

[1] Factors beyond school discipline policy greatly impact both of these trends:

- In 1975 Congress passed Public Law 94-142, which is now known as the Individuals with Disabilities Education Act (IDEA). This law has helped bring hundreds of thousands of students with disabilities from state institutions into public schools, as well as brought millions of students from segregated instructional environments into integrated classrooms. In addition to its many benefits, Public Law 94-142 also could explain some of rise in suspension rates. (https://www2.ed.gov/policy/speced/leg/idea/history.html)
- African-American, American Indian, and Latino students are more likely to be low-income, English Learners, and students with disabilities than students in other subgroups. These students are also less likely to be taught by teachers who share their racial and cultural background. This may explain some of the disproportionality in suspension rates.

The experience of Los Angeles Unified School District (LAUSD), the second largest school district in the country, provides an opportunity to answer that question. In May 2013, the LAUSD school board banned suspensions with the subjective rationale of "defiance," forcing a sudden and precipitous drop in its use. A difference-in-difference research design allows us to estimate how this sudden imposed change in suspension rates impacted academics. The first difference is temporal, comparing academic growth before and after the 2013 policy change. The second difference is geographic, comparing schools within LAUSD to those in the rest of California – i.e. schools that did not receive the "treatment" of a suspension ban. We test three distinct hypotheses, and all of them support the conclusion that the LAUSD suspension ban harmed academic growth.

This paper proceeds in five sections. The first recounts the history of suspension bans. Next is an overview of the competing theories and hypotheses the behavior management literature provides concerning suspension bans. The third section details the data and main analyses used to estimate the impact of LAUSD's policy decision. The forth describes the remaining threats to inference and how additional analyses alleviate many of those concerns. The paper concludes with implications of the findings and next steps for practitioners and researchers.

Unexpected Origins

Suspension bans emerged as a reaction to the zero tolerance approach to school discipline. Zero tolerance policies administer strict punishment for relatively minor rule violations, often regardless of the circumstances and without formal due process (Cerrone 1999). The logic of this approach borrows from the broken window theory in the criminology literature (Wilson and Kelling 1982). Like broken window theory, zero tolerance makes two primary assumptions. One concerns the relationship between minor incidents (e.g. talking rudely to a teacher) and major incidents (e.g. fighting someone). Zero tolerance approaches assume that the presence of minor incidents makes major incidents more likely. The second assumption is that instituting policies of strict punishment for minor violations will make these incidents less frequent. This will occur through some combination of convincing students who get punished to not violate the rule again and convincing other students to never violate the rule at all. If valid, then the zero tolerance approach should lead to fewer minor and major incidents.

Interestingly, zero tolerance policies became widespread in schools partly because of the passage of a federal gun law: the Gun-Free Schools Act of 1994. This law made federal education funding conditional on districts expelling students for a full calendar year if they brought a gun to school. This can be construed as a zero tolerance policy. Having a gun at school is relatively minor compared to brandishing or using a gun at school. This new law strictly punished the presence of guns as an effort to reduce the number of times any guns were brought onto or used on school campuses. Many states and districts used this law as a model to update their systems of discipline, adopting policies of automatic suspensions for a wide variety of relatively minor infractions (Skiba and Knesting 2001). By the early 2000s schools were suspending nearly twice as many students as in the 1970s (Wald and Losen 2003).

Suspension bans are an attempt to counter zero tolerance policies. Instead of automatic suspensions for minor infractions, suspension bans typically forbid schools from suspending students for all but the most serious infractions, such as those involving violence and drugs. LAUSD was the first major district to adopt such a ban for all grades. Surprisingly, the story of that first suspension ban starts with the federal tax reform law of 1986.

The 1986 tax law required nonprofit health organizations to start paying taxes. Blue Cross of California subsequently struggled to compete with for-profit competitors, so in 1996 it converted into an investor-owned for-profit called WellPoint Health Networks, Inc. (Kane 1997). This required the creation of The California Endowment (TCE), a nonprofit charitable organization with an endowment of $2.3 billion. Newspapers across the state expressed concern that TCE's enormous endowment would allow it to dominate policy debates, and that the public would be unable to ensure that TCE acted in its

interest.[2] Perhaps to allay these concerns, the website for TCE emphasizes that its 17-member board is extremely diverse and "is designed to reflect a cross-section of California's people and places."[3]

TCE produced a case study detailing its school discipline efforts in California (Martinez et al 2013). While self-commissioned histories should always be read with caution, the events described suggest that TCE indeed played an unexpected and indispensable role in building the anti-suspension movement in California. In 2010, TCE launched a multi-million dollar effort called Building Healthy Communities to improve the health of 14 areas in California. While collecting input from stakeholders, TCE staff were surprised to hear that the abundant use of school suspensions was harming students' social and emotional health. TCE's statewide policy team looked into school suspensions and found that it was "an issue that framed correctly could have legs in Sacramento" because it was a widespread problem that could be remedied through relatively small changes to state education law (Martinez et al. 2013, p. 7). California *Education Code* Section 48900(k) allows schools to suspend students if they have "[d]isrupted school activities or otherwise willfully defied the valid authority of supervisors, teachers, administrators, school officials, or other school personnel engaged in the performance of their duties." Defiance is the most subjective basis on which schools can issue suspensions; other rationales concern various forms of theft, violence, and possession of illegal substances. Many schools have made frequent use of this flexibility to remove defiant students from school. When the state began collecting suspension data with these categories in 2011-12, it reported over 200,000 based on defiance, constituting 39% of all suspensions.[4]

Having decided to focus on school discipline, TCE used its resources to connect interest groups to one another. In May 2011, TCE convened a discussion of school discipline with community organizers from 8 of the 14 communities along with a statewide advocacy organization called Fight Crime: Invest in Kids. TCE then created the School Discipline Action Team, a coalition of three distinct groups:

1. Community organizers, such as Community Asset Development Re-defining Education (CADRE) and Labor Community Strategy Center (LCSC), who had worked for over a decade with families suffering from zero tolerance policies.
2. Legal advocates, such as Public Counsel and the ACLU, who knew the technical details necessary to know how to change school discipline law.
3. State advocates, Children Now and Fight Crime: Invest in Kids, who were "new to the discipline issue" but also "were sophisticated, repeat players on the statewide scene" (Martinez et al. 2013, p. 8).

The School Discipline Action Team soon drafted 10 bills, but the community organizers lacked the expertise to keep up with the legal and state advocates; "even though CADRE and LCSC were formally involved when the legislative priorities were being hashed out in December 2011 and January 2012, they had limited ability to make substantive contributions" (Martinez et al. 2013, p. 11).

In the final phase, TCE executed a strategic communications plan designed to amplify and coordinate messages from a range of interest groups. TCE created a television commercial that aired in Sacramento as legislators considered the 10 bills. Additionally, TCE paid for a statewide poll about school discipline and strategically released those poll results simultaneously with recent research on suspensions. Seven of the ten bills passed both chambers, leaving it to Governor Jerry Brown to either sign or veto them. Weeks before Governor Brown made his decisions, TCE paid for all the facilities, rental, and travel costs for speakers to attend an event in Los Angeles to highlight the issue of harsh

[2] For example, see Thelen J., "Charity or Advocacy for HMOs?" *The Recorder*, 30 September 1994; Editorial, "Blue Doublecross?" *San Jose Mercury News*, 11 November 1994; Editorial, "The Blue Cross Octopus," *Sacramento Bee*, 6 December1994.

[3] http://www.calendow.org/our-story/#leadership

[4] For simplicity, "suspensions" without any qualifier refers to out-of-school suspensions.

school discipline. Governor Brown ended up signing five of the seven bills into law. One veto was the bill that would have imposed a statewide ban on suspensions based on defiance.

That veto was only a temporary setback. In 2013, some of the community organizing groups that TCE had supported pushed for LAUSD to ban all suspensions based on defiance. The ban was backed by Superintendent John Deasy, an education reformer, and approved by five out of seven members of the school board, which was frequently at odds with the superintendent. This surprising unity reflects the success of TCE's work at making the case for suspension bans in the court of California public opinion generally and among policymakers in particular.

Other districts and state soon started to adopt suspension bans. San Francisco Unified School District instituted its own ban in 2014, followed by Oakland Unified School District as well as grades K-3 in all of California in 2015 (Frey 2015). Since then the policy has spread across the country, with school boards issuing suspension bans for grades K-5 in Oregon[5] and grades K-2 in Texas (Reid-Cleveland 2017) and New York City (Berwick 2017). Additional bans are under consideration in cities such as Pittsburgh (Lindstrom 2017) and Philadelphia (Cline-Thomas and Chang 2017).

It is unclear whether policymakers are adopting suspension bans because they are facing similar incentives, or if policymakers are learning from each other through policy diffusion (Volden et al 2008). If it is policy diffusion, it would most likely be the imitation mechanism, with leaders copying the policy without considering its wider effects (Shipan and Volden 2008). The mechanisms of competition and coercion do not seem to apply, and the learning mechanism makes the assumption that subsequent adopters knew the impact of California's early suspension bans. The policy itself is very simple, which makes it easy to adopt (Makse and Volden 2011). Given the decentralized control of education policy, it is especially easy for state or local school boards to experiment with school discipline policies (Shipan and Volden 2012).

Advocacy groups continue to pressure policymakers to reduce suspensions. The Civil Rights Project out of UCLA is a leader of these efforts, using a multi-pronged strategy that includes appeals to potential legal action, monetary savings, and public exposure. The organization's report about the discipline gap described the huge disparities in out-of-school suspensions as a "potentially unlawful denial of educational opportunity" (Losen et al. 2015). Another report estimated that the lifetime cost of suspensions for one cohort of California high school students was $2.7 billion (Rumberger and Losen 2017). The Civil Rights Project simultaneously published an online dataset of individual districts' suspension rates, cost of suspensions, and potential benefit from discipline reform.[6]

We conclude this section by noting that the federal government has contributed to school discipline reform. Under the Obama administration, agencies pushed hard against racial disproportionality in suspension rates. In 2014, the U.S. Department of Education and Department of Justice published a "Dear Colleague" letter containing a school discipline guidance package.[7] The package emphasized that school discipline must be done without being discriminatory, and the inclusion of the Department of Justice signaled that the federal government would pursue legal action against districts who failed to comply. Indeed, the Department of Education had previously reached voluntary legal settlements with a number of school districts with racially disproportionate suspensions, including LAUSD in 2011 and Oakland Unified in 2012.[8] While the Education and Justice Departments are not

[5] For coverage in local news reports, see: http://gov.oregonlive.com/bill/2015/SB553/

[6] Available here: http://www.fixschooldiscipline.org/costsofdiscipline/

[7] Available here: https://www.ed.gov/news/press-releases/us-departments-education-and-justice-release-school-discipline-guidance-package-

[8] See https://www2.ed.gov/about/offices/list/ocr/docs/investigations/09105001-b2.pdf and https://www.ed.gov/news/press-releases/us-department-education-announces-voluntary-resolution-oakland-unified-school-di

prioritizing these efforts under the Trump administration, anti-suspension information and resources are still available. The Department of Education website still contains a section "Suspension 101" that includes the headlines "Suspensions don't work," "Suspensions have negative consequences," and "There are effective alternatives to suspension."[9]

Theory and Hypotheses

How does the literature expect a suspension ban to impact academic growth? Education researchers have not used the ideal research method of a randomized controlled trial to test behavior management policies. This is partly for ethical reasons: schools and districts cannot randomly suspend only some students for rule violations.[10] This has forced the literature to rely on sub-optimal methods to estimate the impact of a suspension ban.

One strand of the behavior management literature suggests an increase in growth through two mechanisms. First, students who would have been suspended without the policy should now have higher academic growth because they are spending more time learning in school; Arcia (2006) makes this claim based on a matching analysis. Second, even students who would not have been suspended should experience higher growth because they suffer less from the distraction of a punitive environment created by issuing too many suspensions. This second mechanism is based on Perry and Morris's work in *American Sociological Review* (2014), which employs fixed-effect regression models to conclude that giving higher-than-average numbers of suspensions lowers academic growth even for non-suspended students. They draw their theory from the criminology literature, based on parallels to the impact of mass incarceration. Both of these mechanisms would be most prominent in schools with the largest decrease in suspensions.

However, another strand of the literature proposes a different mechanism: that suspensions are a tool teachers can use to remove defiant students, thereby providing more opportunities for most students to learn. By carefully observing classrooms, McFarland (2001) sees that defiant behavior can harm both teachers and other students. Defiant behavior can derail teachers' plans, increase their stress, and – in the most extreme cases – even cause them to leave their positions. Teachers can take actions to prevent most defiant behavior from occurring, but this requires training and practice. If teachers have no tool other than suspensions to deal with defiance, then suddenly removing that tool would lead to classroom management problems. This would result in lost learning opportunities for all students.

In summary, theory provides opposing views about how a sudden decrease in suspensions would impact academic performance. In the context of the LAUSD suspension ban, we can test three distinct hypotheses based on the idea that the suspension ban impacted academics:

H1: Change in academic growth should be higher (lower) in LAUSD than in the rest of California
H2a: The gap found in H1 should be even larger for schools that gave the banned suspensions
H2b: There should be no gap for schools that did not give the banned suspensions
H3: Change in academic growth should be highest (lowest) in schools that used to give the most of those suspensions

Currently, the most methodologically rigorous analysis in the literature argues that natural fluctuations in suspension rates cause subsequent fluctuations in academic performance (Perry and

[9] https://www2.ed.gov/policy/gen/guid/school-discipline/index.html

[10] However, it would be ethically sound for districts to randomly assign schools to one of two (or more) equally promising discipline reform programs.

Morris 2014). This assumes that there are no omitted variables that might cause both changes.[11] However, omitted variables almost certainly are a problem. Factors such as sudden issues in students' home lives or peer relationships – which are impossible to measure precisely – are likely to cause both an increase in suspensions and a decrease in academic growth.

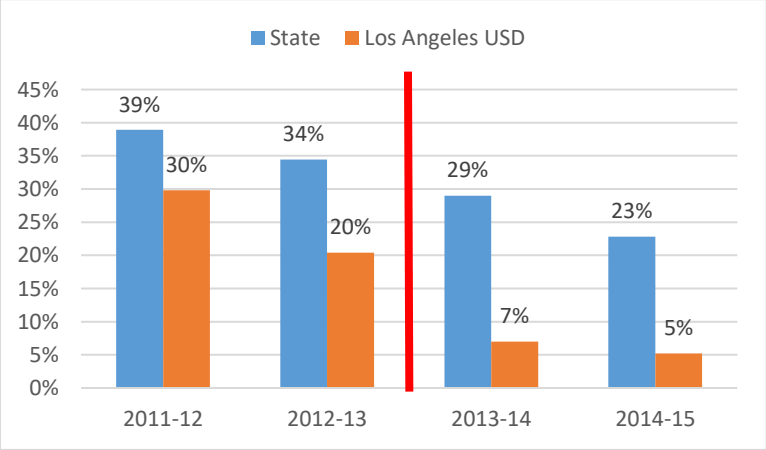
The LAUSD suspension ban provides an opportunity to analyze something closer to an ideal experiment. The data below suggest that many students at LAUSD schools that used to give suspensions based on defiance stopped receiving suspensions after 2013. In other words, these students would have received suspensions had the school board not passed the suspension ban. This mandated decrease in suspensions was not caused by the problematic factors mentioned above, allowing us to avoid concerns of that type of omitted variable bias.

In LAUSD, the same students in the same schools suddenly stopped receiving suspensions for their behavior. Those LAUSD schools compose the treatment group. At the same time, students in schools outside of LAUSD were able to receive suspensions based on defiance; those schools are the control group. While not perfect, this comes closer to the experimental ideal than prior research. This next section uses this research design to estimate the impact of LAUSD's suspension ban on academic growth.

The Impact of LAUSD's Suspension Ban

The data suggests that the LAUSD suspension ban had an enormous impact on suspensions. Graphs 1 and 2A show the number of suspensions based on defiance has dropped steadily state-wide in absolute and relative terms, while the absolute number given for other reasons has decreased slightly.[12] Graphs 1 and 2B show a different story for LAUSD: a precipitous drop in 2013-14 that leads to an extremely small number of suspensions based on defiance. This demonstrates the district-wide response to an explicit school board policy. In May 2013, the LAUSD school board voted to ban suspensions based on defiance during the upcoming school year (Watanabe 2013b). The overall compliance rate was high; from 2,814 suspension based on defiance in 2012-13 to just 305 in 2014-15. In general, schools did not compensate by suspending students for other reasons; the number of those suspensions decreased over the same period at a rate faster than the statewide average.

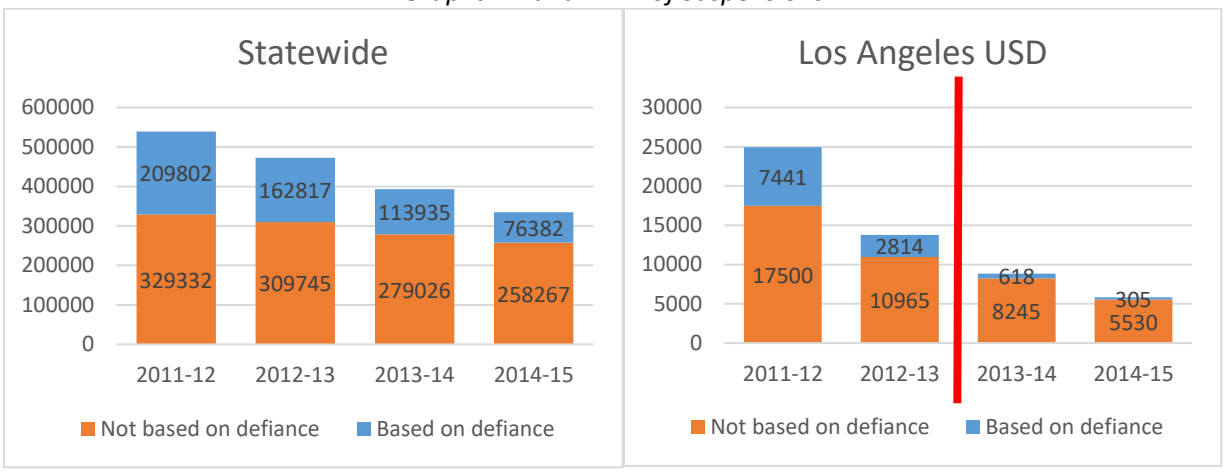
Graph 1: % of Suspensions Based on Defiance



[11] Prior research was cross-sectional, comparing suspension rates to academic achievement (Rausch & Skiba 2004). Even more recent papers on the subject sometimes use correlational evidence to support the claim the suspensions harm overall academic growth (Losen et al. 2015).

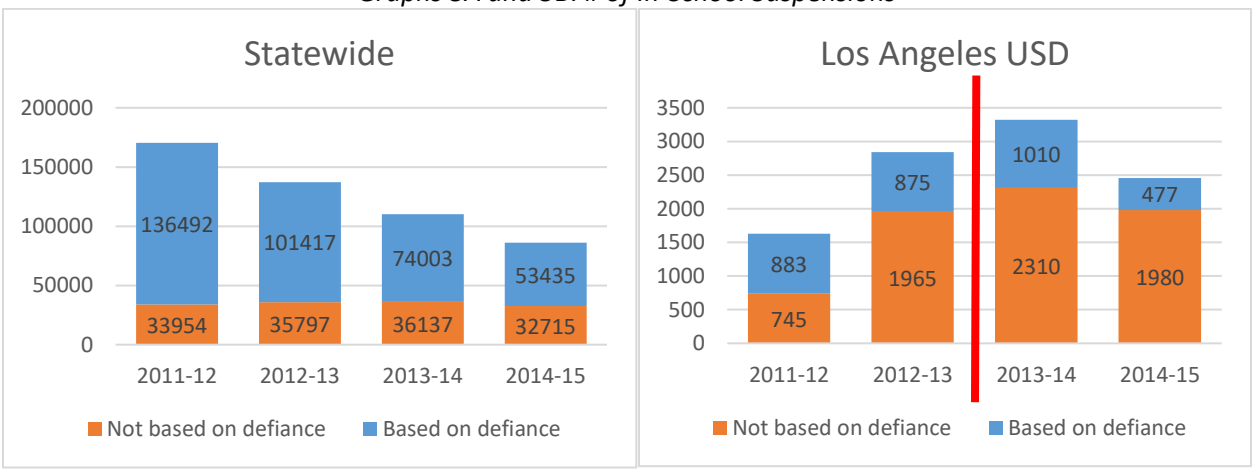
[12] Annually, the state has reported suspension data with the defiance category beginning with the 2011-12 school year.

Graphs 2A and 2B: # of Suspensions



Some LAUSD schools appeared to compensate temporarily by issuing more in-school suspensions. Graph 3B reveals an initial uptick in 2013-14 in-school suspensions, both based on defiance and not. However, by 2014-15 the number of in-school suspensions based on defiance dropped to approximately half the level of 2012-13, and the number given for other reasons returned to the 2012-13 level. If we ignore 2013-14, the change in LAUSD from 2012-13 to 2014-15 mirrors that of the state.

Graphs 3A and 3B: # of In-School Suspensions



Measurement of suspensions becomes more complicated when we shift to the school level. We use two types of independent variables depending on which hypothesis we are testing. For hypotheses one and two, we are comparing schools that experienced LAUSD's suspension ban to California schools that did not. The independent variable is therefore binary: did the California school experience the policy change (i.e. because it is located in LAUSD) or not. The third hypothesis compares LAUSD schools by the extent to which they were impacted by the suspension ban. This forces us to grapple with a variety of data and theoretical limitations described in detail below and in Appendix A. Those constraints lead us to use a measure of the intent to treat: a categorical variable reflecting the number of suspensions based on defiance a school gave in 2013.

Table 1 shows basic information for the groups of schools used to test each hypothesis. Hypothesis one is the broadest, comparing all schools in LAUSD to all California schools outside of LAUSD. The only requirement for inclusion is that schools must have the relevant growth data: sixth grade data in 2011 and 2013, and eighth grade data in 2013 and 2015. Hypothesis two separates these schools based on whether they gave any suspensions based on defiance in 2013, before the LAUSD suspension ban took effect. Hypothesis three looks only at schools within LAUSD. Here we add one exclusion rule; we remove 19 schools that served high school students. We do this because we want our schoolwide suspension data to reflect middle school, the grades where we are measuring academic growth.

Table 1: Information about Schools in Analyses

	Schools	# of Schools	Avg. # 2013 Suspensions	Avg. # 2015 Suspensions	% Low-Income	% English Learner	% Students w/ Disabilities
H1	LAUSD	113	6	1	81%	20%	12%
	Non-LAUSD	1,068	19	9	56%	20%	9%
H2a	LAUSD	79	8	1	83%	21%	12%
	Non-LAUSD	867	23	11	58%	20%	10%
H2b	LAUSD	34	0	0	79%	18%	11%
	Non-LAUSD	201	0	2	47%	19%	8%
H3	LAUSD: 0	29	0	0	77%	18%	10%
	LAUSD: 1-10	49	2	1	80%	19%	12%
	LAUSD: 11+	16	27	3	89%	20%	14%

Note: Demographic variables reflect all tested students in 2015

Available school-level suspension data has several important limitations.[13] While we would ideally like to have data by at the grade level to match academic growth, the state only reports data at the school level. Additionally, redaction prevents us from calculating exact suspension rates for a subset of schools. The state provides suspension data files containing data for each subgroup in each school, but in order to protect student confidentiality it does not report numbers between one and ten. Using the 2013 (i.e. pre-treatment) number of suspensions based on defiance produces the following categorization:

1. None (29 schools)
2. Between 1 and 10 (49 schools)
3. At least 11 (16 schools)

This can be viewed as a measure of the intent to treat. If schools in LAUSD had complied perfectly with the suspension ban, then this categorization would perfectly reflect the absolute change in suspensions. Officially, compliance was high enough to make this almost true. Of the 94 schools in our sample, 67 had zero suspensions based on defiance in 2015, another 20 had one redacted subgroup (i.e. on average 1.5 suspensions each), six schools had between two and twelve suspensions, and one school had thirty-five.

However, there are reasons to believe that the data after 2013 is not accurate. In the first year of implementation, parents claimed that their children were sent home without being officially suspended (Watanabe 2014). School staff became much more likely to call the police in order to deal with defiant students, forcing officers to remind school staff that "willful defiance is not a crime" (Watanabe 2015). There are also a variety of ways to remove students from class but keep them in

[13] For ease of reading, "suspensions" from this point on refer to out-of-school suspensions based on defiance, unless explicitly qualified otherwise.

school, and administrators may not report all those instances as in-school suspensions. The street-level bureaucracy literature is founded upon the notion that the actual implementation of top-down policies is determined largely by the decisions of front line, or street-level, workers (Lipsky 1980). In this case, the street-level workers are school administrators and even teachers who decide how to deal with defiant students. School staff see themselves as accountable to their students, parents, and fellow staff as well as their school board (Hupe and Hill 2007). Staff therefore might report perfect compliance while finding ways to remove defiant students from school.

Importantly, reported compliance is a confounding factor because it impacts the change in suspension rate and could also be related to the change in academic growth. Twelve LAUSD schools reported giving more suspensions based on defiance after the suspension ban. These schools had relatively low growth rates in 2013 and grew even less in 2015 (see Appendix A for more details). Instead of the increase in suspensions causing the drop in academic growth, it seems more likely that these schools suffered from other issues that led to both the increase in suspensions and the drop in academic growth.

Data limitations place a variety of constraints on our ability to measure change in academic growth, the dependent variable. The main issue is that California switched to a new assessment regime immediately after LAUSD enacted its policy. Fortunately, the residual gain model of academic growth does not require that pre- and post-tests be on the same scale. The basic residual gain model is a bivariate regression where the post-test score is the dependent variable and the pre-test score is the independent variable. The regression residuals estimate the extent to which students performed lower or higher than expected given their starting score. This is one form of value-added modeling, which is the best available metric of academic growth in the education literature (e.g. Kogan et al. 2016).[14]

A related challenge is that the assessment regime prior to 2015 gave end-of-course assessments for some high school English classes as well as all Math classes starting with Algebra, which some students took in middle school. We can only calculate the residual gain model of academic growth when all students took the same assessment in the same grade. This prevents measuring growth for high schools at all, or for middle school Math. Elementary would be possible as well, but suspensions occur rarely among students below sixth grade. Therefore, all measures of academic growth are based on middle school English assessments.

Two other issues concern state reporting of academic data. First, California did not publish any statewide assessment results in English or Math for the 2013-14 school year. This means that the pre-test is sixth graders in spring 2013 and the post-test is eight graders two years later in spring 2015. In order to calculate analogous growth prior to the policy intervention, we similarly span two years: sixth graders in spring 2011 to eighth graders in spring 2013. The second issue is that the state only makes data available for entire grades at each school, not individual students. Ideally, pre-test sixth grade scores and eighth grade post-test scores would include the exact same sets of students. Reality is more complicated; some students leave and enter the sixth grade cohort because they change schools or are held back a grade. If the weighted average score of the students who leave the cohort matches the weighted average of the students who enter, then there is no bias. Bias occurs when students who enter and students who leave have different weighted average scores. For example, imagine a school where equal numbers of students leave and enter the sixth grade cohort. Students who leave are

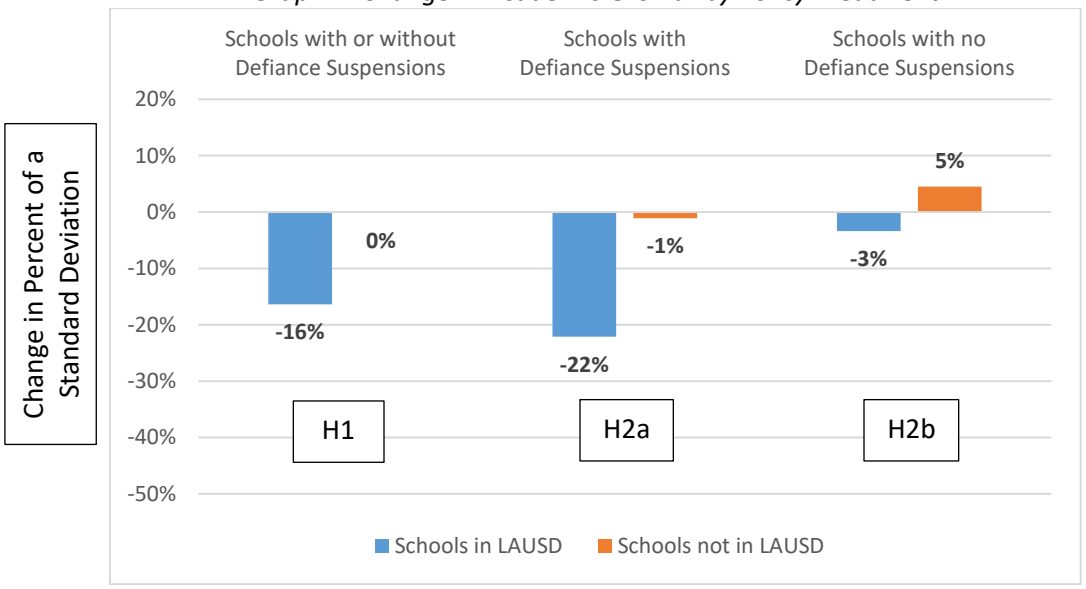
[14] The value-add literature has primarily focused on its controversial use in teacher evaluation. Using value-add estimates for entire grades avoids some concerns, such as sorting difficult-to-teach students into particular teachers' classes (Rothstein 2009). However, even grade level value-add measures can experience significant variation across time (Goldhaber and Hansen 2008) and can be sensitive to the assessment used (Lockwood et al. 2007). On the positive side, value-add measures are highly correlated with principal evaluations of teachers (Kimball et al. 2004; Jacob and Lefgren 2008).

relatively high-achieving while students who enter are relatively low-achieving. In this case, our growth measure would underestimate the amount of academic improvement that really occurred. While California does not report what percent of students remain in a sixth grade cohort from year to year, it reported what percent of students are continuously enrolled at a school from early October to spring testing as recently as 2013. The median score was always near 95%, reflecting the fact that most schools have relatively stable student populations.

The primary analysis is a difference-in-difference research design. The first difference is temporal: outcomes before the suspension ban compared to outcomes after the suspension ban. In this case, the dependent variable is change in middle school academic growth. The second difference is the comparison between various groups of schools, depending on the particular hypothesis under consideration. While H1 and H2 compare LAUSD schools to schools in other parts of California, H3 compares groups of schools within LAUSD.

To test H1, we compare all LAUSD middle schools to all other middle schools statewide. Graph 4 shows that LAUSD schools experienced a 16% standard deviation decrease while other California middle schools remained essentially static. However, this comparison includes schools that were not impacted by the suspension ban because they did not utilize defiance suspensions. The middle columns support H2a by revealing a bigger gap between LAUSD (-22%) and the rest of the state (-1%). The right-hand columns show very slight changes in academic growth for schools – in and out of LAUSD – that did not give defiance suspensions. This supports H2b, providing confidence that the gap we see between LAUSD and the rest of the state is driven by this suspension ban as opposed to other possible factors. Bivariate regressions reveal that the difference between LAUSD and non-LAUSD schools is almost statistically significant for H1 (p=0.13) and is significant for H2a (p=0.09), supporting the strand of the literature arguing that a sudden drop in suspensions harms academic growth.

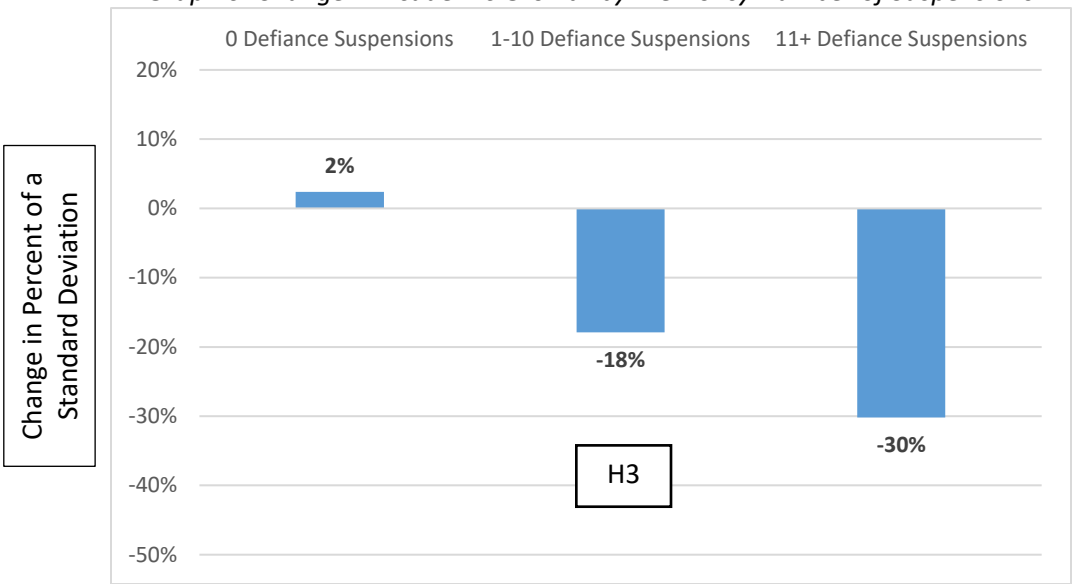
Graph 4: Change in Academic Growth by Policy Treatment



Looking within LAUSD reveals a linear relationship between the intent to treat and academic growth. Schools with no suspensions in 2013 had almost no change in growth. Schools with one to ten suspensions experienced an 18% standard deviation drop, while schools with at least eleven suspensions in 2013 experienced a 30% decrease in growth. This pattern perfectly fits the expectations of hypothesis three. The small numbers of schools in each category prevent any of the differences from being statistically significant. Combined with hypothesis 2, however, these findings somewhat alleviate

concerns of omitted variable bias: that some other factor about LAUSD caused the drop in academic growth. In order for an omitted variable to explain the story, it would need to both be related to changes in academic growth and the number of suspensions based on defiance schools issued in 2013.

Graph 6: Change in Academic Growth by Pre-Policy Number of Suspensions



Limitations

This paper falls short of the ideal of an experimental design leading to statistically significant results. However, additional analyses alleviate the main threats to inference. One major concern is measurement error in the outcome, change in academic growth. Each measurement of growth covers two years and reflects one cohort of students. Student mobility causes grade level cohorts on which growth is calculated change by unknown amounts over time at different schools. Academic growth from 2013 to 2015 spans two different assessment systems and standards, leading to concerns that we may be measuring variations in teacher preparation for this change more than changes in student learning.

Two analyses address this issue. One involves converting the academic performance of each school in the state into a percentile from 1 to 100 in both 2013 and 2015. This is the closest we can get to a consistent measure of achievement given the change in assessment systems during this time. We then calculate the change in academic percentile from 2013 to 2015 for each school. Replicating graphs 5 and 6 with this academic measure reveals the same trends reported in this paper. The other analysis concerns Oakland Unified School District and is described below.

A second major concern is omitted variable bias. As explained above, the difference-in-difference methodology partly addresses this concern. By looking at the change over time, we hold all time-invariant factors constant. The top-down nature of the suspension ban provides additional assurances; the intended drop in suspensions is unrelated to sudden changes in students' lives that would also cause a drop in academic growth. Additionally, we run seven regressions to see if school-level traits can explain the strongest finding, hypothesis 2A. We control for prior academic growth, racial demographics, non-racial demographics, average parent education, and then all those factors combined. The variable for being in LAUSD always has a negative coefficient (i.e. LAUSD schools experienced decreased academic growth), and is similar in size and significance across most models. The exceptions are models 3 and 6, which include non-racial demographics: the percent of students who are low-income, English Learners, and students with disabilities. The rows for H2A in Table 1 show that

LAUSD has relatively high percentages of students in all three of these non-racial demographic categories. It is possible that these factors caused approximately half of LAUSD's relative decrease in academic growth. However, it is also possible that the correlation between these factors and being in LAUSD is coincidental.

Table 2: OLS Regressions as Robustness Tests for Hypothesis 2A

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
In LAUSD	-0.227* [0.132]	-0.256** [0.106]	-0.135 [0.139]	-0.252* [0.138]	-0.242* [0.147]	-0.103 [0.112]	-0.204* [0.109]
Additional Independent Variables	None	Prior Growth (2011 to 2013)	Non-Racial Demo-graphics (% with Disabilities, % Low-Income, % English Learner)	Racial Demo-graphics (% African American, % Asian, % Latino, % Other)	Average Parent Education	All	All but Non-Racial Demo-graphics
Observations	946	946	946	946	942	942	942
Adjusted R^2	0.002	0.360	0.007	0.003	0.001	0.384	0.372

Notes: The outcome variable is change in growth from 2011-2013 to 2013-2015. Observations include all schools with change in growth data and no suspensions based on defiance in 2013: 79 in LAUSD and 867 not in LAUSD. Demographic data reflects all test-takers at each school in spring of 2015. Average parent education is based on an annual school-administered survey; the mean response rate in the sample is 90%, but 11% of schools have less than a 75% response rate and 4 schools have no responses at all.

The final concern is selection bias: LAUSD schools were not randomly selected to impose a suspension ban. We are assuming that LAUSD schools are comparable to all other California schools. However, LAUSD is different from the rest of the California in many ways, and perhaps some of those unique characteristics allowed the suspension ban to harm academic growth. If another school district implemented a suspension ban, would academic growth be similarly hurt?

The experience of Oakland Unified School District suggests the answer is yes. Oakland banned suspensions in 2015, after the first year of testing had occurring in California's new assessment system. Oakland schools grew 44% of a standard deviation less than the rest of the state by spring of 2016 (-5.13 scale score points / 11.74 standard deviations). This difference is statistically significant (p=0.000). Even more troubling, it occurred despite the fact that Oakland spent over two million dollars during 2015-16 training teachers in behavior management! We see the same pattern in a much simpler measure of academic growth, change in the percent of students at or above the state standard for their grade level. While the state had 5% more students reaching the standard in spring 2016 in English and 4% more in Math, Oakland saw only 2% growth in both subjects. We cannot conduct a full difference-in-difference analysis for Oakland because we lack data to measure the change in academic growth. It could be that Oakland would have had a relatively low growth rate even without the suspension ban. However, the analyses of both LAUSD and Oakland complement one another and make it increasingly likely that suspension bans have a significant and negative causal impact on academic growth.

Implications and Conclusion

Despite lingering uncertainty, the evidence in this paper is important for three primary reasons. First, it is closer to the experimental ideal than the current literature. It is a significant improvement on

work that correlates suspension rates and achievement (Losen et al. 2015). It does not rely on matched comparisons between students who get suspended and students who do not, as does Arcia (2006). Nor does it rely on the assumption that natural fluctuations in suspension rates cause subsequent changes in academic performance, as does Perry and Morris (2014). This paper relies on the LAUSD board decision to ban suspensions based on defiance to impose a sudden, intended shift in suspension rates among a particular subset of California schools. The standard of quality for social science research is not perfection, but providing new information with the best available method (Gerring 2012).

This has implications for the theoretical disagreement within the literature. None of the analyses support the more prominent strand of the literature claiming that sudden reductions in suspensions will cause academic growth. In contrast, all of the analyses provide suggestive evidence in favor of the strand of the literature claiming that sudden reductions in suspensions will reduce academic growth. Perry and Morris' (2014) finding that reducing suspensions would improve academic growth for non-suspended students appears to be driven by omitted variable bias. Unmeasured factors that caused drops in suspension rates in their sample also caused increases in academic growth. Arcia's (2006) finding that suspensions academically harm suspended students may or may not suffer from similar bias. Our analysis of LAUSD lacks the data granularity to differentiate between types of students. It could be that the suspension ban in fact improved academic growth for students who would have been suspended and simultaneously harmed academic growth for all other students. Future research could look for varied impacts across groups such as suspended students, English Learners, or students with disabilities.

Student-level data also could enable a more accurate estimate of the LAUSD suspension ban's impact. Ideally, the cohorts used to measure two-year growth would only include students continuously enrolled in schools during that time. Additionally, exact counts of the numbers of students – ideally in those cohorts, not the entire school – who received out-of-school suspensions for defiance would allow the calculation of 2013 suspension rates for the same groups of students for whom we measure academic growth. Student-level data would enable use of a multilevel model and make it more likely that estimates are statistically significant.

This should give pause to the variety of actors promoting suspension bans. At best, the lack of statistical significance in two robustness tests of Hypothesis 2A means that LAUSD's suspension ban may have had no causal impact. There is no evidence allowing us to say that the ban improved academic growth, and the totality of evidence heavily favors the conclusion that the ban harmed academic growth. Those of us troubled by the negative consequences of suspensions need to find other ways to reduce their use. Bans had the benefit of being simple to implement, had virtually no immediate financial cost, and produced immediate results. More gradual approaches to suspension reduction will be relatively complicated, cost time and money (e.g. for staff training), and may take several years to see results. Nevertheless, the available evidence makes the gradual approach much preferable.

The second reason this evidence is important is that the impact on academic growth might be substantial. A full school year of academic growth is roughly equivalent to one standard deviation; the analyses suggest the suspension ban had an impact of approximately 20% of a standard deviation on schools that gave those suspensions. If true, this would be an enormous impact. This would be a larger impact than shifting from a bottom quartile teacher to a top quartile teacher (2012 Gathering Feedback for Teaching).[15] The same suspension ban would have been imposed statewide in 2012 if California Governor Gerry Brown had not vetoed it (Watanabe 2013a). Although there is some uncertainty around this 20% estimate, the fact that it could have such a large negative impact should make people very

[15] The Measures of Effective Teaching (MET) project rigorously evaluated the performance of nearly 3,000 teachers. It found that teacher performance is most accurately predicted by a combination of classroom observations, student surveys, and value added measures from standardized tests.

hesitant to encourage the adoption of suspension bans. Even significant amounts of money simultaneously spent on teacher training in behavior management, as in Oakland, does not appear to prevent the academic harm.

The final reason this evidence is important is that school discipline is an active area of policy debate and experimentation. Betsy Devos, President Donald Trump's pick for Education Secretary, has held meetings to discuss whether to change the Department of Education's stance on school discipline. Conservative groups are advocating that the federal government should resume its limited role of investigating particular complaints of unfair practices. This would be quite a departure from the Obama-era policy of pressuring school districts to reduce suspensions if rates between racial groups were not equitable. This paper indirectly weighs in on this debate: if the federal government wants school districts to reduce suspensions, districts should also evaluate their changes to ensure that their efforts do not have the unintended consequence of harming academic growth. Additionally, districts would be wise to choose gradual approaches rather than initiatives that would impose sudden reductions in suspension rates.

What should districts do that have already implemented suspension bans? LAUSD School board member Richard Vladovic voted to ban suspensions based on defiance "as an experiment, saying he would be 'the first to stop it' if it proved disruptive to learning" (Watanabe 2013c). This brings up an important point: just because a suspension ban decreases academic growth does not mean that reversing a ban will cause an increase. Unfortunately, it is probably easier to harm academic growth than to help it. Also, it is possible that training in restorative justice or other behavior management practices might cause an increase in academic growth without having to resume giving suspensions for defiance. Districts would be wise to try a variety of more gradual options, evaluate their impacts, and then make an informed choice as to how to proceed.

This highlights the broader need for policy evaluation at the school district level. School board members and district leaders rarely know with certainty how a policy is likely to impact academic growth, and they often do not conduct evaluations of the policies they implement. An array of factors contribute to this problem, ranging from the relatively weak research base in education to the political cost of having to admit that past decisions led to bad outcomes. We are far from making evaluation a routine component of district policy decisions. The aim of this paper is to nudge us in that direction. Future researchers may find ways to reduce suspensions that also increase academic growth. Even more important would be if future researchers regularly help to inform school district leaders as they make important decisions concerning the education of our children.

References

(2007). History: Twenty-five years of progress in educating children with disabilities through IDEA. U.S. Office of Special Education programs, accessed https://www2.ed.gov/policy/speced/leg/idea/history.pdf.

(2012). Gathering Feedback for Teaching. *Bill and Melinda Gates Foundation*, MET Project research paper, accessed file:///C:/Users/dzarecki/Downloads/MET_Gathering_Feedback_Research_Paper1.pdf.

(2014). U.S. Departments of Education and Justice Release School Discipline Guidance Package to Enhance School Climate and Improve School Discipline Policies/Practices. Jan 8, accessed https://www.ed.gov/news/press-releases/us-departments-education-and-justice-release-school-discipline-guidance-package-.

Arcia, Emily. (2006). Achievement and Enrollment Status of Suspended Students: Outcomes in a Large, Multicultural School District. Education and Urban Society 38(3):359–69.

Berwick, Carly. (2016). "Ban school suspensions!" *The Week*, Aug 9.

Cerrone, K. M. (1999). The Gun-Free Schools Act of 1994: Zero Tolerance Takes Aim at Procedural Due Process. *Pace Law Review*, 20, 131.

Charis, Kimberly, Losen, D. J. (2017). Fifth Indicator: School Climate and Student Discipline. *National Association of State Boards of Education*, Policy Update 24(4).

Christle, C. A., Jolivette, K., & Nelson, C. M. (2005). Breaking the school to prison pipeline: Identifying school risk and protective factors for youth delinquency. *Exceptionality*, *13*(2), 69-88.

Cline-Thomas and David Chang. (2017). Should the Philly School District End Suspensions for Elementary School Students? *NBC10*, May 3.

Crone, D. A., Hawken, L. S., & Horner, R. H. (2010). *Responding to problem behavior in schools: The behavior education program*. Guilford Press.

Frey, Susan. (2015). Oakland ends suspensions for willful defiance, funds restorative justice. *EdSource*, May 14.

Gerring, John. (2012). *Social Science Methodology: A Unified Framework*, 2nd edition. Cambridge University Press.

Goldhaber, D. D., & Hansen, M. (2008). *Is it Just a Bad Class?: Assessing the Stability of Measured Teacher Performance*. Center on Reinventing Public Education.

Hupe, P., & Hill, M. (2007). Street-Level bureaucracy and public accountability. *Public Administration*, 85(2), 279-299.

Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of labor Economics*, *26*(1), 101-136.

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, *79*(4), 54-78.

Kogan, V., Lavertu, S., & Peskowitz, Z. (2016). Performance federalism and local democracy: Theory and evidence from school tax referenda. *American Journal of Political Science*, *60*(2), 418-435.

Liaupsin, C. J., Umbreit, J., Ferro, J. B., Urso, A., & Upreti, G. (2006). Improving academic engagement through systematic, function-based intervention. *Education and Treatment of Children*, 573-591.

Lindstrom, Natasha. (2017). Pittsburgh Public Schools examines suspension practices. *TribLive*, June 22.

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

Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, V. N., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, *44*(1), 47-67.

Losen, Daniel. (2011). Discipline Policies, Successful Schools, and Racial Justice. UCLA: The Civil Rights Project / Proyecto Derechos Civiles.

Losen, Daniel J.; Hodson, Cheri L.; Keith II, Michael A.; Morrison, Katrina; & Belway, Shakti. (2015). Are We Closing the School Discipline Gap?. *K-12 Racial Disparities in School Discipline*. UCLA: The Civil Rights Project / Proyecto Derechos Civiles.

Losen, Daniel J.; Keith II, Michael A.; Hodson, Cheri L.; Martinez, Tia E.; & Belway, Shakti. (2015). Closing the School Discipline Gap in California: Signs of Progress. UCLA: The Civil Rights Project / Proyecto Derechos Civiles.

Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg, A. B. (2005). Whole-school positive behaviour support: effects on student discipline problems and academic performance. *Educational Psychology*, *25*(2-3), 183-198.

McFarland, D. A. (2001). Student resistance: How the formal and informal organization of classrooms facilitate everyday forms of student defiance. *American Journal of Sociology*, *107*(3), 612-678.

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

Putnam, R. P., Horner, R. H., & Algozzine, R. (2006). Academic achievement and the implementation of school-wide behavior support. *Positive Behavioral Interventions and Supports Newsletter*, *3*(1), 1-6.

Rausch, M. K., & Skiba, R. (2004). Disproportionality in School Discipline among Minority Students in Indiana: Description and Analysis. Children Left Behind Policy Briefs. Supplementary Analysis 2-A. *Center for Evaluation and Education Policy, Indiana University*.

Reid-Cleveland, Keith. 2017. Texas Governor Greg Abbott signs bill to ban out-of-school suspensions for young students. *Black Youth Project*, June 13.

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education*, *4*(4), 537-571.

Rumberger, R. W., & Losen, D. J. (2017). The Hidden Costs of California's Harsh School Discipline: And the Localized Economic Benefits from Suspending Fewer High School Students. *Civil Rights Project-Proyecto Derechos Civiles*.

Shipan, C. R., & Volden, C. (2008). The mechanisms of policy diffusion. *American Journal of Political Science*, 52(4), 840-857.

Shipan, C. R., & Volden, C. (2012). Policy diffusion: Seven lessons for scholars and practitioners. *Public Administration Review*, 72(6), 788-796.

Skiba, R. J., & Knesting, K. (2001). Zero tolerance, zero evidence: An analysis of school disciplinary practice. *New Directions for Student Leadership*, *2001*(92), 17-43.

Skiba, R. J., & Sprague, J. (2008). Without Suspensions. *Educational Leadership*.

Volden, C., Ting, M. M., & Carpenter, D. P. (2008). A formal model of learning and policy diffusion. *American Political Science Review*, 102(3), 319-332.

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

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

Watanabe, Teresa. (2013a) LAUSD board could ban suspensions for 'willful defiance.' *Los Angeles Times*, May 12.

Watanabe, Teresa. (2013b) L.A. Unified bans suspension for 'willful defiance.' *Los Angeles Times*, May 14.

Watanabe, Teresa. (2013c) L.A. schools will no longer suspend a student for being defiant. *Los Angeles Times*, May 15.

Watanabe, Teresa. (2014) L.A. Unified suspension rates fall but some question figures' accuracy. *Los Angeles Times*, May 31.
Watanabe, Teresa. (2015) Why some LAUSD teachers are balking at a new approach to discipline problems. *Los Angeles Times*, Nov 07.

Appendix A: Analysis using Estimates for Change in Suspension Rate

The 2014-15 data for LAUSD is either zero or redacted, with just four exceptions.[16] Supplemental manual data collection allows us to obtain exact suspension rates for all but 83 LAUSD schools.[17] This reveals that the 83 remaining schools have an average of 1.5 suspensions each.[18] This includes 20 schools in the analyses for hypothesis three, so we run robustness tests that exclude these schools.

There is a bit more uncertainty about the number of suspensions in 2012-13. Data files contain exact information for schools that report either zero or more than ten suspensions based on defiance for all subgroups. The more difficult cases are schools where some or all of their suspension data is redacted: 459 redacted data points account for 1166 suspensions. Manual data collection allows us to know the exact number of suspensions – totaling 674 – for another 107 LAUSD schools. The 204 schools with one redacted subgroup therefore share the remaining 492 suspensions (1166 minus 674) – an average of 2.4 suspensions each. This includes 30 schools in the analyses for hypothesis three, so we run robustness tests excluding those schools. The larger number of schools with unknown suspension counts (204 vs. 83) combined with those schools' higher average number of suspensions (2.4 vs. 1.5) results in a little more uncertainty surrounding the estimates for 2012-13 than for 2014-15.

The uncertainty motivates us to measure "change in suspension rate" categorically rather than continuously. Replacing unknown redacted data with the average number of suspensions per subgroup would allow us to estimate a continuous variable, but this measure would contain errors for schools with unknown redacted data. One approach uses those estimated suspension rates to place LAUSD middle schools into one of one of four categories:

1. Increased suspension rate (12 schools)
2. Maintained suspension rate (25 schools)
3. Decreased suspension rate <=1% (44 schools)
4. Decreased suspension rate >1% (13 schools)

For 13 LAUSD schools in our sample, we are uncertain whether they experienced an increase or decrease in suspensions because they had one redacted subgroup in one year and between 1 and 10 suspensions in the other year. We run robustness tests without these schools and find very similar results.

There is significant overlap between this categorization and the one (in the paper) based on the number of suspensions given in 2013. All but 1 of the 13 schools that decreased suspension rates more than 1% had at least eleven suspensions in 2013. All but 3 of the 44 schools that decreased 1% or less had fewer than eleven suspensions in 2013. All 25 schools that had no change in suspension rate had no suspensions in 2013. The main difference is that the first categorization identifies and groups the 12 schools that experienced increases in suspension rates; 6 had one redacted subgroup (i.e. probably one or two suspensions), five had between two and twelve reported suspensions, and one had thirty-five suspensions in 2013.

---

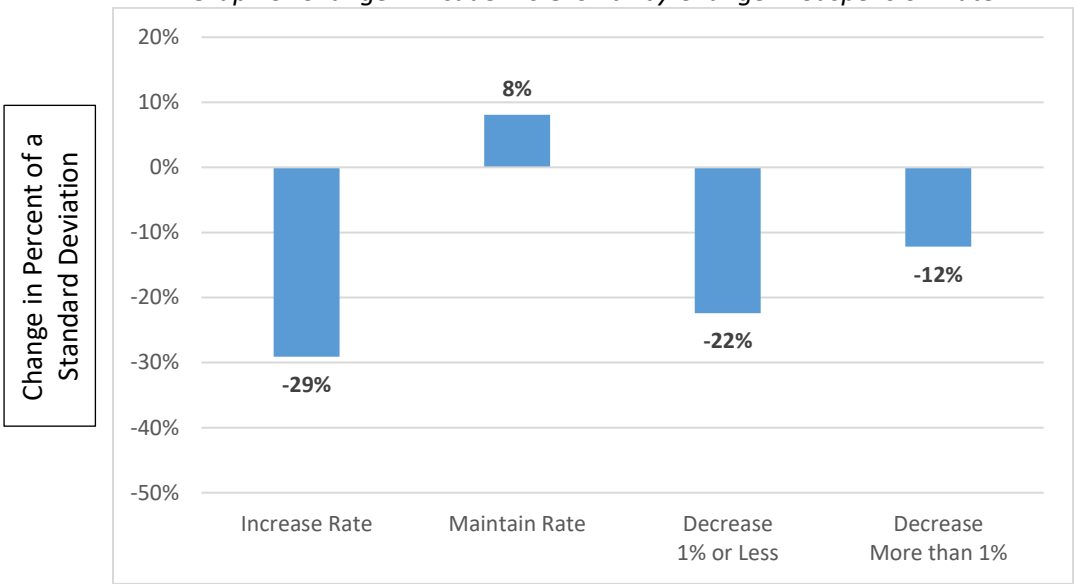
[16] Four LAUSD schools – only one included in the analysis, Edwin Markham Middle – had at least one subgroup with over ten suspensions based on defiance in 2014-15, for a combined total of 75 suspensions. Subtracting these 75 suspensions from the 305 we know occurred district-wide leaves 230 possible suspensions.

[17] Schoolwide suspension counts are only available from individual school pages a state website (not in a dataset format), and it is redacted whenever reporting it would reveal a redacted subgroup. Looking up each of the twenty-one schools that had two or three redacted subgroups allows us to explain 106 of the remaining 230 suspensions that occurred district-wide (see above footnote).

[18] The remaining 124 suspensions must be spread across eighty-three schools that have only one redacted subgroup each. While it is possible that a very small number of these schools have up to ten suspensions, mathematical constraints assure us that the vast majority of these school have only one or two.

To test hypothesis three, we conduct analyses of schools within LAUSD. That strand of the literature predicts that the decrease in academic growth should be correlated to the decrease in suspension rate. The first categorization – based on changes in suspension rates – show a non-linear relationship between suspensions and growth. Schools that maintained a suspension rate saw an increase in academic growth. Schools with a decrease of 1% or less had a 22% drop in growth. Schools with at least a 1% decrease in suspension rate experienced a decrease approximately half the size: 10% of a standard deviation. The pattern for these three right-hand columns does not perfectly fit H3, but it is close.

Graph 5: Change in Academic Growth by Change in Suspension Rate



The largest decrease – 29% of a standard deviation – was among the dozen schools that had an increase in suspension rate. Four of these schools had 0 defiance suspensions in 2013, seven of them had between 1 and 10, and one school had 40. Although this finding is not driven by a particular outlying school, evidence does not suggest that giving more suspensions in 2015 caused this academic decline. Most of these schools had extremely small changes in suspension rate, and the five schools with changes above 0.2% experienced almost no change in academic growth (-5%). Additionally, it is noteworthy that these twelve schools had the lowest 2013 growth rate of all the groups of LAUSD schools we analyze, as shown in Table 2. This suggests a different explanation: that at least some of these schools had other problems which led to both their non-compliance with the suspension ban and a further decrease in their academic growth rate.

Table 2: Academic Growth Rates by Change in Suspension Rate

	Academic Growth 2011 to 2013	**Academic Growth 2013 to 2015**
Increased Suspension Rate	-21%	-51%
Maintained Suspension Rate	5%	13%
Decreased Suspension Rate 1% or Less	-10%	-32%
Decreased Suspension Rate More than 1%	-9%	-21%