Education Reform Faculty and Graduate Students Publications

Education Reform

3-31-2017

Do School Discipline Policies Treat Students Fairly? A Second Look at School Discipline Rate Disparities

Kaitlin Anderson
Michigan State University, ande2018@msu.edu

Gary W. Ritter
University of Arkansas, Fayetteville

Citation

Anderson, K., & Ritter, G. W. (2017). Do School Discipline Policies Treat Students Fairly? A Second Look at School Discipline Rate Disparities. *Education Reform Faculty and Graduate Students Publications.* Retrieved from https://scholarworks.uark.edu/edrepub/39

WORKING PAPER SERIES

Do School Discipline Policies Treat Students Fairly?

A Second Look at School Discipline Rate Disparities

Kaitlin P. Anderson

Dr. Gary W. Ritter

March 31, 2017

EDRE Working Paper 2015-11

Do School Discipline Policies Treat Students Fairly?

Evidence from Arkansas

Kaitlin P. Anderson
University of Arkansas
211 Graduate Education Building
Fayetteville, AR 72701
Email: kaitlina@uark.edu
Phone: (479) 575-3172
Fax: (479) 575-3196

Gary W. Ritter, PhD
University of Arkansas
207 Graduate Education Building
Fayetteville, AR 72701
Email: garyr@uark.edu)
Phone: (479) 575-4971
Fax: (479) 575-3196

Abstract

Much work has documented that African-American students are more likely to receive expulsions and suspensions than their white peers. These disparities are troubling, but researchers and policymakers need more information to fully understand this issue. We use three years of student-level discipline data for an entire state to assess whether non-white students are receiving different disciplinary consequences from their white peers in the same schools, for similar infractions and with similar behavioral history. We find that Black students received more severe (longer) punishments than their White peers in the state for the same types of infractions. These differences are due primarily to school-level differences, but even within the same schools, Black students receive slightly longer punishments than their White peers.

Since the early 1990s, many schools have adopted zero tolerance policies in response to fears of violence. This approach removes students from school for violations ranging from serious offenses like violent behavior to less serious offenses such as dress code violations or truancy (Losen & Skiba, 2010; Skiba, 2014; Skiba & Peterson, 1999;). While it may be necessary, in extreme cases, to remove a student from campus, many fear this movement has gone too far. Zero tolerance policies and exclusionary discipline such as expulsions and suspensions are associated with lower academic achievement (Beck & Muschkin, 2012; Raffaele-Mendez, 2003; Raffaele-Mendez, Knoff, & Ferror, 2002; Skiba & Rausch, 2004), school dropout (American Academy of Pediatrics, 2013; American Psychological Association, 2008; Ekstrom, Goertz, Pollack, & Rock, 1986), and involvement in the juvenile justice system (American Academy of Pediatrics, 2013; Balfanz, Spiridakis, Neild, & Legters, 2003; Fabelo et al., 2011; Nicholson-Crotty, Birchmeier, & Valentine, 2009).

One particularly troubling by-product of the increased use of exclusionary discipline is the growing evidence that such disciplinary practices are employed disproportionately for students from marginalized groups. Numerous researchers have documented differences in suspension rates between White students and students of color (Anyon et al., 2014; Losen, Hodson, Keith, Morrison, & Belway, 2015; Losen & Skiba, 2010; Sartain et al., 2015; Skiba et al., 2014; Skiba, Michael, Nardo, & Peterson, 2002; Skiba et al., 2011; Welch & Payne, 2010). In this study, we contribute to this growing base of evidence by assessing the extent to which Black students in Arkansas, over the past several years, have received more severe consequences than White students – despite being cited for similar infractions. This analysis makes a unique contribution both by controlling for the specific infractions leading to the disciplinary consequences (relatively few studies in the existing literature connect infractions to consequences) and by using days of suspension as the consequence measure rather than simply the likelihood of being suspended. While it is certainly helpful to know if Black students are more likely –

all else equal – to receive exclusionary discipline, it is also important that we are aware of any disparities in the severity of the consequences given.

In the next section, we set the context for our study by presenting the evidence from the literature on racial disparities in student discipline in two categories of studies. First, we discuss national studies that have generally relied on school-level data and provided only an overview of the consequences levied on students of different races. Because these studies are unable to connect consequences with the associated infraction referral, many questions are left unanswered. We then consider a second set of studies that have investigated the student and school characteristics associated with racial disparities in discipline within particular states or districts.

National Overviews of Disciplinary Disproportionalities

Most recently, Losen et al. (2015) from the Civil Rights Project at UCLA, published a comprehensive report asking "Are We Closing the School Discipline Gap"? The authors focused on OSS rates in every U.S. school district during the 2011-12 school year. The data revealed an increase in suspensions over the past 40 years, as well as growth in the suspension rate gap between White students and students of color. In 1972-73, only 6% of Black students were suspended, as compared to 3% of White students (and 3% of Hispanic students). By 2011-12, 16% of Black students were suspended; this rate was more than twice as great as for Hispanic students (7%) and more than three times as great as for White students (5%). The authors also found significant variability in rates across districts and schools, indicating that local policies influence discipline outcomes.

Earlier, Losen and Skiba (2010) utilized data from the Civil Rights Data Collection (CRDC) to analyze suspension rates for students in more than 9,200 middle schools across the nation from the years 2002 to 2006. Losen and Skiba (2010) focused on middle school, claiming that the relatively low number of suspensions in elementary school may mask disparities in middle and secondary school.

This analysis revealed stark racial gaps in suspensions: while only 10% of White male students in middle school were suspended in 2006, 28% of Black male students were suspended in that same year.

Overall, these and other analyses confirm that there are systematic racial disparities in OSS, but they provide little information related to the factors (including student behavior) that drive these disparities. Moreover, students face numerous consequences other than OSS, which these national studies have not investigated.

Studies Examining the Drivers of Racial Discipline Gaps

Many researchers have gathered school-level and even student-level data to better understand the variables underlying the disparities in disciplinary outcomes (Anyon et al., 2014; Sartain et al., 2015; Skiba et al., 2002; Skiba et al., 2011; Skiba et al., 2014; Welch & Payne, 2010). Two high-profile studies described in this section concluded that differences in disciplinary practices and culture between schools fostered racial discipline gaps. (Sartain et al., 2015; Welch & Payne, 2010). That is, the schools Black students attended were the same schools that engaged in high levels of exclusionary discipline and thus, overall, disproportionate numbers of Black students faced school suspension.

Welch and Payne (2010) used school level data from a national survey of principals in middle schools and high schools to ask whether school leaders in buildings serving more Black students would be more likely to use punitive discipline. In this analysis, the authors sought to extend the "racial threat hypothesis" from criminal justice research into a school context and found that principals in schools with higher proportions of Black students were more likely to self-report that they employ punitive disciplinary styles. The weakness of this study, however, was its reliance on self-reports rather than actual disciplinary outcomes. Moreover, the data in this study were school-level and did not indicate whether Black students in particular are punished more frequently.

Researchers from the Consortium on Chicago School Research used actual student behavior data from roughly 85,000 high school students in Chicago in 2013-14 to further examine school level differences in suspensions (Sartain et al., 2015). The student-level data employed in the analysis did allow the researchers to look at within-school racial differences, but they concluded that the primary driver of the differences was the fact that Black students attended schools, on average, that issued more suspensions. Like the Welch and Payne (2010) study, this analysis is limited because the authors did not consider the actual infractions committed by students; nevertheless, these two studies provide evidence of the fact that principals across the country are more likely to favor punitive discipline in schools serving many Black students and that, in at least one large city, such schools are home to greater numbers of student suspensions.

While it is helpful that researchers have documented the different disciplinary cultures and practices prevalent in schools serving mostly Black students, the findings are not surprising to those studying school discipline. Over the past 15 years or so, a number of researchers have examined infraction-level data to further explore other factors that might help explain the differential rates of punishment faced by Black students. For example, Skiba et al. (2002) used student-level data on more than 11,000 students from 19 middle schools in one of the largest US school districts in 1994-95 to ask what factors drive discipline disproportionalities by analyzing the reasons for the referrals. Skiba et al. (2002) discovered that, while White students were more likely to be referred for objective infractions like smoking or vandalism, Black students were more likely to be referred for subjective offenses such as disrespect and noise. Overall, Black students faced consequences more often than White students because they were referred to the office more often, but this does not indicate whether Black students were being treated unfairly if the disparities in disciplinary consequences were due to more frequent

misbehavior by Black students. Of course, the fact that Black students were more likely to be referred for subjective offenses leaves open the possibility of unequal treatment by school leaders.

In a later study, Skiba et al. (2011) were able to dig a little deeper on the question of unequal treatment in an analysis of student-level data from 364 elementary and middle schools across the United States in 2005-06. Using logistic regression and multinomial logistic regression, the authors found that Black students are more likely than White students to be referred for a variety of disciplinary infractions, and that for the same infraction types, Blacks were significantly more likely to be given OSS or even expulsion. Thus, even after accounting for recorded infraction, Black students were more likely to receive exclusionary discipline.

Similarly, Anyon et al. (2014) used student level data from Denver Public Schools in 2011-12, and found that Black, Latino, Native American, and Multiracial students have increased risk, relative to White students, of receiving an office referral. Further, conditional on the referral reason (infraction type), Black students and Multiracial students were more likely than their White peers to receive out of school suspensions. Another contribution of Anyon et al.'s (2014) work, is the potential usefulness of restorative practices and in-school suspensions as alternatives to out-of-school suspensions. Still, this work is limited in its scope, as it focused on a single year in a single district, and also did not include detailed infraction-level data, as in the current study.

This initial round of studies using infraction-level data supported an unfortunate but unsurprising conclusion: marginalized students were not only punished more often, but also received more severe consequences for similar infractions. It is still unclear, however, whether these disparities occurred because of different disciplinary practices employed in schools serving mostly marginalized students, or because different groups of students were treated differently within the same schools. Only studies that include infraction-level data along with school characteristics or school fixed effects allow

researchers to understand the extent to which disciplinary disparities are occurring within or between schools. Thus far, we are aware of only one such study.

Skiba et al. (2014) began to address this gap by introducing school characteristics into a three-level HLM (Hierarchical Linear Modeling) analysis predicting student punishment as a function of infraction type. Using information from all students in the disciplinary database in a single Midwestern state in 2007-08, the authors found that even after controlling for type of infraction, Black students remained more likely to receive OSS, but statistically as likely to receive an expulsion. This analysis extended beyond prior work with the inclusion of school-level characteristics, such as student race and poverty and the principal's attitude toward discipline. In the third level analysis, the race of the individual student is no longer significant; school-level variables, including the concentration of Black students in the school, drive the severity of the punishments allocated. Thus, these results are consistent with the "racial threat hypothesis" in schools suggested by Welch and Payne (2010).

Overall, the evidence gathered thus far indicates clear racial disparities with respect to exclusionary discipline. Indeed, the Office for Civil Rights recently demonstrated nationwide racial disparities in discipline: Black students in 2011-12 made up only 16% of the students represented in the Civil Rights Data Collection (CRDC), but 33% of students suspended once, 42% of those suspended more than once, and 34% of students expelled. While the evidence on the existence of disparities is strong, less is known about what actually drives these disparate outcomes. Are Black students simply confined to schools with more severe disciplinary practices, or are Black students treated differently than white students within the same school walls despite being cited for identical school offenses?

While one study (Skiba et al., 2014) does find that school factors such as the concentration of Black students are a main driver of the severity of punishments administered, that single study has

6

limitations One potential weakness of the study is the setting – a single state that serves relatively few poor students (fewer than 40%) and very few Black students (8%). Moreover, given the importance of the question, policymakers need more information than what they can glean from a single study. More generally, there is still little evidence on the within-school racial disparities in disciplinary outcomes, for students committing similar types of infractions and with similar behavioral histories. Our study will contribute new evidence and enhance our collective understanding of disciplinary disparities facing Black students in US schools.

In addition, there is little known about how the severity of the punishment, as measured by number of days excluded from school, correlate with the racial background of the student? We focus in this study primarily on the length of suspension as a measure of the severity of a consequence, since suspensions and other learning time lost due to disciplinary processes are associated with lower academic achievement (Davis & Jordan, 1994; Public Agenda, 2004; Scott & Barrett, 2004), which is consistent with the mass of literature on the fact that increased opportunity for learning is associated with high achievement and large academic achievement gains (Brophy, 1988; Greenwood, Horton, & Utley, 2002; Hattie, 2002; Reynolds & Walberg, 1991; Wang, Haertel, & Walberg, 1997, among others). The length of a disciplinary punishment is readily interpretable and is also a detailed measure of a student's removal from the traditional classroom environment.

Research Questions

This study will attempt to answer two research questions:

1) Are certain racial subgroups receiving stricter (longer) punishments, controlling for the infraction committed, than their White grade-level peers *throughout the state*?

2) After controlling further for the school that a student attends, are certain subgroups of students still more likely to receive longer punishments? That is, do different students *in the same school* receive differing punishments for the same types of infractions?

Data and Sample

The study uses three years of de-identified student demographic and disciplinary data from all Arkansas K-12 schools provided by the Arkansas Department of Education (2010-11 through 2012-13). The demographic data include gender, race, grade level, special education status, limited English proficiency-status, and free-and-reduced-lunch (FRL) eligibility. The discipline information includes indicators for 19 infraction codes and 13 consequence codes, the date of the infraction, and the length of the consequence (number of days). To simplify the analysis, we collapsed infraction types involving handguns, rifles, and shotguns into a single category, resulting in 17 distinct categories. Furthermore, 13 consequence categories were collapsed into 7 (in school suspension (ISS), OSS, expulsion, referral to an alternative learning environment (ALE), corporal punishment, no action, and other).[1]

The unit of analysis for this study is the student-infraction level, so students can and often do have multiple observations within the same year. In some cases, there were exact duplicate records in the dataset, which would indicate an over reporting of a single incident. After removing duplicate entries (same student, date, infraction type, consequence, etc.), 651,804 total observations remain over the three-year period. 61,054 were excluded from the analytic sample in our preferred models due to

---

[1] Our measure of out-of-school suspension includes two separately reported OSS types: Out-of-School Suspension (when the incident did not result in physical injury) and Out-of-School Suspension (when the incident did result in physical injury). Our measure of expulsion includes five separately reported expulsion types: Expelled, Expelled for Weapons (as defined by Federal, State, and Student Discipline Policy), Expelled for Drugs (does not include alcohol or tobacco), Expelled for dangerousness (the incident did not result in physical injury), and Expelled for dangerousness (the incident resulted in physical injury). Our measure of ALE referrals includes two separately reported consequence types: Alternative Learning Environment (full year) and Alternative Learning Environment (less than one year). The other three consequence categories are Corporal Punishment, No Action, and Other.

missing an indicator for the number of days of punishment meted out, leaving 590,750 infractions in the analytic sample. The breakdown by infraction type and consequence can be seen in Tables 1 and 2.

Limitations of the Data

The data utilized in this analysis are individual-level administrative data collected by the state. To be included in this analysis, a student must have had a reported disciplinary infraction, so the dataset does not include every child in the state. For each recorded incident, we have both an infraction and a consequence type. The state-level data only includes certain types of infractions and consequences, so some categories used at a local level are coded as "other" at the state level. There is an "other" category for both infractions and consequences, but they are not always used together. As a result, a large number of cases are coded as "other" in either the infraction committed, the consequence received, or both. Over 95% of the "Other (Non-Specified)" consequences can be attributed to four types of infractions: disorderly conduct, insubordination, other (non-specified) infractions, and truancy. In the primary analysis, we limit the impact of this uncertainty because by only including in our sample the "other" consequences that included a number of days. Robustness checks using different assumptions are provided in the Appendices as well.[2] The exclusion of other (non-specified) consequences without a number of days listed is justified because these other (non-specified) consequences tend to be less severe and less exclusionary than a suspension.[3]

---

[2] 52% of incidences of "Other (Non-Specified)" consequence were dropped if and only if the data did not include a number of days of punishment for this consequence. If number of days was included, these infractions were included. These exclusions account for only about 9% of the total infractions in the dataset, and were disproportionately from the Northwest (46%) and Central (27%) regions of the state. Two districts make up about 59% of these infractions. See Appendix A for a robustness check.

[3] Conversations with the Arkansas Department of Education Assistant Commissioner for Research and Technology, Eric Saunders, indicates that the majority of these other consequences are detentions, bus suspensions, parent/guardian conferences, Saturday school, or warnings.

Finally, we should also acknowledge there are still key factors that we cannot observe in this dataset (or in any dataset). To illustrate the sequence of events connected to school disciplinary infractions, we walk through a simple example.

1) A student verbally abuses a teacher in front of a school administrator

2) The administrator records an insubordination infraction for that student

3) The administrator issues a 2-day ISS to the student

Most empirical disciplinary analyses employ data from event 3. Our analysis is distinguished from many others in that we also have data on event 2 (the specific citation). Unfortunately, our dataset provides no information on the actual actions of the students and or administrators involved; we are unable to observe the first event from our fictitious example above. As a result, we don't observe misbehaviors that were never recorded, and we do not know the severity of the infraction committed. Certain infraction categories are more subjective than others. For example, the three biggest categories of infractions (disorderly conduct (30.2%), insubordination (26.2%), and other non-specified (20.6%)) can include a wide range of behaviors and are subject to teacher or administrator discretion. Indeed, some have argued that administrator discretion in the use of citations for subjective offenses may be one cause of disparate disciplinary outcomes (Gregory & Weinstein, 2008; Skiba et al., 2011; Skiba et al., 2002). Nevertheless, our ability to at least control for the type of recorded infraction (as well as past infraction history) enhances our ability to identify per infraction disparities.

Another limitation of these data is that they are reported by schools, collected by the Arkansas Department of Education, and then provided to us for research purposes; however, the data have not gone through an external validation process. While it is possible (and even likely) that school districts are reporting in slightly different ways, we can reduce the impact of such differences through the use of school-fixed effects. Furthermore, these limitations exist in most large-scale studies on school discipline, because there is no practical way to ensure perfectly consistent reporting across districts.

With these limitations in mind, we now turn to the analytic methods we use to explore the relationship between race and disciplinary outcomes.

Method

Research Question 1: Are certain racial subgroups receiving stricter (longer) punishments than their White grade-level peers *throughout the state*?

To generate a first estimate of the relationship between race and discipline outcomes, we utilize Stata to estimate a multiple regression model predicting the days of punishment as a function of race and several control variables that could reasonably be expected to affect the level of punishment. We include grade level indicators because exclusionary discipline is more common in higher grades, so we only want to compare outcomes for student within grade. In addition, we include an indicator for special education status, because special education students are limited to ten consecutive days of OSS and face other unique circumstances that might reasonably influence the disciplinary practices for such students (Arkansas Department of Education, 2008).

An important contribution of this analysis is that we are able to attach specific consequences to particular infraction types and the student's behavioral history and control for this in our models. We include indicators for each of the 17 infraction categories plus indicators for the order of infractions. The order of infractions indicators include indicators for whether the infraction was that student's first, second, third, fourth, fifth, sixth, or seventh or more consequence within the same school in the same year (student-school-year), plus an indicator for the last infraction for that student-school-year. Given that the last infraction is more likely to be at the end of the year or perhaps the final infraction before a student is expelled or transfers out of a school, it makes sense to control for this variable. Further, in order to compare only similar consequence types, indicators for the type of consequence (ISS, OSS, expulsion, corporal punishment, etc.) are included. We utilize pooled ordinary least squares (OLS) regression with heteroskedastic robust standard errors clustered at the student level (Huber, 1967;

11

White, 1980). Since the unit of analysis is the infraction, students may be included more than once, even within a given year, and clustered standard errors (at the student level) allows for correlation among multiple outcomes for an individual student.

Another key feature of this study is our ability to consider the severity of each consequence, as measured by the number of days. Defining and operationalizing this outcome variable (days of punishment), however, took careful consideration due to the types of consequences that exist. Each consequence is characterized by both a type (ISS, OSS, expulsion, etc.) and an amount (such as a number of days in some cases). For some punishments (ISS, OSS, referral to an ALE, expulsion), having a number of days of punishment is clearly logical. For others, such as corporal punishment, *instances* are more important than a number of days. An "other" (non-specified) consequence is ambiguous.

To create a measurable outcome variable, we focus solely on the number of days as a proxy for lost instructional time and we ignore the possibility that different types of consequences may vary in "severity".. This adds simplicity in interpretation, but also recognizes that it is unclear how to lay out these punishments in an ordinal way. For example, OSS may be considered more severe than ISS to school personnel, but to a student, it may be viewed as a vacation from school and actually preferred; to be certain, both options clearly remove the student from the classroom.

Corporal punishment presents a unique challenge for our "number of days" outcome. In the primary model, we include over 93,000 corporal punishment cases as outcomes where the days of punishment is equal to zero days, but given that we include consequence type indicators, corporal punishment cases don't help us identify our coefficient of interest, yet these cases do assist with precision in our models. In addition, our consequence type indicators essentially control for differences in length of consequence that are confounded with the type of consequence. For example, expulsions

and referrals to alternative learning environments are very different kinds of consequences, and tend to also have many more days recorded with the consequence.[4]

Furthermore, since each unit of analysis is a single observation, it is important to note that the coefficients we estimate are on a *per infraction* basis. This means that we are estimating the impact of race on the number of days punished *per infraction*, conditional on being in the infraction database to begin with.

The baseline model is:

$$days\_pun_{it} = \beta_0 + \beta_1 race_{it} + \beta_2 SpEd_{it} + \beta_3 grade_{it} + \beta_4 infraction\_type_{it}$$
$$+ \beta_5 infraction\_order_{it} + \beta_6 consequence\_type_{it} + u_{it}$$

Where:

$days_{pun_{it}}$ = number of days of punishment (instances of corporal punishment = 0 day)

$race_{it}$ = a vector of race indicators for Black, Hispanic, Native American, Asian/Pacific Islander, and Two or More Races (White as baseline)

$SpEd_{it}$ = an indicator for special education student (1) or not (0)

$grade_{it}$ = a vector of grade indicators (with 8th grade as the baseline)

$infraction\_type_{it}$ = a vector of infraction indicators (disorderly conduct as the baseline)

$infraction\_order_{it}$ = a vector of indicators for first, second, third, fourth, fifth, sixth, seven or more infraction (student by school by year) plus last infraction (student by school by year)

$consequence\_type_{it}$ = a vector of consequence indicators (ISS as the baseline)

In cases where the outcome variable (days punished) is missing, the preferred model drops observations if the consequence was something that should reasonably have days attached (ISS, OSS, expulsion, ALE, or other). Dropping these observations reduces the analytic sample in our preferred

---

[4] We also tested how much our results differ if we do not include these consequence type controls, and the estimates in our preferred models are more conservative (lower) than if we don't control for consequence type.

model from 651,804 to 590,750 (see Table 2). The majority of cases dropped were other (non-specified) outcomes (57,751), followed by ALE (1,991).[5] For example, if an incident of OSS was recorded, but without a days of punishment attached, in the primary model we just drop this observation, since the dependent variable is missing. All observations with a consequence of "no action" are coded with zero days, as this is an indication that no consequence was given. As noted earlier, all instances of corporal punishment are coded as zero days of punishment. The primary goal was to predict the amount of missed instructional time, so corporal punishment is not treated as a full day of missed time in the preferred model.[6]

The assumptions previously described built the preferred model, however, we provide results under various assumptions in Appendix A. In addition, since school practices related to recording these types of data may well be related to whether this dependent variable is missing, we have somewhat more confidence in the within-school estimates (which include school fixed effects).

In sum, this model asks whether certain racial groups are receiving stricter (longer) punishments than their white peers (in the same grade and with the same special education status) in the state for the same infraction type and with similar infraction history. Next, we ask whether certain subgroups of students receive longer punishments than their peers *within the same school.*

Research Question 2: After controlling further for the school that a student attends, are certain subgroups of students still more likely to receive longer punishments? That is, do different students *in the same school* receive differing punishments for the same types of infractions?

---

[5] Of the 61,054 infractions dropped from the preferred model, relative to what we would expect from the larger sample, the infractions that were dropped were more likely to be for lower grade levels (kindergarten through fifth), more likely to be for non-FRL and non-Black students, and were more likely to occur in the third year of our sample (2012-13). Certain regions of the state were also more likely to not report the number of days for some of the consequence types that typically have days attached. Overall, however, 88% of the school-by-year combinations had no missing day issues, so these are concentrated in a small number of schools.

[6] Corporal punishment is still viewed by many as detrimental to students, regardless of the fact that a student is not missing class for days at a time as a result (American Psychological Association, 2015; Council of Europe, 2007; Flynn, 1999; Welsh et al., 1997).

There are several reasons why discipline policies and outcomes should be compared within a specific school location, which justifies the use of school-fixed effects. First of all, policies are set locally, and will vary across the state. Even within a district, application of policies can vary by school due to differences in interpretation or implementation. Furthermore, school-fixed effects control for other school-level characteristics that could be associated with disciplinary outcomes (school climate, poverty levels, region, urban or rural classification, demographics, grade levels served, etc.). Most importantly, however, these fixed effects allow us to ask our second fundamental question.

In Question 1, we asked whether certain types of students receive longer punishments for the same type of infraction *across the state*. It is possible that the disparities, if present, are the result of certain races being more likely to attend schools with longer punishments on average. Alternatively, it could be that, even within the same school, certain groups of students receive disproportionately severe punishments. To disentangle this relationship, Model 2 incorporates the same covariates in Model 1, with the addition of school level fixed effects.

Model 2, for research question 2, including the addition of school-fixed-effects (λ_s) is:

$$days_{pun_{it}} = \beta_0 + \beta_1 race_{it} + \beta_2 SpEd_{it} + \beta_3 grade_{it} + \beta_4 infraction_{type_{it}} + \beta_5 infraction_{order_{it}}$$
$$+ \beta_6 consequence_{type_{it}} + \lambda_s + u_{it}$$

Researchers and policy makers may question whether apparent racial disparities might be correlated with income disparities. Thus, to test whether the driver of disproportionalities was actually economic disadvantage, we conduct parallel analyses in which we include a proxy for family income (free- and reduced- price lunch eligibility). These analyses test whether race still predicts disciplinary outcomes, above and beyond socio-economic status. If we find that the racial disparities are reduced, but there are disproportionalities based on FRL-status, it simply indicates that disproportionalities

15

exist, but can be explained by demographic characteristics other than just race. Therefore, we include a robustness check with models that include an indicator for a student's FRL status.

Results

Research Question 1: Are certain racial subgroups receiving stricter (longer) punishments than their White grade-level peers *throughout the state*?

First, we compute racial differences in the number of days of consequence, controlling for other factors. The two non-White subgroups that we focus on are Black and Hispanic, which make up about 21% and 10% of the state's enrollment, respectively. Other races (Asian/Pacific Islander, Native American, and two or more races) make up a relatively small proportion of the state (less than 5%), so we do not discuss the results for these subgroups separately. It is important to note that, while Black students make up 21% of the state's public school enrollment, they represent almost 44% of the disciplinary incidents included in our dataset. On the other hand, Hispanic students are more proportionally represented in the disciplinary data; they make up 10% of public school enrollment and 9% of the disciplinary incidents.

The first two columns in Table 3 correspond to the two research questions laid out. The results in column 1 of Table 3 indicate the state level disparities by race, and the results in column 2 indicate the disparities within school. The dependent variable is the total number of days of punishment, so our results indicate that Black students receive about 0.33 extra days of punishment *per incident* relative to White students in the state, holding the other covariates fixed. Hispanic students receive about 0.15 days *less* per incident relative to White students in the state.

We also test whether the inclusion of economic disadvantage in our model mitigates the racial impact on disciplinary outcomes (columns 3 and 4 of Table 3). Even controlling for FRL-eligibility as, the coefficients on Black and Hispanic are relatively unchanged both in terms of significance and magnitude, indicating that racial-disparities exist even above and beyond socioeconomic status.

Differences by Infraction Type

Analyzing observations of a single infraction category at a time serves as a robustness check for Model 1, as well as a potential indicator of where to focus policies to correct for the disparities. Subjective offenses (e.g. disorderly conduct and insubordination) have been linked to disparate application of discipline (Gregory & Weinstein, 2008; Skiba et al., 2011; Skiba et al., 2002), so we can compare disparities in these categories with those in more objective categories such as truancy.

The results in Table 4 indicate that the longer punishments for Black students relative to their White peers varies in magnitude but is still statistically significant for each of the five most frequent infraction categories. We see disparities of about one third of a day for both disorderly conduct and insubordination (two subjective infractions), and an even larger disparity in the outcomes for fighting infractions. On the other hand, for the very objective infraction of truancy, we see that the racial disparity is much smaller (.08). In addition, Hispanic students tend to receive shorter punishments for disorderly conduct, other (non-specified) infractions, and insubordination.

Research Question 2: After controlling further for the school that a student attends, are certain subgroups of students still more likely to receive longer punishments? That is, do different students *in the same school* receive differing punishments for the same types of infractions?

Column 2 of Table 3 includes the same covariates as in Column 1, but with the addition of school fixed effects. In column 2, we find that the within school differences are much smaller than those in Column 1. Black students receive about 0.07 extra days of punishment *per incident* relative to White students in the same school, and there is not a statistically significant difference between these outcomes for Hispanic and White students within the same school.

We also compare how the coefficients on Black and Hispanic change with the addition of school-fixed effects. Table 5 compares the coefficients on Black (panel A) and Hispanic (panel B) from Model 1 (without school-fixed effects) and Model 2 (with school-fixed effects) for the overall dataset and for

17

each of the five most common infractions, separately. Once school-fixed effects are employed and we examine within-school differences, the coefficient on Black always diminishes in magnitude and in some cases loses statistical significance or becomes negative. For example, Black students received about 0.72 more days of punishment than White students across the state for fighting, but only 0.19 more days than White students within the same school. Statistically significant disproportionalities between Black students and White students at the state level for disorderly conduct and other (non-specified) infractions both lost statistical significance after the inclusion of school fixed effects. For Hispanic students, this trend is even more pronounced – all statistically significant differences at the state-level disappear in within-school analyses. Thus, any differences observed related to Hispanic students are due to differences in schools attended by Hispanic students.

Robustness Checks

Our dependent variable in the primary model was operationalized with a focus on days of punishment, driven by concern over missing instructional time, along with the belief that more days out of school might decrease a student's engagement with the school. In this model, we include days of punishment for ISS and referrals to ALE in addition to more clearly exclusionary consequences such as OSS and expulsions. However, it could be that an ISS, for example, occurs without much loss of instruction. Appendix A presents the results of a robustness check that treats only the most clearly exclusionary discipline consequences (expulsion and OSS) as missed instructional days, and codes all other types of consequences including even things that have a number of days associated such as ISS and ALE referrals as zero days. The results from these models are in column 3 and 4 of Appendix A, which can be compared to the main results from Table 3 (which are reproduced in columns 1 and 2 of Appendix A). We see that the Black-White disparities both at the state-level and within-school remain significant but diminishes in magnitude, and the Hispanic-White disparity at the state-level is now zero.

Lastly, we conduct a robustness check for our assumptions dealing with missing data. There were over 57,000 infractions in the original dataset that coded as an "other/non-specified" consequence, but with no days of punishment. In the primary analysis, since the outcome did not include a number of days, and because they are generally relatively less severe consequences,[7] these were dropped, however we perform a robustness check for this decision. The results are in columns 5 and 6 of Appendix A. The substantive conclusion is generally the same as in our preferred models from Table 3 (reproduced in columns 1 and 2 of Appendix A), however, the coefficients on Black and Hispanic are slightly smaller.

As an additional robustness check, separate analyses were run for eighth through twelfth grade and kindergarten through seventh grade, and the results were robust to this decision as well.[8]

Discussion

There are racial disparities in the length of disciplinary consequences imposed on students across the state of Arkansas. While our analyses confirm that fact, the real contribution of this study is measuring the different treatment of students with school fixed effects and controls for behavior. We find that, even after controlling for student infractions, there are large racial disproportionalities in the length of punishments across the state. However, the within-school disparities are much smaller. The good news is that, for the most part, educators within the same schools are not systematically administering more severe punishments to Black students for similar offenses. The bad news is that, statewide, Black students do systematically attend schools with more severe disciplinary responses.

So what is driving these between school differences? One possible explanation for the difference between the state-level and school-level estimates of the Black-White disparity, for example, is that Black students could be disproportionately attending schools with stricter discipline

---

[7] Conversations with the Arkansas Department of Education Assistant Commissioner for Research and Technology, Eric Saunders, indicates that the majority of these other consequences are detentions, bus suspensions, parent/guardian conferences, Saturday school, or warnings.

[8] These results can be made available at the reader's request.

policies. In fact, during this three year time period, schools with the greatest proportion of Black students did indeed hand out more suspensions and longer suspensions.[9] It is possible as well that principal attitudes within different schools drive disciplinary practices. Raush and Skiba (2005) concluded that principals who blamed behavioral problems on poverty and poor parenting tended to employ more suspensions (Raush & Skiba, 2005). Similarly, Skiba et al. (2003) found greater use of preventative measures and lesser use of exclusionary measures in schools with principals who self-reported a belief that suspension and expulsion were unnecessary. Further, in the national report, *Opportunities Suspended* (Advancement Project/Civil Rights Project, 2000)*,* investigators suggested that the use of suspensions was closely aligned with the support for zero tolerance policies and procedures. With particular respect to racial disparities in disciplinary outcomes, Skiba et al. (2014) find that principal perspectives on discipline are also related to racial disparities in suspension and expulsion. Thus, it is possible that principals in schools with more Black students have particular beliefs about disciplinary strategies. In fact, the analysis of Welch and Payne (2010) revealed that school leaders serving higher proportions of Black students preferred more punitive disciplinary styles.

Overall, then, our results are mostly consistent with the only other study to date that attempted to disentangle between-school and within-school drivers of racial disparities in disciplinary consequences (Skiba et al., 2014). In short, Black students overall are subject to stricter consequences, but the source of these differences appear to be differing schoolwide practices. In some ways, this finding might be viewed as a positive one for policymakers seeking remedies for these disparities. Rather than having to venture into every school building and work with school leaders who are treating Black students differently than their White peers for similar offenses, policymakers can focus their

---

[9] To test whether this, we split the schools into quintiles by the percentage of students who are Black and analyzed the frequency and duration of various consequences within each quintile.

strategies on the school sites where overly strict discipline is being practiced. And, using available data in many states, policymakers should be able to identify these schools.

In terms of remedies, the infraction-specific models can provide state and local policy makers with an indication of which infraction types are most likely to result in disparate consequences. For example, the Black-White disparities within schools appear to be driven by incidences of fighting or insubordination. While in general, Hispanic students received shorter punishments than White students, Hispanic students do receive longer punishments than White students for fighting. Moreover, qualitative research on principal attitudes or other school-level factors could add to the knowledge base about why these disparities are occurring and how they might be minimized. In addition, subjective offenses (such as disorderly conduct and insubordination) have been linked to disparate application of discipline (Gregory & Weinstein, 2008; Skiba et al., 2011; Skiba et al., 2002). We see some evidence consistent with this given our results in Table 4, which shows large disproportionalities for Black students relative to White students at the state level, but these disparities are much smaller or insignificant within schools (Table 5).

Most importantly, if policymakers can identify the schools that are meting out particularly high levels of exclusionary discipline, the next step is to offer strategies or resources for more effective discipline. Unfortunately, to date, there is a relatively thin literature on effective alternatives, but we discuss some of the available resources and research on alternatives in the next section.

In 2014, the Department of Education and the Department of Justice released a package of resources and informational materials to assist with school climate and discipline (US Department of Education and US Department of Justice, 2014). In addition, many states have revised their laws, and many school districts have changed their policies to reduce reliance on out-of-schools suspensions for more minor offenses or for certain types of offenses. According to Steinberg and Lacoe (2017), as of May 2015, 22 states the District of Columbia had revised their laws in order to limit exclusionary

discipline or implement more supportive and preventative strategies for at-risk students. In addition, as of the 2015–16 school year, 23 of the nation's 100 largest school districts changed policies to require non-punitive discipline strategies and/or limit suspension use (Steinberg & Lacoe, 2017). There is also some promising evidence that revising student codes of conduct in this way can reduce the use of suspensions for minor offenses and limit the length of suspensions may be effective (Lacoe & Steinberg, 2017; Mader et al., 2016) and at little cost to school climate (Mader et al., 2016).

Non-experimental evidence shows some reduction in referrals or suspensions and expulsions within certain programs such as Response to Intervention (RTI), which attempts to prevent recidivism by responding to behavioral issues as they arise, finds that office referrals decreased following RTI implementation (Fairbanks et al., 2007), restorative justice (Fronius et al., 2016), or some combination of these two strategies (Collins-Ricketts & Rambo, 2015).

Further, there is some experimental evidence on School-Wide Positive Behavioral Interventions and Supports (SWPBIS a.k.a. PBIS) that indicates implementation of a PBIS framework has a variety of positive impacts such as decreases in office referrals (Flannery et al., 2014), and improvement in student perceptions of school safety and test scores (Horner et al., 2009). There is also some experimental evidence that the My Teacher Partner Program (MTP), which encourages focuses on teacher interactions with students reduced suspensions in the treatment group (Gregory et al., 2014).

Finally, policymakers and researchers alike may view our current lack of knowledge alongside our growing interest in addressing this problem as an opportunity. Right now, there are practitioners and program representatives supporting a set of potentially effective interventions. At the same time, there are schools struggling with overly severe disciplinary policies and disparate disciplinary outcomes. This situation is ripe for policymakers to bring together practitioners and researchers to field numerous trials of these potentially effective strategies and rigorously study the results. If done well,

systematically employing more preventative and intensive support structures and rigorously studying

the outcomes could ultimately reduce the disproportionalities in total referrals as well. Most

importantly, in the long run, this strategy might lead to improved outcomes for marginalized students.

References

Advancement Project/Civil Rights Project. (2000). *Opportunities suspended: The devastating consequences of zero tolerance and school discipline.* Cambridge, MA.

American Academy of Pediatrics. (2013). Out-of-School Suspension and Expulsion. *Pediatrics, 131(3).*

American Psychological Association. 2008. Are Zero Tolerance Policies Effective in Schools? An Evidentiary Review and Recommendations. *American Psychologist, 63(9),* 852-862.

American Psychological Association. (2015). *Corporal Punishment.* Retrieved from http://www.apa.org/about/policy/corporal-punishment.aspx

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

Arkansas Advocates for Children and Families. (2013). Keeping Kids in Class: Fixing Racial Disparities in School Discipline. Retrieved from http://www.aradvocates.org/publications/keeping-kids-in-class-fixing-racial-disparities-in-school-discipline/

Arkansas Department of Education. (2008). *Special Education and Related Services 11.00 Discipline Procedures.* Retrieved from https://arksped.k12.ar.us/rules_regs_08/1.%20SPED%20PROCEDURAL%20REQUIREMENTS%20AND%20PROGRAM%20STANDARDS/11.00%20DISCIPLINE%20PROCEDURES.pdf

Balfanz, R., Spiridakis, K., Neild, R. C., & Legters, N. (2003). High-poverty secondary schools and the juvenile justice system: How neither helps the other and how that could change. *New Directions for Youth Development,* No. 99.

Beck, A. N. & Muschkin, C. G. (2012). The Enduring Impact of Race: Understanding the Disparities in Student Disciplinary Infractions and Achievement. *Sociological Perspectives, 55(4),* 637-662.

Brophy, J. (1988). Research linking teacher behavior to student achievement: Potential implications for instruction of Chapter 1 students. *Educational Psychologist, 23*(3), 235-386.

Children's Defense Fund. (1975). School Suspensions: Are they Helping Children? Washington Research Project, Cambridge, MA. Retrieved from http://files.eric.ed.gov/fulltext/ED113797.pdf

Collins-Ricketts, J. E. & Rambo, A. (2015). The PROMISE program case examples: From get tough to solution building. *International Journal of Solution-Focused Pracctices, 3*(2), 17-22.

Council of Europe. (2007). *Abolishing corporal punishment of children: Questions and answers.* France: Council of Europe Publishing.

Davis, J. E., & Jordan, W. J. (1994). The effects of school context, structure, and experiences on Black males in middle and high schools. *Journal of Negro Education, 63*, 570-587.

Ekstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who Drops Out of High School and Why? Findings from a National Study. *The Teachers College Record, 87(3),* 356-373.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking Schools' Rules: A Statewide Study of How School Discipline Relates to Students' Success and Juvenile Justice Involvement.* The Council of State Governments Justice Center & Public Policy Research Institute.

Fairbanks, S., Sugai, G., Guardino, D. & Lathrop, M. (2007). Response to intervention: Examining classroom behavior support in second grade. *Exceptional Children, 73*(3), 288-310.

Flannery, K.B., Fenning, P., McGrath Kato, M., & McIntosh, K. (2014). Effects of School-Wide
Positive Behavioral Interventions and Supports and fidelity of implementation on problem
behavior in high schools. *American Psychological Association, 29*(2), 111-124.

Flynn, C. P. (1999). Exploring the link between corporal punishment and children's cruelty to animals.
Journal of Marriage and the Family, 61, 971-981.

Fronius, T., Persson, H., Guckenberg, S., Hurley, N., & Petrosino, A. (2016). *Restorative justice in
U.S. schools: A research review.* West Ed.

Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic engagement: Current perspectives
on research and practice. *School Psychology Review, 31*, 328-349.

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2014). The promise of a
teacher professional development program in reducing racial disparity in classroom
exclusionary discipline. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable
remedies for excessive exclusion* (pp. 166–179). New York: Teachers College Press.

Gregory, A., & Weinstein, S. R. (2008). The discipline gap and Blacks: Defiance or cooperation in the
high school classroom. *Journal of School Psychology: 46,* 455–475.

Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational
Research*, *37*, 449-482.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009).
A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior
support in elementary schools. *Journal of Positive Behavior Interventions, 11*(3), 133-144.

Huber, P. J. (1967). *The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions*,
Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1,
221–233. Berkeley: University of California Press.

Imich, A. J. (1994). Exclusions from school: Current trends and issues. *Educational Research, 36 (1),* 3-11.

Lacoe, J. R., & Steinberg, M. P. (2016). Rolling back zero tolerance: The effect of discipline policy reform on suspension usage, school climate, and student achievement. Manuscript in preparation.

Losen, D. & Gillespie, J. (2012). Opportunities Suspended: The Disparate Impact of Disciplinary Exclusion from School. Civil Rights Project, UCLA.

Losen, D., Hodson, C., Keith, M., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* The Civil Rights Project, UCLA.

Losen, D. & Skiba, R. (2010). *Suspended Education: Urban Middle Schools in Crisis.* Report: The Civil Rights Project at UCLA, The Equity Project at the Indiana University Center for Evaluation and Education Policy, and the Southern Poverty Law Center.

Mader, N., Sartain, L., & Steinberg, M. (2016). When suspensions are shorter: The effects on school climate and student learning. Manuscript in preparation.

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the Impact of School Discipline on Racial Disproportion in the Juvenile Justice System. *Social Science Quarterly, 90(4),* 1003-1018.

Public Agenda. (2004). *Teaching interrupted: Do discipline policies in today's public schools foster the common good?* Retrieved from http://www.publicagenda.org/files/teaching_interrupted.pdf

Raffaele-Mendez L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development,* No. 99.

Raffaele-Mendez, L. M., Knoff, H. M., & Ferron, J. M. (2002). School Demographic Variables and Out-of-School Suspension Rates: A Quantitative and Qualitative Analysis of a Large, Ethnically Diverse School District. *Psychology in the Schools, 39(3),* 259-277.

Rausch, M. K., & Skiba, R. J. (2005). The academic cost of discipline: The contribution of school
    discipline to achievement. Paper presented at the annual meeting of the American Educational
    Research Association, Montreal, Quebec.

Reynolds, A. J., & Walberg, H. J. (1991). A structural model of science achievement. *Journal of
    Educational Psychology, 83*, 97-107.

Sartain, L., Allensworth, E. M., & Porter, S. *with* Levenstein, R., Johnson, D. W., Huynh, M. H.,
    Anderson, E., Mader, N., & Steinberg, M. P. (2015). *Suspending Chicago's Students:
    Differences in Discipline Practices across Schools.* The University of Chicago Consortium on
    Chicago School Research.

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

Skiba, R. (2014). The Failure of Zero Tolerance. *Reclaiming Children and Youth, 22*(4), 27-33.

Skiba, R., Horner, R., Chung, C. G., Rausch, M. K., May, S., & Tobin, T. (2011). Race is Not Neutral:
    A National Investigation of Black and Latino Disproportionality in School Discipline. *School
    Psychology Review, 40(1),* 85-107.

Skiba, R., Michael, R., Nardo, A., & Peterson, R. (2002). The Color of Discipline: Source of Racial
    and Gender Disproportionality in School Punishment. *The Urban Review, 34(4),* 317-342.

Skiba, R. & Peterson, R. (1999). ZAP Zero Tolerance. *Phi Delta Kappan, 80,* 372-76, 381-82.

Skiba, R. & Rausch, M. K. (2004). The Relationship between Achievement, Discipline, and Race: An
    Analysis of Factors Predicting ISTEP Scores. Bloomington: Center for Evaluation and
    Education Policy.

Skiba, R., Simmons, A., Staudinger, L., Rausch, M., Dow, G., & Feggins, R. (2003). Consistent

Removal: Contributions of School Discipline to the School-Prison Pipeline. Paper presented at

the School to Prison Pipeline Conference, Harvard Civil Rights Conference.

Skiba. R. J, Chung, C., Trachok, M., Baker, T., Sheya, A., & Hughes, R. (2014). Parsing Disciplinary

Disproportionality: Contributions of Infraction, Student, and School Characteristics to Out-of-

School Suspension and Expulsion. *American Educational Research Journal, 51(4),* 640-670.

Steinberg, M. P. & Lacoe, J. (Winter 2017). What do we know about school discipline reform?

Education Next, *17*(1).

US Department of Education and US Department of Justice. (2014). Dear Colleague Letter on the

Nondiscriminatory Administration of School Discipline. Retrieved from:

http://www2.ed.gov/about/offices/list/ocr/letters/colleague-201401-title-vi.html

US Department of Education's Office for Civil Rights. (March 2014). *Civil Rights Data Collection*

Data Snapshot: School Discipline. Issues Brief No. 1. Retrieved from:

http://ocrdata.ed.gov/Downloads/CRDC-School-Discipline-Snapshot.pdf

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). Learning influences. In H. J. Walberg & G. D.

Haertel (Eds.), *Psychology and educational practice* (pp. 199-211). Berkeley: McCutchan.

Welch, K. & Payne, A. N. (2010). Racial threat and punitive school discipline. *Social Problems, 57(1),*

25-48.

Welsh, W., Jenkins, P., & Greene, J. (1997). *Building a Culture and Climate of Safety in Public*

Schools in Philadelphia: School-Based management and Violence Reduction. Philadelphia:

Center for Public Policy.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for

Heteroskedasticity. *Econometrica: 48,* 817–838.

Table 1: Arkansas K-12 Discipline Data by Infraction (2010-11 to 2012-13)

	Analytic Sample	
	Number	*%*
Disorderly Conduct	173,400	29.4%
Insubordination	151,091	25.6%
Other (Non-specified)	122,925	20.8%
Fighting	49,128	8.3%
Truancy	46,410	7.9%
Bullying	16,506	2.8%
Tobacco	8,995	1.5%
Student Assault	6,996	1.2%
Drugs	6,390	1.1%
Miscellaneous*	8,909	1.5%
Total	**590,750**	**100.0%**

*Includes: Vandalism, Knife, Staff Assault, Alcohol, Gangs, Guns, Club, Explosives

Table 2: Arkansas K-12 Discipline Incidents by Consequence (2010-11 to 2012-13)

	# of Disiplinary Incidents		Analytic Sample*	
	Number	*%*	**Number**	*%*
In-School Suspension	261,759	40.2%	260,766	44.1%
Out-of-School Suspension	175,012	26.9%	174,850	29.6%
Other (Non-Specified)	111,785	17.2%	54,034	9.1%
Corporal Punishment	93,222	14.3%	93,222	15.8%
No Action	5,159	0.8%	5,159	0.9%
Referral to Alternative Learning Environment	3,021	0.5%	1,030	0.2%
Expulsion	1,846	0.3%	1,689	0.3%
Total	**651,804**	**100.0%**	**590,750**	**100.0%**

*Analytic sample excludes observations dropped for missing outcome variable, as noted in Methods section

Table 3: OLS Regression of Days Punished Per Infraction, Arkansas K-12 (2010-11 to 2012-13)

	(1)	(2)	(3)	(4)
Black	0.329 **	0.072 **	0.319 **	0.071 **
	(0.018)	(0.021)	(0.018)	(0.021)
Hispanic	-0.145 **	0.012	-0.156 **	0.011
	(0.032)	(0.031)	(0.032)	(0.031)
Asian	-0.007	0.106	-0.013	0.106
	(0.093)	(0.092)	(0.093)	(0.092)
Native American	-0.142 **	0.035	-0.148 **	0.035
	(0.052)	(0.048)	(0.052)	(0.048)
Two or More Races	-0.048	0.019	-0.054	0.019
	(0.038)	(0.037)	(0.038)	(0.037)
SPED Status	-0.099 **	-0.099 **	-0.101 **	-0.099 **
	(0.023)	(0.020)	(0.023)	(0.020)
FRL Status			0.056 **	0.003
			(0.014)	(0.014)
Expulsion	53.36 **	53.23 **	53.36 **	53.23 **
	(1.860)	(1.850)	(1.860)	(1.850)
Constant	2.078 **	2.334 **	2.044 **	2.332 **
	(0.040)	(0.043)	(0.041)	(0.044)
Grade Indicators	X	X	X	X
Infraction Type Indicators	X	X	X	X
Infraction Order Indicators	X	X	X	X
Consequence Indicators	X	X	X	X
School Fixed Effects		X		X
Observations	590,750	590,750	590,750	590,750
Adjusted R-Squared	0.3749	0.3994	0.3632	0.3882

Notes Coefficients with robust standard errors, clustered at the student level in parentheses.
**p<.01, *p<.05
Dependent variable = count of number of days of punishment.
  Preferred model: dropped if missing days punished (except No Action = 0 and Corporal Punishment = 0).
Baseline grade: 8th. Baseline infraction: disorderly conduct. Baseline consequence: ISS.
Infraction order indicators include: indicators for first, second, third, fourth, fifth, sixth, and seventh or more, plus an indicator for last infraction (baseline = 6th infraction).

Table 4: OLS Regression of Days Punished by Infraction, Arkansas K-12 (2010-11 to 2012-13)

	Disorderly Conduct	**Insubordination**	**Other**	**Fighting**	**Truancy**
Black	0.366 **	0.331 **	0.160 **	0.716 **	0.083 **
	(0.029)	(0.021)	(0.030)	(0.049)	(0.029)
Hispanic	-0.279 **	-0.125 **	-0.210 **	0.162	0.0864
	(0.033)	(0.048)	(0.038)	(0.086)	(0.073)
Asian	-0.239	-0.160	-0.099	0.213	0.296 **
	(0.141)	(0.112)	(0.072)	(0.140)	(0.081)
Native American	-0.262 **	0.078	-0.168 **	-0.173	-0.196 *
	(0.050)	(0.140)	(0.055)	(0.194)	(0.093)
Two or More Races	-0.141 **	-0.049	-0.055	0.100	-0.067
	(0.050)	(0.048)	(0.060)	(0.130)	(0.060)
SPED Status	-0.180 **	-0.050	0.01	-0.154 **	0.033
	(0.032)	(0.030)	(0.047)	(0.055)	(0.047)
Expulsion	46.95 **	34.55 **	61.43 **	39.28 **	27.53 **
	(4.259)	(4.868)	(6.381)	(5.836)	(6.781)
Constant	1.975 **	1.909 **	1.806 **	3.233 **	2.060 **
	(0.053)	(0.051)	(0.047)	(0.121)	(0.053)
Grade Dummies	Y	Y	Y	Y	Y
Infraction Order Dummies	Y	Y	Y	Y	Y
Consequence Dummies	Y	Y	Y	Y	Y
School fixed effects	N	N	N	N	N
Observations	173,400	151,091	122,925	49,128	46,410
Adjusted R-Squared	0.2999	0.2434	0.3647	0.2695	0.2216

Notes Coefficients with robust standard errors, clustered at the student level in parentheses.
**p<.01, *p<.05
Dependent variable = count of number of days of punishment.
Preferred model: dropped if days punished was missing (except No Action set to 0 and Corporal Punishment set to 0)
Baseline grade: 8th. Baseline consequence: ISS.
Infraction order dummies include: dummies for first, second, third, fourth, fifth, sixth, and seventh or more, plus a dummy for last infraction (baseline = 6th infraction).

Table 5: Race Coefficients from Model 2 and Model 3, Overall and for 5 Most Frequent Infractions
Panel A: Coefficents on Black   **Panel B: Coefficients on Hispanic**

	State-Level	With School-Fixed Effects		State-Level	With School-Fixed Effects
Overall	**0.329** **	**0.072** **	**Overall**	**-0.145** **	**0.012**
	-0.018	**(0.021)**		**(0.032)**	**(0.031)**
Disorderly Conduct	0.366 **	0.0478	Disorderly Conduct	-0.279 **	-0.046
	-0.029	(0.031)		(0.033)	(0.028)
Insubordination	0.331 **	0.078 **	Insubordination	-0.125 **	0.045
	-0.021	(0.026)		(0.048)	(0.050)
Other	0.160 **	0.003	Other	-0.210 **	-0.035
	(0.030)	(0.028)		(0.038)	(0.037)
Fighting	0.716 **	0.187 **	Fighting	0.162	0.133
	(0.049)	(0.055)		(0.086)	(0.072)
Truancy	0.083 **	-0.073 *	Truancy	0.086	0.125
	(0.029)	(0.035)		(0.073)	(0.083)

Notes: The overall effects come from Table 3, Columbs 2 and 3, which both control for the type of infraction.
**p<.01, *p<.05
Dependent variable = count of number of days of punishment.
Both sets of infraction-specific models control for race, special education status, grade, infraction order,
and type of consequence received.
The second column in each panel also includes school-fixed effects.

Appendix A: Robustness Checks

Table A1: OLS Regression of Days Punished on Race, Arkansas K-12 (2010-11 to 2012-13)

	Primary Models (From Table 3)		Robustness Check #1 (All except OSS and Expulsion = 0 Days)		Robustness Check #2 (Missing or Zero days of "Other" Imputed as 1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.329 **	0.072 **	0.174 **	0.057 **	0.274 **	0.055 **
	(0.018)	(0.021)	(0.013)	(0.015)	(0.016)	(0.019)
Hispanic	-0.145 **	0.012	0.020	0.010	-0.116 **	0.006
	(0.032)	(0.031)	(0.020)	(0.019)	(0.028)	(0.027)
Asian	-0.007	0.106	0.058	0.090	0.009	0.100
	(0.093)	(0.092)	(0.067)	(0.070)	(0.080)	(0.080)
Native American	-0.142 **	0.035	0.004	0.001	-0.098 *	0.036
	(0.052)	(0.048)	(0.031)	(0.031)	(0.046)	(0.044)
Two or More Races	-0.048	0.019	0.071 **	0.043	-0.041	-0.013
	(0.038)	(0.037)	(0.023)	(0.022)	(0.032)	(0.034)
SPED Status	-0.099 **	-0.099 **	-0.109 **	-0.108 **	-0.094 **	-0.095 **
	(0.023)	(0.020)	(0.014)	(0.014)	(0.021)	(0.019)
Expulsion	53.36 **	53.23 **	50.85 **	50.76 **	53.4 **	53.29 **
	(1.860)	(1.850)	(1.756)	(1.747)	(1.861)	(1.850)
Constant	2.078 **	2.334 **	0.078 **	0.164 **	2.033 **	2.266 **
	(0.040)	(0.043)	(0.030)	(0.034)	(0.037)	(0.040)
Grade Indicators	X	X	X	X	X	X
Infraction Type Indicators	X	X	X	X	X	X
Infraction Order Indicators	X	X	X	X	X	X
Consequence Indicators	X	X	X	X	X	X
School Fixed Effects		X		X		X
Observations	590,750	590,750	651,804	651,804	648,501	648,501
Adjusted R-Squared	0.3749	0.3994	0.4434	0.4531	0.3661	0.3884

Primary Model: dropped if missing days punished (except No Action = 0 and Corporal Punishment = 0). In this primary model, instances of missing or zero days for "other/non-specified infractions" were dropped.

Robustness Check #1: All consequences except OSS and Expulsion coded as zero days. Instances of missing or zero days for "other/non-specified infractions" were included as zero days.

Robustness Check #2: Days of punishment imputed = 1 if there was an instance of missing or zero days for an "other/non-specified" infraction.

Notes: Coefficients with robust standard errors, clustered at the student level in parentheses.

**p<.01, *p<.05

Dependent variable = count of number of days of punishment.

Baseline grade: 8th. Baseline infraction: disorderly conduct. Baseline consequence: ISS.

Infraction order indicators include: indicators for first, second, third, fourth, fifth, sixth, and seventh or more, plus an indicator for last infraction (baseline = 6th infraction).