Disparate Use of Exclusionary Discipline: Evidence on Inequities in School Discipline from a U.S. State

Kaitlin P. Anderson
&
Gary W. Ritter
University of Arkansas Fayetteville
United States

Abstract: There is much discussion in the United States about exclusionary discipline (suspensions and expulsions) in schools. According to a 2014 report from the U.S. Department of Education's Office for Civil Rights, Black students represent 15% of students, but 44% of students suspended more than once and 36% of expelled students. This analysis uses seven years of individual infraction-level data from public schools in Arkansas. We find that marginalized students are more likely to receive exclusionary discipline, even after controlling for the nature and number of disciplinary referrals, but that most of the differences occur across rather than within schools. Across the state, black students are about 2.4 times as likely to receive exclusionary discipline (conditional on reported infractions and other student characteristics) whereas within school, this same conditional disparity is not statistically significant. Within schools, the disproportionalities in exclusionary discipline are driven primarily by non-race factors such as free- and reduced-price lunch (FRL) eligibility and special education status. We find, not surprisingly, that schools with larger

proportions of non-White students tend to give out longer punishments, regardless of school income levels, measured by FRL rates. Combined, these results appear to indicate multiple tiers of disadvantage: race drives most of the disparities across schools, whereas within schools, FRL or special education status may matter more.

Keywords: discipline policy; school discipline; exclusionary discipline; race; disproportionalities

Uso desigual del castigo de exclusión: Evidencias sobre las desigualdades en castigo escolar de un estado de los EE.UU.

Resumen: Hay mucha discusión en los Estados Unidos sobre el castigo de exclusión (suspensiones y expulsiones) en las escuelas. De acuerdo con un informe de 2014 de la Oficina de Derechos Civiles del Departamento de Educación de los Estados Unidos, los estudiantes negros representan el 15% de los alumnos, pero el 44% de los estudiantes suspendidos más de una vez y el 36% de los estudiantes expulsados. Este análisis utiliza siete años de datos individuales de nivel de infracción de escuelas públicas en Arkansas. Hemos comprobado que los estudiantes marginados tienen más probabilidades de recibir un castigo de exclusión, incluso después de controlar la forma de ser y el número de referencias disciplinarias, pero que la mayoría de las diferencias se producen a través de las escuelas. En todo el estado, los alumnos negros tienen cerca de 2,4 veces más posibilidades de ser propensos a recibir castigo de exclusión (condicional en infracciones relatadas y otras características del alumno), mientras que dentro de la escuela, esta misma disparidad condicional no es significativamente significativa . En las escuelas, las desproporcionalidades en el castigo de exclusión son impulsadas principalmente por factores no raciales, tales como elegibilidad para almuerzo gratis y precio reducido (FRL) y status de educación especial. Hemos descubierto, no sorprendentemente, que las escuelas con mayor proporción de estudiantes no blancos tienden a dar castigos más largos, independientemente de los niveles de ingreso escolar, medidos por las tasas FRL. Combinados, estos resultados parecen indicar múltiples capas de desventajas: la raza impulsa la mayoría de las disparidades en las escuelas, mientras que en las escuelas, el FRL o el estado de educación especial pueden importar más.

Palabras-clave: política de castigo; castigo escolar; exclusion; raza

Utilização desigual da punição de exclusão: Evidências sobre as desigualdades em punição escolar de um estado dos EUA

Resumo: Há muita discussão nos Estados Unidos sobre a punição de exclusão (suspensões e expulsões) nas escolas. De acordo com um relatório de 2014 do Escritório de Direitos Civis do Departamento de Educação dos Estados Unidos, os estudantes negros representam 15% dos alunos, mas 44% dos estudantes suspensos mais de uma vez e 36% dos estudantes expulsos. Esta análise usa sete anos de dados individuais de nível de infração de escolas públicas no Arkansas. Verificamos que os estudantes marginalizados têm maior probabilidade de receber punição de exclusão, mesmo depois de controlar a maneira de ser e o número de referências disciplinares, mas que a maioria das diferenças ocorre através de do que dentro das escolas. Em todo o estado, alunos negros têm cerca de 2,4 vezes mais chances de ser propensos a receber punição de exclusão (condicional em infrações relatadas e outras características do aluno), enquanto que dentro da escola, esta mesma disparidade condicional não é significativamente significativa. Nas escolas, as desproporcionalidades na punição de exclusão são impulsionadas principalmente por fatores não-raciais, tais como elegibilidade para almoço grátis e preço reduzido (FRL) e status de educação especial. Descobrimos, não surpreendentemente, que as escolas com maior proporção de estudantes não-brancos tendem a dar punições mais longas, independentemente dos níveis de renda escolar, medidos pelas taxas FRL. Combinados, esses resultados parecem indicar múltiplas camadas de

desvantagens: a raça impulsiona a maioria das disparidades nas escolas, enquanto nas escolas, o *FRL* ou o status de educação especial podem importar mais.
Palavras-chave: política de punição; punição escolar; exclusão; raça

Background on Issues in School Discipline

Since the early 1990s, many schools across the United States have adopted zero-tolerance and other harsh disciplinary policies in response to fears of violence in schools. The zero-tolerance philosophy is an approach that removes students from school for a variety of violations, ranging from actual serious offenses like violent behavior to dress code violations or truancy (Losen & Skiba, 2010; Skiba, 2014; Skiba & Peterson, 1999). While it is necessary for school leaders to do what is reasonable to maintain a positive learning environment and ensure the safety of the school community, these so-called zero-tolerance policies have been opposed by a growing number of researchers and observers who fear that this movement has gone too far.

Opponents of harsh disciplinary practices have voiced numerous concerns. First, there is some evidence that these policies do not have the hoped-for deterrent effect. For example, Curran (2016) recently found that state zero-tolerance laws are not associated with decreases in problem behaviors as perceived by principals. In fact, there is evidence that school suspension predicts higher rates of misbehavior and suspensions in the future, rather than reducing misbehaviors (Costenbader & Markson, 1998; Raffaele-Mendez, 2003; Tobin, Sugai, & Colvin, 1996). Moreover, critics fear that zero tolerance might have other unintended negative consequences (Skiba, 2014). Zero-tolerance policies and exclusionary discipline practices, such as expulsions and suspensions, have been associated with lower academic achievement (Beck & Muschkin, 2012; Raffaele-Mendez, 2003; Raffaele-Mendez, Knoff, & Ferron, 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, & Letgers, 2003; Fabelo et al., 2011; Nicholson-Crotty, Birchmeier, & Valentine, 2009).

This active opposition to exclusionary discipline has made an impact and influenced some high-profile changes in school disciplinary practices. Chicago public schools enacted a policy in 2012 to reduce the length of student suspensions, and researchers from the Consortium on Chicago School Research have been analyzing the impacts (Sartain et al., 2015). In September 2014, California became the first state in the nation to enact limits of student suspension for minor misbehaviors (Public Counsel, 2014). One of the nation's largest school districts, Miami-Dade, also eliminated out-of-school suspensions (OSS) ahead of the 2015-16 school year (O'Connor, 2015). In Seattle, the School Board declared a one-year moratorium on suspensions for elementary students in September 2015 (Cornwell, 2015).

Perhaps a key reason that disciplinary policies have been revised is the concern that zero-tolerance policies and exclusionary practices have been applied disproportionately to students from marginalized backgrounds. A 2014 national report from the U.S. Department of Education's Office for Civil Rights focused on the racial disparity in exclusionary disciplinary policies. The authors reported that although Black students represent only 15% of students across the nation, 35% of students suspended once are Black, 44% of students suspended more than once are Black, and 36% of expelled students are Black. Indeed, over the past decade (and beyond), numerous researchers have documented differences in suspension rates between White students and non-White students across the nation (Children's Defense Fund, 1975; Costenbader & Markson, 1998; Losen & Gillespie, 2012; Losen, Hodson, Keith, Morrison, & Belway, 2015; Losen & Skiba, 2010; Raffaele-

Mendez, 2003; Skiba et al., 2014; Skiba et al., 2011; Skiba, Michael, Nardo, & Peterson, 2002; Sullivan, Klingbeil, & Van Norman, 2013). In addition, non-White students were more likely to receive suspensions for relatively subjective offenses, such as disrespect; the result is that non-Whites were disproportionately missing school time, often for non-violent or even trivial reasons (Skiba, Michael, Nardo, & Peterson, 2002).

In reaction to these circumstances, there is a growing, but still sparse, research base examining the racial disparities in the incidence of exclusionary discipline in schools across the country. Some studies rely on aggregate school- or district-level data and thus do not connect the actual student infractions to the disciplinary consequences (Children's Defense Fund, 1975; Losen, 2015; Losen & Skiba, 2010); such studies are informative but do not shed light on whether students are being treated unfairly. Others have utilized student-level data, but focus on disproportionalities in outcomes, without connecting them to the type or severity of infraction recorded (Raffaele-Mendez, 2003; Sullivan et al., 2013). Some more recent studies do utilize student-level or infraction-level datasets to address a more important issue: whether particular groups of students treated differently for committing the same type of infraction (Skiba et al., 2014; Skiba et al., 2011; Skiba et al., 2002). While several of these analyses move the knowledge base forward on the question, these studies are hampered by a variety of issues such as limited samples of students - one study (Skiba et al., 2002) involved only middle schools in a single district. In addition, certain studies do not incorporate school-level information or school fixed effects to assess whether disparities exist within certain types of schools (Skiba et al., 2011; Skiba et al., 2002).

Thus, in this article, we examine all disciplinary infractions and the resulting consequences for all K-12 students in a single U.S. state over a seven-year time period. We are able to connect individual student characteristics to specific infractions and to the resulting consequences. Using this rich dataset, we can carefully examine disparities in disciplinary outcomes by race and other student characteristics, while controlling for the infraction committed and for school attended. By identifying the extent to which students of different racial groups are punished more or less severely for the same offenses, even within the same schools, we hope to make a meaningful contribution to the growing evidence base on this important and timely issue.

The rest of this article proceeds as follows. In Section II, we present the literature on the topic of disparities in school discipline, and articulate our research questions. In Section III, we describe our data and sample. Section IV outlines our analytic methods, Section V presents the results, and in Section VI, we conclude with some discussion of our results.

Evidence from the Literature

We describe the relevant research in two sections. First, we present the evidence on the racial disparities in student discipline on a national level. Studies addressing this broad question generally rely on school-level data and provide only an overview of the consequences levied on students of different races. While these analyses are certainly important, they leave many questions unanswered because they do not examine the drivers of these differences. For example, if particular groups of students are punished more severely than others for serious but similar infractions, this is likely an indication of implicit or explicit bias in disciplinary practice at the school. Thus, the second set of studies we present are particularly informative as they investigate the student- and school-level characteristics that are associated with the racial disparities in discipline.

National Overviews of Disciplinary Disproportionalities

In 2015, Dan Losen and colleagues from the Civil Rights Project at UCLA, published a comprehensive report asking "Are We Closing the School Discipline Gap?" The authors focused on out-of-school suspension rates in every school district in the nation through the 2011-12 school year. The data revealed the overall increase in suspensions over the past 40 years, as well as the increasing gap in the suspension rates for White students and students of color. In 1972-73, only 6% of Black students were suspended during the year, 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%). Moreover, the authors also examined rates within states and districts and found much variability, indicating that district and school policies could strongly influence exclusionary discipline outcomes.

Several years earlier, Losen teamed with noted discipline researcher Russell Skiba on a national study of suspension rates in middle school, using an earlier 2006 version of the Civil Rights Data Collection (CRDC). In this study, the authors analyzed suspension rates for students in more than 9,200 middle schools across the nation, as well as a sub-sample from 18 large urban districts, from the years 2002 to 2006 (Losen & Skiba, 2010). This analysis also revealed stark racial gaps in suspensions; for example, 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. In the urban sub-sample district-level analysis, the authors found many schools in which more than one out of every three students in a particular racial group had been suspended during the year.

Overall, these and other analyses confirm that there are indeed systematic racial disparities in out-of-school suspensions. But, what factors drive these disparities? And do these differences persist even after controlling for infractions and referrals? In the next section, we summarize the emerging research literature addressing these questions. While we have not conducted a full systematic review of the literature, we searched thoroughly for literature on racial disparities in school discipline, with a focus on the use of exclusionary discipline, and used a snowball search to identify additional studies to include. We do not include theoretical or philosophical arguments for or against exclusionary discipline, but rather focus on studies that quantitatively assess the number of infractions or incidences of disciplinary consequences and the demographic characteristics of the students receiving these consequences. In general, we focus on articles since the year 2000.

Studies Examining the Drivers of Racial Discipline Gaps

In Chicago, where there has been a great deal of focus on exclusionary discipline in recent years, researchers from the Consortium on Chicago School Research scanned discipline data from roughly 85,000 high school students in the district in 2013-14 (Sartain et al., 2015). Using descriptive analyses, the authors have shown that Black students were three times as likely as Hispanic students to be suspended, and four times as likely as White and Asian students. While there was some evidence of students of different racial backgrounds systematically receiving more suspensions within the same schools, the primary driver of the differences was the school. That is, Black students attended schools, on average, that reported larger numbers of suspensions. While this investigation did consider some factors that play a role in the disparities, the authors were unable to account for the infractions allegedly committed by the students. Moreover, suspensions were the only consequence analyzed here. Nevertheless, this study moved the field forward by putting forth the idea that differing school environments or practices may be one driver of the racial discipline gap.

Welch and Payne (2010) further examined what drives the discipline gap by considering the "racial threat hypothesis" from criminal justice research. The authors posited that school leaders in

buildings serving more Black students would be more likely to use punitive discipline and less likely to use restorative approaches. Analyzing data from a 1998 nationally representative survey of students and school personnel in 294 public middle schools and high schools, the authors used multivariate regression techniques and found that principals in schools with higher proportions of Black students were more likely to report the use of punitive disciplinary styles. Next, the authors considered the influence of differential behavior by different groups of students by controlling for student reports of delinquency and teacher reports of school safety. This study suggests that students in schools serving high concentrations of Black students may well be subject to stricter discipline measures despite similarly safe and orderly environments. The weakness here, of course, is that the study is based on self-reports of disciplinary strategies rather than on actual disciplinary outcomes; moreover, the data are all school-level and do not indicate whether Black students themselves are punished more severely or more frequently.

The studies discussed up to this point do not provide much information related to the causes of the observed disproportionalities. The disproportionalities may be due to more frequent misbehavior by Black students or a greater willingness of school staff to refer these students to the office for subjective offenses. While many of the studies described in the previous section utilized student-level data, other researchers have advanced the field by using infraction- and referral-level data to further analyze the disciplinary outcomes for certain infraction types.

Russell Skiba and a variety of colleagues have published studies that assess the drivers of actual racial disparities in discipline. First of all, Skiba et al. (2002) used student-level data on more than 11,000 students from 19 middle schools in one of the largest U.S. school districts in 1994-95 to explore what factors drive discipline disproportionalities. While this analysis did not consider the variation in disciplinary strictness between schools, the authors did pay attention to infraction type and assessed whether differential bad behavior might play a role by analyzing the reasons for the disciplinary referrals. Specifically, the authors found that White students were more likely to be referred to the office for objective infractions such as smoking or vandalism while Black students were more likely to be referred for more subjective offenses such as disrespect and noise. Thus, the authors concluded that Black students were not more "disruptive," but they have also shown that the disproportionalities were indeed due to greater numbers of office referrals rather than greater severity of punishment (race had no impact on the length of punishment, given the referral).

Skiba et al. (2011) investigated the issue more deeply using student-infraction-level data from 364 elementary and middle schools across the United States using School-wide Positive Behavior Supports in 2005-06. Using logistic regression and multinomial logistic regression, the authors found that (1) Black students are more likely than White students to be referred to the office for a large variety of disciplinary infractions, and that (2) for the same referred infractions, Black students in all grades were significantly more likely to be given out-of-school suspension or even expulsion. Thus, even after accounting for the reported infraction type, Black students were more likely to be given exclusionary discipline. The only gap in this analysis is that there is no control for school effects; so, we do not know if the disparate strictness is occurring within school or between schools.

Next, Skiba et al. (2014) used Hierarchical Linear Modeling to predict punishment as a function of infraction type and incorporated a third level to the model by incorporating school characteristics. Using information from all students in the disciplinary database in a single Midwestern state in 2007-08, the authors found that the odds of being suspended or expelled were predictably influenced by the severity of the infraction. Importantly, even after controlling for the infraction, Black students remained more likely to be given out-of-school suspensions, but were no more likely to be expelled. This analysis extends beyond the prior work due to the inclusion of level three, in which school-level characteristics, such as student race and poverty and the principal's

attitude toward discipline, are incorporated into the model. In this third level analysis, the race of the individual student was no longer significant; school-level variables, including the concentration of Black students in the school, drove the severity of the punishments allocated. Thus, these results are consistent with the "racial threat hypothesis" in schools suggested by Welch and Payne (2010). One potential weakness of the Skiba et al. (2014) study is the setting and context – the data represent a single year in a single U.S. state that serves relatively few FRL-eligible students (fewer than 40%) and very few Black students (8%). Our current study expands on this work by incorporating seven years of student-level panel data in a state that contains a more diverse population (about 21% Black and about 10% Hispanic).

Overall, a vast amount of evidence indicates there are racial disparities with respect to exclusionary discipline outcomes. Indeed, the Office for Civil Rights has recently demonstrated nationwide racial disparities in rates of suspensions and expulsions, and moreover, a couple of recent studies have concluded that Black students have been given disproportionate consequences for the infraction committed. However, it is still not clear whether in most cases, this disparity is due to students being treated differently in the same school or to the fact that Black students attend systematically different schools where the disciplinary practices are abnormally strict. To date, the most thorough analysis conducted to assess the extent to which non-White students have been more severely punished for similar disciplinary referrals, considering also whether these disparities occur within schools and across schools, has been published by Skiba et al. (2014).

Thus, although questions surrounding the sources of disciplinary disparities are critically important, the best evidence to date comes from a single school year in a single state serving relatively few FRL-eligible and Black students. Therefore, we believe it is valuable to conduct such analyses in additional settings, ideally with greater levels of student diversity and a longer study period. Our current study expands on previous work by accounting for specific infraction information (type, frequency, etc.) and school-level fixed effects whenever possible, using multiple years of data within a single U.S. state serving a student population that is approximately 60% low-income (FRL-eligible) and 20% Black.

Research Questions Guiding this Study

First, across the state, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White, low-income students, special education students, or English language learners?

Second, within schools, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White students?

Finally, what are the school characteristics that are associated with harsher (longer) disciplinary consequences?

Data and SampleArkansas Student Sample

First, it is important to note how closely the patterns in the Arkansas data utilized in this study mirror the OCR data mentioned previously. In Table 1, we calculate the percentage of students in various subgroups, the percentage of students who received OSS at least once who were in these subgroups, and the percentage of students who were expelled in these subgroups.[1] The odds of a student in a given subgroup being in a consequence category (e.g. expelled) is the percentage of

---

[1] We report differences between Black and White students, Hispanic and White students, ELL and non-ELL students, special education and regular education students. The Office for Civil Rights does not report disciplinary rates for FRL and non-FRL students separately, and the Arkansas dataset we used did not include gender, so those differences are not reported here.

expelled students in that group divided by the percentage of total students in that group. For example, White students represent 65% of students in the state of Arkansas, and 38% of students receiving OSS, so the odds ratio is equal to (0.38/0.65) or approximately 0.58. Odds of less than one indicate that a certain group is underrepresented in a certain category, relative to their prevalence in the state, and odds of greater than one indicate that a certain group is overrepresented in a certain category.

Then, we calculate disparities (relative odds) between groups, which can be compared across different subgroups. In terms of the disparities for Black students, relative to White students, the Arkansas disparities are larger than the nationwide disparities for OSS, but smaller for expulsions. We can also see that overall, the Black-White disparities are much larger than any other disparities, including those for special education students relative to non-special education students. Interestingly, in both the OCR data (nationally) and in the Arkansas data, based on the odds, Hispanic students and English language learners are somewhat underrepresented in these types of exclusionary discipline practices. However, when we compare the relative odds of Hispanic students to White students, there are still disparities, at least in the OCR data. Arkansas Hispanic students are under-represented, even relative to White students, at least in terms of expulsions. While Arkansas is only one of 50 states in the United States, the similarities in these patterns (particularly with regard to the Black-White disparity) indicate that many of the findings of the current study may be relevant for many other parts of the nation as well.

Data and Descriptive Statistics

The study uses seven years of de-identified demographic and disciplinary data from all K-12 schools in Arkansas provided by the Arkansas Department of Education (2008-09 through 2014-15). The student demographic data include race, grade, special education status, limited English proficiency status, and free-and-reduced-lunch (FRL) eligibility. Discipline data include indicators for 19 infraction types and 13 consequences, the date of the infraction, and the length of the consequence. To simplify the analysis, we have collapsed infractions involving handguns, rifles, and shotguns into a single category, resulting in only 17 distinct categories. Furthermore, 13 consequence categories are collapsed into 7 (in school suspension (ISS), OSS, expulsion, referral to an alternative learning environment (ALE), corporal punishment, no action, and other).[2] An alternative learning environment is an "an alternate class or program within a public school or school district that affords all students an environment that seeks to eliminate barriers to learning for any student whose academic and social progress is negatively affected by the student's personal characteristics or situation" (AR Code § 6-48-104, 2015).

---

[2]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.

Table 1

Comparison of Arkansas Disciplinary Data and Office for Civil Rights National Data

		White		Black		
		% of Group	Odds	% of Group	Odds	Disparity
Arkansas (2008-09 to 2014-15)	% Enrollment	65%		21%		
	% Receiving OSS	38%	0.58	54%	2.53	4.32
	% Expelled	48%	0.75	44%	2.06	2.75
OCR (2011-12)*	% Enrollment	52%		16%		
	% Receiving OSS	35%	0.67	38%	2.38	3.56
	% Expelled	36%	0.70	36%	2.25	3.20
		White		Hispanic		
		% of Group	Odds	% of Group	Odds	Disparity
Arkansas (2008-09 to 2014-15)	% Enrollment	65%		10%		
	% Receiving OSS	38%	0.58	6%	0.61	1.05
	% Expelled	48%	0.75	6%	0.57	0.76
OCR (2011-12)*	% Enrollment	52%		24%		
	% Receiving OSS	35%	0.67	22%	0.91	1.37
	% Expelled	36%	0.70	22%	0.90	1.28
		Non-English Language Learner (Non-ELL)		English Language Learner (ELL)		
		% of Group	Odds	% of Group	Odds	Disparity
Arkansas (2008-09 to 2014-15)	% Enrollment	93%		7%		
	% Receiving OSS	96%	1.03	4%	0.57	0.55
	% Expelled	96%	1.03	4%	0.55	0.53
OCR (2011-12)*	% Enrollment	90%		10%		
	% Receiving OSS	94%	1.04	6%	0.60	0.57
	% Expelled	95%	1.06	5%	0.50	0.47

Table 1 (Cont'd)

Comparison of Arkansas Disciplinary Data and Office for Civil Rights National Data

		Regular Education		Special Education		
		% of Group	Odds	% of Group	Odds	Disparity
Arkansas (2008-09 to 2014-15)	% Enrollment	89%		11%		
	% Receiving OSS	81%	0.91	19%	1.69	1.85
	% Expelled	81%	0.91	19%	1.76	1.94
OCR (2011-12)*	% Enrollment	88%		12%		
	% Receiving OSS	78%	0.89	22%	1.83	2.07
	% Expelled	81%	0.92	19%	1.58	1.72

		Non-FRL-Eligible		FRL Eligible		
		% of Group	Odds	% of Group	Odds	Disparity
Arkansas (2008-09 to 2014-15)	% Enrollment	40%		60%		
	% Receiving OSS	21%	0.52	79%	1.32	2.52
	% Expelled	21%	0.53	79%	1.31	2.48
OCR (2011-12)*	% Enrollment	N/A		N/A		
	% Receiving OSS	N/A	N/A	N/A	N/A	N/A
	% Expelled	N/A	N/A	N/A	N/A	N/A

Note: All percentages reflect the number of students receiving either OSS or expulsion at least once in a school year that are within a certain subgroup. Odds for white students are the percent of students suspended or expelled divided by the percent of enrollment. Values over one indicate over-representation. Disparities (relative odds) are calculated as the odds for one group divided by the odds for another group. These indicate whether a subgroup is over-represented relative to another subgroup. Values greater than one indicate over-representation. Values less than one indicate under-representation. Office for Civil Rights race breakdowns reflect the race/ethnic composition of students without disabilities and students with disabilities served under IDEA. Special-education students only include those with an IEP, under the IDEA. Does not include handicapped students under Section 504. The Office for Civil Rights does not report disciplinary rates for FRL and non-FRL students separately.

Table 2

Infraction Types, By Year from 2008-09 to 2014-15 (Arkansas)

	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	Total	% of Total
Disorderly Conduct	54,641	51,027	48,765	51,539	42,575	57,750	63,533	369,830	29.7%
Other	31,871	28,639	26,481	31,858	35,024	60,600	95,733	310,206	24.9%
Insubordination	47,273	46,151	45,765	38,798	34,759	43,068	51,200	307,014	24.7%
Fighting	12,378	12,456	12,471	12,136	12,434	13,128	14,576	89,579	7.2%
Truancy	9,968	11,834	11,734	10,465	9,407	12,914	14,987	81,309	6.5%
Bullying	3,455	4,099	4,363	4,483	4,515	5,496	5,856	32,267	2.6%
Tobacco	2,218	2,253	1,973	1,920	1,977	2,482	2,837	15,660	1.3%
Student Assault	1,856	1,820	1,615	1,645	2,007	2,153	2,232	13,328	1.1%
Drugs	944	996	954	1,146	1,259	1,295	1,511	8,105	0.7%
Vandalism	962	833	909	689	736	1,084	1,087	6,300	0.5%
Knife	401	419	384	396	443	532	497	3,072	0.2%
Staff Assault	292	312	277	314	354	350	487	2,386	0.2%
Alcohol	294	299	325	289	309	353	416	2,285	0.2%
Gangs	361	339	177	107	131	103	113	1,331	0.1%
Explosives	49	57	60	50	42	53	40	351	0.0%
Club	21	21	49	45	42	53	57	288	0.0%
Guns	38	18	32	26	35	33	62	244	0.0%
Total	167,022	161,573	156,334	155,906	146,049	201,447	255,224	1,243,555	100.0%
% of Total	13.4%	13.0%	12.6%	12.5%	11.7%	16.2%	20.5%	100.0%	

The unit of analysis is the student-infraction level, so students can and often do have multiple observations within the same year. After removing duplicate entries (same student, discipline date, infraction type, consequence type, etc.), 1,243,555 total observations remain over the seven-year period. These observations were recorded for 240,999 individual students, which would represent about 35% of the individual students expected to attend Arkansas schools during this time period (thus, the other 65% of students in the state's public schools received no disciplinary referrals or consequences during this time period.) The breakdown by infraction and consequence, by year, can be seen in Tables 2 and 3. The vast majority of infractions (79.4%) are relatively subjective consequences such as disorderly conduct (29.7%), other infractions not specified in these categories (24.9%), and insubordination (24.7%).

Table 3 shows the trends in the reported types of disciplinary consequences. We include out-of-school suspensions, expulsions, and referrals to an Alternative Learning Environment as exclusionary, given that they remove a student from the traditional learning environment, and in the case of expulsions and ALE, for long periods of time. In-school suspensions are considered non-exclusionary as the student remains in the school building, continues to receive assignments from their regularly assigned teacher, and then returns to the same classroom after a relatively short suspension (generally one to two days). The trend over time has been a decrease in exclusionary discipline as a proportion of total infractions (about 25% all disciplinary consequences were exclusionary in 2008-09 compared to only about 19% in 2014-15), but much of this is due to large increases in the use of other non-specified consequences. While we have concerns about the uncertainty within this other non-specified category, the vast majority of these other non-specified outcomes are non-exclusionary.[3] Expulsions and no actions are consistently rare, and ISS was the largest category in each year, until 2014-15, in which the other (non-specified) category was the most common. The number of incidences of the other (non-specified) consequence category grew by over 300% between 2008-09 and 2014-15.

To simplify interpretation of the infraction categories, we create categories based on the type and length of consequences typically received for each infraction type. We present the percentage of incidences that result in exclusionary discipline (expulsion, out-of-school suspension, or referral to an Alternative Learning Environment), as well as the number of days of suspension or expulsion that typically result. Table 4 indicates the descriptive statistics used in the creation of these categories. To group these infractions, we consider, simultaneously, the percentage of incidences of that infraction type that result in exclusionary discipline, as well as the typical number of days of exclusionary discipline that results. At the same time, we consider infractions that are similar in nature (for example, substances that are not only illegal to have at school, but also illegal for even 18 year olds, such as drugs and alcohol, are somewhat different than tobacco and thus grouped as such). Further, the cut-off between "major" and "minor" non-violent offenses, for example, is primarily based on the likelihood of each offense resulting in exclusionary discipline; there is a break in the pattern where the three "major" non-violent offenses result in the student being excluded approximately 30% of the time while "minor" offenses lead to exclusionary discipline around 20% of the time.

---

[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.

J

Table 3
Consequence Types, By Year from 2008-09 to 2014-15 (Arkansas)

	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14	2014-15	Total	% of Total
Exlusionary Discipline									
Out-of-School Suspension	41,348	39,613	36,780	37,791	40,233	42,290	47,853	285,908	23.0%
ALE	918	794	621	253	317	586	538	4,027	0.3%
Expulsion	135	322	193	95	200	249	165	1,359	0.1%
Total Exclusionary	42,401	40,729	37,594	38,139	40,750	43,125	48,556	291,294	23.4%
% of Annual Total	25.4%	25.2%	24.0%	24.5%	27.9%	21.4%	19.0%	23.4%	23.4%
Non-Exclusionary Discipline									
In-School Suspension	63,018	64,760	60,052	62,532	63,019	74,169	92,084	479,634	38.6%
Other	23,120	23,858	27,600	26,482	21,850	62,972	92,865	278,747	22.4%
Corporal Punishment	36,484	30,732	29,311	27,760	19,142	19,746	19,571	182,746	14.7%
No Action	1,999	1,494	1,777	993	1,288	1,435	2,148	11,134	0.9%
Total Non-Exclusionary	124,621	120,844	118,740	117,767	105,299	158,322	206,668	952,261	76.6%
% of Annual Total	74.6%	74.8%	76.0%	75.5%	72.1%	78.6%	81.0%	76.6%	76.6%
Total	167,022	161,573	156,334	155,906	146,049	201,447	255,224	1,243,555	100.0%
% of 7 Year Total	13.4%	13.0%	12.6%	12.5%	11.7%	16.2%	20.5%	100.0%	

Interestingly, in Table 4, we see that exclusionary discipline is not even used in all gun infractions. While expulsion is allowed for any student who brings a firearm or other weapon to school, the superintendent also has discretion to modify this requirement on a case-by-case basis, which appears to be happening in over 20% of gun cases (AR Code § 6-18-507, 2015). The ability for school district leaders to adjust consequences on a case-by-case basis is perhaps further evidence that there are opportunities for disproportionalities in discipline outcomes to occur, even for infraction types in which we expect near universal exclusion.

The state only codes certain types of infractions and consequences, so some categories used at a local level are coded as "other" at the state level. As a result, a large number of cases can be coded as "other" in either the infraction committed, the consequence received, or both. In the next section, we describe the analytic methods we employ to analyze these data and examine any possible disparities in disciplinary practices.

Table 4
Category Groups (Based on Percent Exclusionary and Number of Days of Exclusion)

	% Resulting in Exclusionary Discipline	Typical Number of Days of Exclusion
Guns	77.5	11.8
Drugs and Alcohol	87.8	8.8
Drugs	88.2	9.0
Alcohol	86.4	8.0
Major Violence/Weapons	75.1	5.2
Club	83.0	4.0
Knife	74.9	5.8
Staff Assault	74.4	4.7
Minor Violence/Weapons	59.3	3.6
Gangs	63.6	5.4
Fighting	60.8	3.5
Student Assault	49.2	3.9
Explosives	47.6	4.5
Major Non-Violent	30.3	3.1
Tobacco	35.4	3.3
Vandalism	32.1	4.1
Bullying	27.5	2.8
Minor Non-Violent	19.2	3.2
Disorderly Conduct	20.4	3.6
Insubordination	18.7	2.7
Other	18.2	3.2
Truancy	12.0	2.9

Analytic Method

In our straightforward descriptive analyses presented in the previous section, we have described how frequently students of various subgroups are cited for various types of infractions as well as how frequently students in these subgroups receive various types of consequences. Next, we use logistic regression and aggregated residual techniques to address our three primary research questions.

Research Question 1: Across the state, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White students, low-income students, special education students, or English language learners?

We begin by testing whether students of various subgroups are more or less likely to receive exclusionary discipline, controlling for the type of infraction committed. We first analyze these disparities at a state level. Any disparities we find at this level could be due to differences across districts or schools, within district, or within school. We utilize logistic regression to predict whether certain types of students are more likely to receive exclusionary discipline (expulsion, OSS, or referral to an ALE), rather than another consequence (ISS, corporal punishment, no action, or other). Whether or not a student receives exclusionary discipline (E_1) is defined as:

$$E_1 = \left\{ \begin{array}{l} 1 \; if \; E_1^* > 0 \\ \\ 0 \; if \; E_1^* \leq 0 \end{array} \right.$$

$$E_1^* = \beta_0 + \beta_1 V_{is} + \beta_2 InfCat_i + \beta_3 InfOrder_i + \beta_4 Grade_i + \beta_5 Year_i + \varepsilon_i$$

Where V_i is a vector of the student-level demographic indicators (some combination of race, FRL-eligibility, special education status and LEP-status), $InfCat_i$ is a vector of 7 infraction categories, grouped by severity, $InfOrder_i$ is a vector of indicators for whether the infraction was the first, second, third, etc., for that student that year (a total of 10 indicators for 1-9 and 10 or more), $Year_i$ is a vector of school-year indicators, and ε_i is the infraction-level idiosyncratic error (clustered at the student level).

In this first analysis, no school-level indicators or covariates are included, so it is considered a model of state-wide racial or other disparities in disciplinary outcomes, conditional on similar infraction types, infraction history, and grade level.

Research Question 2: Within schools, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White students, low-income students, special education students, or English language learners?

Next, we seek to understand the disparities within schools, rather than across schools. We utilize similar logistic regression as in Research Question 1, but with the addition of school fixed effects. This within-school analytic strategy is motivated by work by Anderson & Ritter (forthcoming) who find that most of the disparities in the length of punishments (e.g. number of days of suspensions) at the state level diminishes when school fixed effects are included, indicating that most of the disparities are across schools rather than within schools. If, in the current study, the disparities diminish when school fixed effects are included in our models, this would indicate that a

great deal of the variation exists between schools. Thus, we also ask question three which seeks to disentangle the particular school characteristics driving these differences.

Research Question 3: What are the school characteristics that are associated with harsher (longer) disciplinary consequences?

To address whether certain types of schools are more likely to assign disproportionately long punishments for similar types of infractions, we use a two-stage residuals analysis approach. In the first stage, we predict the number of days of exclusionary discipline as a function of information related to the reported infraction that could reasonably predict the type or length of consequence received, as well as the cumulative number of reported infractions associated with that student during the same school year. In this first stage, we do not include any student demographic information other than grade level, which could be associated with the type or severity of consequence used. Our first stage model utilizes ordinary least squares regression, with heteroskedastic-robust standard errors clustered at the student level (Angrist & Pischke, 2009; Huber, 1967; Rogers, 1993; White, 1980). The first stage model is:

$$DaysPunished_i = \beta_0 + \beta_1 InfCat_i + \beta_2 InfOrder_i + \beta_3 Year_i + \beta_4 GradeLevel_i + \varepsilon_i$$

Where i indexes at the infraction level, $DaysPunished_i$ is the total number of days of punishment, $InfCat_i$ is a vector of infraction categories, which can be defined two ways (using all 17 categories, or our 7 infraction types, grouped generally by severity), $InfOrder_i$ is a vector of indicators for whether the infraction was the first, second, third, etc., for that student that year (a total of 10 indicators for 1-9 and 10 or more), $Year_i$ is a vector of school-year indicators, $GradeLevel_i$ is a vector of grade-level indicators, and ε_i is the infraction-level idiosyncratic error (clustered at the student level). In our primary model, we focus on days of exclusionary discipline (expulsion, OSS, or referral to an ALE) associated with a given infraction, with all other consequence types coded as zero days[4].

These residuals generated by the OLS model are then averaged at a school-by-year level to produce a measure of whether a school, on average, meted out longer punishments (residuals greater than 0) or shorter punishments (residuals less than 0), relative to the state average, for a similar type of infraction for a student in the same grade with a similar number of past disciplinary infractions. We refer to this average school-level residual as the School Severity Index (SSI). The school-by-year SSI values are estimated using a school-level random effects model, which shrinks the estimates towards zero for schools with relatively few observations. Schools with positive SSI values tend to give out longer punishments, and schools with negative SSI values tend to give out shorter punishments, relative to the state average for similar infraction observations.

In the second stage, we predict the SSI as a function of school-level demographic characteristics to assess which school characteristics are associated with disciplinary practices:

$$SSI_s = \beta_0 + \beta_1 X_s + \beta_2 Year_s + \varepsilon_s$$

Where s indexes at the school level, X_s is a vector of school level characteristics (log of enrollment, an indicator for region, an indicator for open-enrollment charter schools, indicators for elementary, middle, high school, or other school grade-level types, and the percentage of the student

---

[4] Days of exclusion were at most 365 days. The average expulsion was 18.4 days, the average OSS was 3.3 days, and the average ALE was 10.6 days.

population that is FRL-eligible, of a certain race, receiving special education services, limited English proficient (LEP), or gifted and talented), $Year_s$ is a vector of school-year indicators, and ε_s is the school-level idiosyncratic error.

Next, we present our findings, beginning with some brief descriptive statistics and ultimately walking through the results of each of three research questions.

Results

Initial descriptive analyses focused on the frequency of both consequence types and infractions for different subgroups of students. In Figure 1, it is easy to see that non-White students are disproportionately receiving all types of consequences. On average, each year, there are 29.6 in-school suspensions for every 100 Black students, but only 9.9 in-school suspensions for every 100 White students. Each year, there are 24.6 out-of-school suspensions for every 100 Black students, but only 4.3 for every 100 White students. Thus, a ratio-based measure of the Black-White disparity in ISS indicates that Black students are about three times as likely to receive OSS as White students (29.6 divided by 9.9). For other consequence types such as referrals to ALE, this ratio is about 9.5 times, or for OSS, 5.7 times.
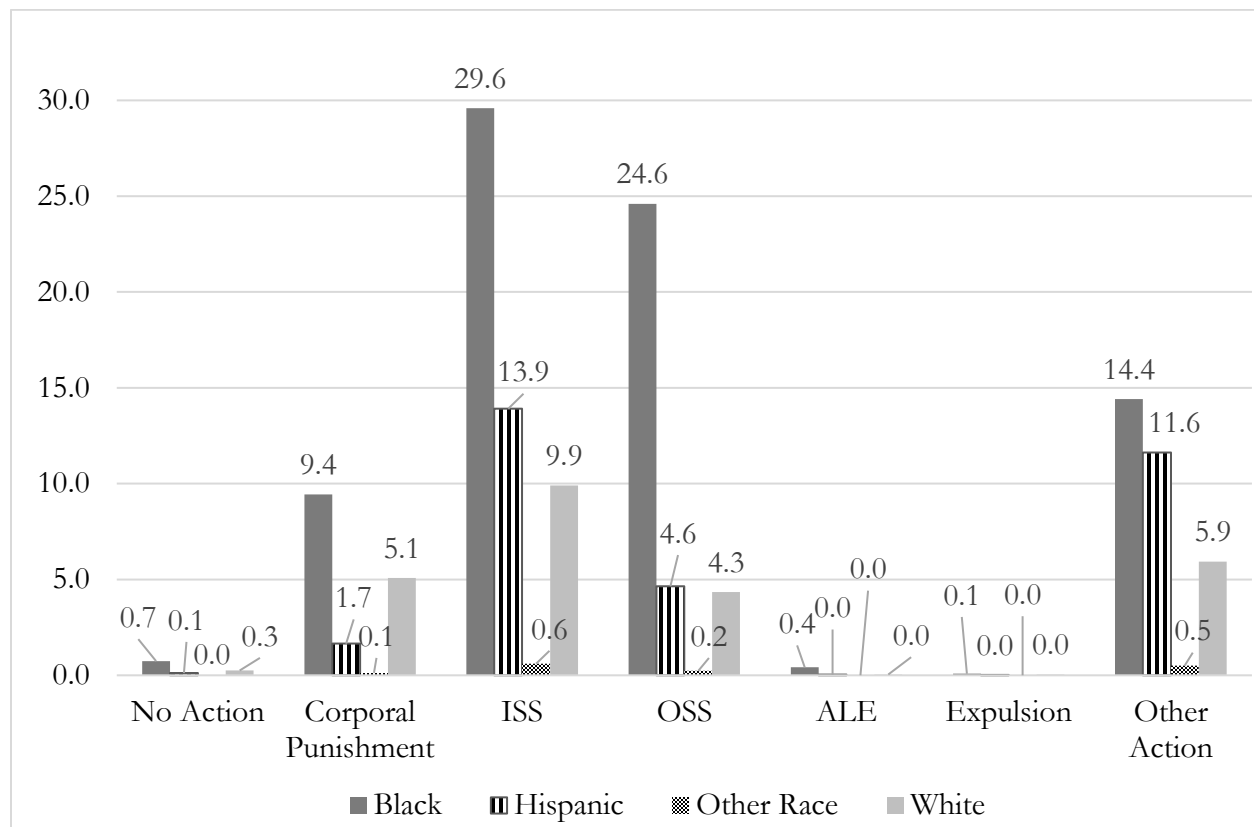
Figure 1. Disciplinary Consequences by Racial Subgroup (Annual Incidences per 100 Students, 2008-09 to 2014-15)

Looking just at the disparities in Figures 1, one might come to the quick conclusion that students are being treated unfairly, but it is also important to connect consequences to the infractions for which the students were referred. Thus, we next consider whether there are disparate rates of referrals for certain types of infractions, and indeed, we see that there are

disproportionalities at this level. This does not, however, rule out the possibility that disparities may still exist *within* each infraction type, which we address with Research Questions 1 and 2.

First, a key take-away point from Figure 2 is that the vast majority (almost 80%) of incidences are minor, non-violent offenses (disorderly conduct, insubordination, and other). A second point is that Black students are three times more likely than White students to be referred for misbehavior but are nearly six times more likely to be given out-of-school suspensions (24.6 versus 4.3 incidences per 100 students, in Figure 1). These data do indicate that Black students are being referred for discipline more often, but this only accounts for about half the difference in the rate of out-of-school suspensions. Our analyses in the next section, using logistic regression to examine incident-level data, helps us to identify more clearly whether there are disparities that still exist conditional on students' reported behavior.
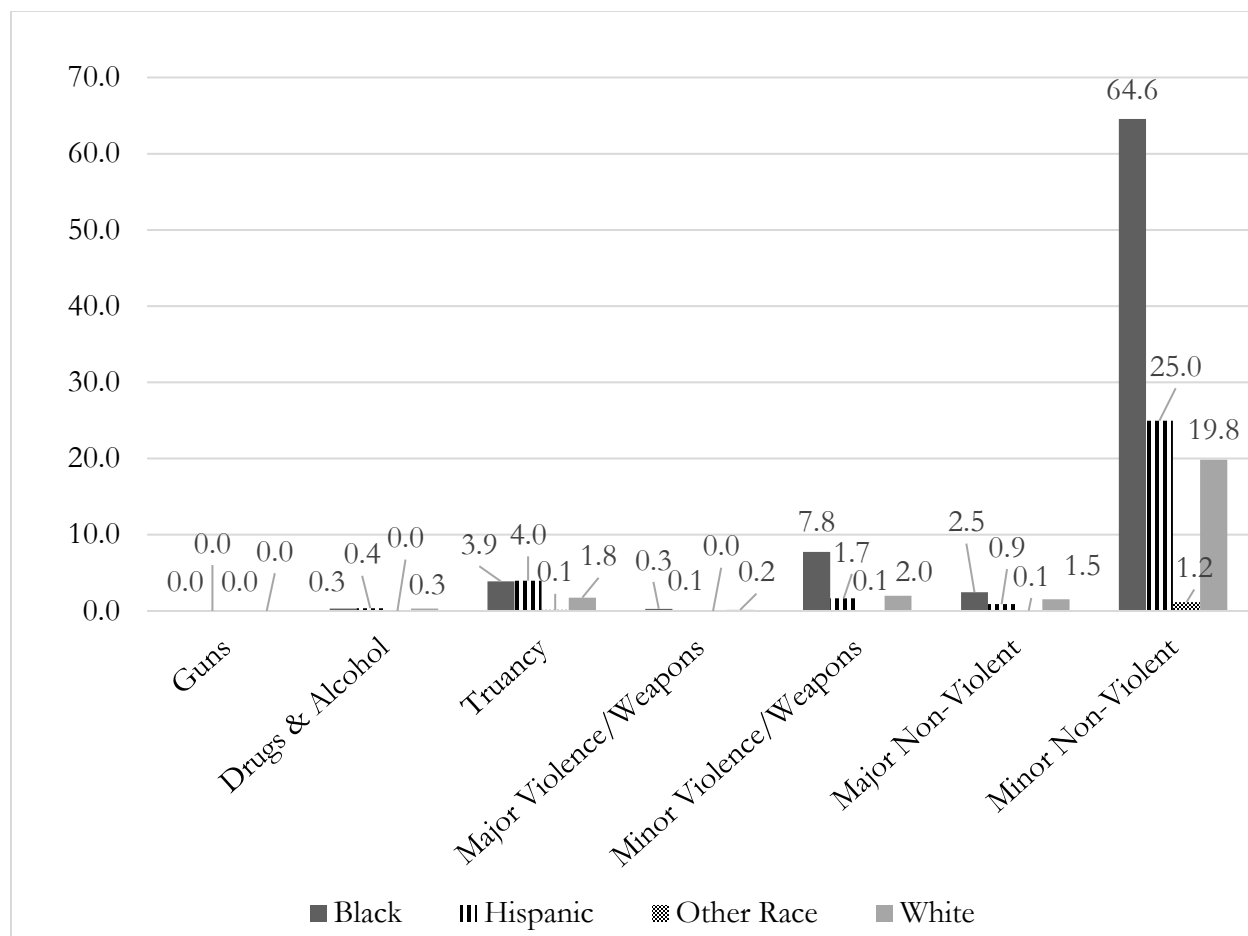
Figure 2. Referrals by Infraction Type and Race (Annual Incidences per 100 Students, 2008-09 to 2014-15)

Research Question 1: Across the state, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White students?

Logistic regression is used to determine the disparities in the likelihood of exclusionary discipline, controlling for the type of infraction committed, the infraction history of the student, and the student's grade level. No school-level factors are taken into account, so this model indicates the

extent to which different subgroups of students across the state are disproportionately exposed to exclusionary practices. Any differences by subgroup we find at this level could be due to differences at a variety of levels (across districts or schools, within district, or within school).

Table 5
Logistic Regression of Exclusionary Discipline (Arkansas State, 2008-09 to 2014-15)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	2.215***		2.132***	2.471***				2.378***
	(0.0206)		(0.0202)	(0.0238)				(0.0233)
Hispanic	0.795***		0.838***	0.888***				0.897***
	(0.0159)		(0.0217)	(0.0173)				(0.0231)
Other Minority	0.854***		0.878***	0.912**				0.920**
	(0.0329)		(0.0346)	(0.0345)				(0.0356)
FRL-Eligible		1.475***	1.224***		1.518***			1.232***
		(0.0135)	(0.0115)		(0.0145)			(0.0120)
Special Education			1.106***			1.068***		1.090***
			(0.0130)			(0.0126)		(0.0129)
LEP			0.860***				0.534***	0.922**
			(0.0292)				(0.0134)	(0.0307)
Grade Indicators	Y	Y	Y	Y	Y	Y	Y	Y
School Year Indicators	Y	Y	Y	Y	Y	Y	Y	Y
Infraction Types				Y	Y	Y	Y	Y
Infraction Order Indicators				Y	Y	Y	Y	Y
Constant	0.351***	0.363***	0.297***	0.226***	0.240***	0.347***	0.348***	0.191***
	(0.0720)	(0.0732)	(0.0611)	(0.0498)	(0.0526)	(0.0752)	(0.0753)	(0.0423)
Wald Chi-Squared	11,571	5,645	12,146	76,215	73,372	2,321	71,778	76,398
Pseudo R-Squared	0.036	0.012	0.037	0.118	0.092	0.089	0.091	0.119

Note: Heteroskedastic-robust standard errors in parentheses, clustered at the student level; * *p*<0.1, ** *p*<0.05, *** *p*<0.01. In all models, number of observations = 1,243,555, number of students = 240,999. Infraction types are indicators for different groups of infractions: guns, drugs and alcohol, truancy, major violence/weapons, minor violence/weapons, and major non-violent. Baseline infraction category is Minor Non-Violent Infractions.

Relative risk ratios from several logistic regressions are indicated in Table 5. It is important to note that all models in Table 5, using infraction-level disciplinary data, are conditional on a student being referred for some infraction, so even without controlling for infraction, we can see that, holding constant that a student was referred for any misbehavior, we get a better picture of disciplinary disparities than with just comparing raw numbers of suspensions and expulsions as in Figures 1 and 2. In columns 1-3 of Table 5, we present the results of relatively naïve models that are contingent only upon the student being referred for some disciplinary infraction, but without accounting for infraction type. The primary results, based on models in which we control for the type of infraction committed and for the number of infractions committed by the student during the school year, are presented in columns 4-8.

If disciplinary consequences were handed out evenly across various subgroups of students, we would expect to see relative risk ratios for each indicator (e.g. Black) equal to one. The results in

Table 5, column 4 indicate that Black students are almost 2.5 times as likely to receive exclusionary discipline as their White peers in the same grade for similar types of infractions, with a similar number of previous infractions that year. Hispanic students as well as other non-White students are somewhat less likely than their White peers to receive exclusionary discipline. In fact, these figures may represent a lower bound of the actual racial disparities in punishments, as they focus only on the instances of students being assigned to exclusionary discipline, but ignore possibly larger disparities in the number of days punished per suspension.

Looking at columns 5, 6, and 7 of Table 5, instead of testing disproportionalities in exclusionary discipline using race indicators, we use other indicators of a student's disadvantaged status (FRL-eligibility, Special Education status, or English proficiency). FRL-elgible students are about 1.5 times as likely as their non FRL-eligible peers in the state to receive exclusionary discipline. Special education students are slightly more likely to receive exclusionary discipline, and students with Limited English Proficiency are about half as likely to receive exclusionary discipline. The model in Column 8 includes the full combination of control variables.

It is interesting that the racial disparities, indicated by the relative risk ratios on Black, Hispanic, and other non-White groups, are quite similar between columns 1 and 4 and columns 3 and 8. The disparities based on FRL-status, indicated by the relative risk ratios on FRL-Eligible are also similar between columns 2 and 5. This result indicates that that the racial disparities in the use of exclusionary discipline are not driven solely by the types of infractions students committed. While the specific type of infraction, controlled for in columns 4-8, does help to explain whether a student receives exclusionary disipline (higher pseudo R-squared), the relative risk ratios for various racial subgroups do not decline much with the inclusion of infraction-type controls. The stability of these results gives us some confidence in our key finding, that Black students are more than twice as likely to receive exclusionary discipline after being referred for identical infractions. While the pseudo R-squared values are somewhat low, they are not equivalent to the R-squared found in OLS regression, so while higher values represent better model fit, they cannot be interpreted exactly the same as the R-squared found in OLS regression (Institute for Digital Research and Education, 2017).[5]

The results for research question 1, discussed previously, are only representative of disparities in disciplinary outcomes across the state. It could be that most of these disparities only occur across schools, or it could be, instead, that disparities also exist within schools. In the next section, we utilize school fixed effects to assess what disproportionalties exist, if any, in disciplinary outcomes for students within the same schools.

Research Question 2: Within schools, what, if any, disproportionalities exist in the use of exclusionary discipline for non-White students, low-income students, special education students, or English language learners?

In this section, logistic regression is again used to assess whether student demographic factors are associated with higher rates of exclusionary discipline, this time for students within the same schools. Relative risk ratios from several logistic regressions, all including school fixed effects, are indicated in Table 6. The results in column 1 indicate that Black students are only slighly more likely to receive exclusionary discipline, relative to their White peers within the same schools. Larger disparirities can be seen based on whether the student is FRL-eligible (column 2) or receiving Special Education services (column 3). This result indicates, perhaps, the multiple tiers of privilege or disadvantage – that Black students are disproportionately exposed to exclusionary discipline as a

---

[5] Stata's default pseudo R-squared value is McFadden's R-squared, which is equivalent to 1 minus the ratio of the log likelihood of the full model to the log likelihood of a simple intercept model (Institution for Digital Research and Education, 2011).

function of the school that they attend, but that within schools, other factors such as poverty or special education status influence the likelihood of a student receiving exclusionary consequences.

Table 6
Logistic Regression of Exclusionary Discipline within Schools (Arkansas, 2008-09 to 2014-15)

	(1)	(2)	(3)	(4)	(5)
Black	1.035***				1.007
	(0.0109)				(0.0108)
Hispanic	0.935***				0.949**
	(0.0156)				(0.0205)
Other Minority	1.011				1.023
	(0.0325)				(0.0332)
FRL-Eligibile		1.165***			1.157***
		(0.0104)			(0.0105)
Special Education			1.191***		1.180***
			(0.0115)		(0.0115)
Limited English Proficient				0.910***	0.935***
				(0.0180)	(0.0243)
Grade Level Indicators	Y	Y	Y	Y	Y
School Year Indicators	Y	Y	Y	Y	Y
Infraction Types	Y	Y	Y	Y	Y
Infraction Order Indicators	Y	Y	Y	Y	Y
School Fixed Effects	Y	Y	Y	Y	Y
Constant	0.0812***	0.0706***	0.0818***	0.0820***	0.0703***
	(0.0234)	(0.0205)	(0.0239)	(0.0237)	(0.0205)
Observations	1,236,401	1,236,401	1,236,401	1,236,401	1,236,401
Number of Students	239,202	239,202	239,202	239,202	239,202
Model Chi-Squared	132,531	132,473	131,941	132,507	132,333
Pseudo R^2	0.324	0.325	0.325	0.324	0.325

Note: Heteroskedastic-robust standard errors in parentheses, clustered at the student level; * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Infraction types are indicators for different groups of infractions: guns, drugs and alcohol, truancy, major violence/weapons, minor violence/weapons, and major non-violent. Baseline infraction category is Minor Non-Violent Infractions.

The coefficients on the indicator for Black students is smaller in the school fixed effects models (indeed, in the most full model, the odds ratio is not statistically different from one), relative to the models without school fixed effects, indicating that the racial disparities in exclusionary discipline are driven almost entirely by differences across schools rather than within schools. Because these analyses revealed that the between school differences are so important, in the following section, we test which characteristics of schools drive these differences.

Research Question 3: What are the school characteristics that are associated with harsher (longer) disciplinary consequences?

Since there are larger racial disparities across the state than specifically within schools, it could be that there are differences in the disciplinary policies and practices at the types of schools that serve large proportions of non-White students. We test this by creating a School Severity Index (SSI) for each school using the residuals from an infraction-level model predicting the length of exclusionary punishments of various types. In this model, consequences other than exclusionary discipline (e.g. in-school suspension, corporal punishment, no action, or other actions) are coded as zero days of punishment, but are not removed from the model. The residuals are averaged at a school level to generate the school SSI: a positive SSI indicates that a school tends to give out longer (more exclusionary) punishments for similar types of infractions. A negative SSI indicates shorter (less exclusionary) punishments. These SSIs were created using school random effects to account for the noisy measures within schools with fewer disciplinary incidences by allocating greater weight to the schools with larger sample sizes and more precise measures.

The SSI for each school are then regressed on a variety of school-level characteristics. The results in Table 7 are based on SSIs created in the first-stage using the days of exclusionary punishment. Other types of consequences are included as zero days. Importantly, the R-squared values in the models with the race percentage variables (columns 2, 3, 5, and 6) have about 2.5 times the predictive power of those without the race percentage variables (columns 1 and 4). Therefore, the racial breakdown of schools appears to be an important factor in explaining disciplinary outcomes within schools. The results here are stable across models and consistent with our earlier findings: Schools serving greater proportions of Black students had higher scores in the severity index and thus longer punishments; schools serving greater percentages of Hispanic students had lower scores.

Other variables in Table 7 also appear to have significant relationships to the severity of punishment. Open-enrollment charter schools, all else equal, appear to give out somewhat harsher punishments (an extra 0.4 to 1.3 days of punishment, per infraction, depending on the model). Importantly, the charter school coefficient is much lower in the models including controls for the school racial demographics than in the models without these variables. Open-enrollment charter schools in the state are primarily clustered in urban areas and serve a larger proportion of Black students (41%) than the state average (21%), and fewer White students (46%, relative to 63%), so without controlling for these racial demographics, the charter school variable is confounded with racial demographics as well. It is possible that charter schools, conditional on student demographics, may use harsher punishments, if, for example, they focus on a so-called "no excuses" model as in the highly successful KIPP Charter Network (Arkansas had two locations in 2014-15, the last year in our dataset). In addition, evidence from Michigan indicates that students may seek out charter schools if they are already having disciplinary issues or other problems in the traditional public school district (Horn & Miron, 2000).

The coefficients on middle school and high school may be surprising. Based on these coefficients alone, middle schools and high schools appear to be administering relatively less severe consequences than elementary schools in the state, for the same types of infractions. While these coefficients are statistically significant, the magnitude is somewhat small (about 0.1 days, per infraction). To interpret this relationship, it is also important to note that the SSI is the average residual from a model predicting the severity (number of days) of a punishment, as a function of a variety of things, including grade level. Therefore, the SSI, in some ways, is already accounting for the increasing severity by grade-level, so we just treat these school-type variables as control variables.

Table 7
School Characteristics Associated with Harsher Punishments (Dep Var = School Severity Index based on days of exclusionary discipline, units= number of days)

	(1)	(2)	(3)	(4)	(5)	(6)
Log (School Enrollment)	0.144***	-0.0202	-0.079***	0.144***	-0.0209	-0.073***
	(0.0192)	(0.0166)	(0.0187)	(0.0191)	(0.0166)	(0.0187)
Sch. % Black		1.392***	1.586***		1.377***	1.552***
		(0.0355)	(0.0449)		(0.0355)	(0.0450)
Sch. % Hispanic		-0.800***	-0.658**		-0.887***	-0.761***
		(0.266)	(0.265)		(0.266)	(0.266)
Sch. % Other Minority		1.716***	1.444***		1.762***	1.518***
		(0.274)	(0.275)		(0.274)	(0.276)
Sch. % FRL	1.106***		-0.554***	1.118***		-0.499***
	(0.0642)		(0.0775)	(0.0641)		(0.0777)
Sch. % Gifted & Talented	2.934***	2.093***	1.908***	2.925***	2.083***	1.919***
	(0.185)	(0.170)	(0.172)	(0.185)	(0.170)	(0.173)
Sch. % Special Education	-0.144	-0.112	-0.0269	-0.178	-0.143	-0.0636
	(0.168)	(0.155)	(0.154)	(0.168)	(0.155)	(0.155)
Sch. % LEP	0.410***	1.901***	2.116***	0.433***	2.012***	2.207***
	(0.117)	(0.308)	(0.308)	(0.116)	(0.308)	(0.309)
Open-Enrollment Charter	1.294***	0.541***	0.389***	1.292***	0.542***	0.408***
	(0.0884)	(0.0825)	(0.0859)	(0.0883)	(0.0826)	(0.0860)
Middle School	-0.148***	-0.085***	-0.098***	-0.145***	-0.082***	-0.094***
	(0.0306)	(0.0283)	(0.0282)	(0.0306)	(0.0283)	(0.0283)
High School	-0.115***	-0.0567**	-0.080***	-0.115***	-0.0578**	-0.079***
	(0.0272)	(0.0250)	(0.0251)	(0.0271)	(0.0250)	(0.0251)
Other School Type	0.680***	0.487***	0.431***	0.690***	0.500***	0.449***
	(0.0980)	(0.0905)	(0.0903)	(0.0978)	(0.0906)	(0.0905)
School Year Indicators	Y	Y	Y	Y	Y	Y
Constant	-1.771***	-0.446***	0.195	-1.777***	-0.434***	0.141
	(0.140)	(0.106)	(0.141)	(0.140)	(0.106)	(0.141)
Observations	6,871	6,891	6,871	6,871	6,891	6,871
R-squared	0.096	0.236	0.241	0.098	0.233	0.238

Note: Standard errors in parentheses; * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Models 1-3 use SSI created with all 17 infraction categories in the first stage. Models 4-6 use SSI created with the 7 infraction groups in the first stage. In the first stage, SSI were created using school random effects. Baseline school type is Elementary.

Overall, models 1, 2, and 3 are very similar to models 4, 5, and 6, respectively; the only difference is the number of infraction categories used to generate the SSI. The fullest models are presented in columns 3 and 6. In addition to the results related to the racial composition of the schools, there are few other interesting relationships. For example, schools serving greater numbers of Gifted and Talented students and Limited English Proficient students, also appear to give out somewhat harsher punishments. After racial characteristics are controlled for, the coefficient on the Free and Reduced lunch variable is actually negative – schools serving more students in poverty are less harsh in doling out consequences. The phenomenon here appears to be that the racial characteristics, which are correlated with the poverty rates, are driving the school-level differences in disciplinary outcomes. In fact, this seems consistent with our overall finding that Black students across the state face more severe disciplinary responses than do their peers. We will delve deeper into this finding in the section that follows on robustness checks.

Robustness Checks

As a robustness check, we also ran models excluding the log of school enrollment. We do not present the tables summarizing these results here, but in all cases, in the models without school size accounted for, the disparities are similar but slightly higher.[6] In addition, we also conducted robustness checks using days of *any* type of punishment (not just exclusionary). The results for the primary variable of interest (School percent Black) are generally similar. There are some differences, however, in terms of the coefficients on the School percent Hispanic, which has a significantly negative relationship with SSI when created using only exclusionary discipline, but a non-significant relationship in terms of days of any type of consequence. This indicates that, all else equal, while there is no relationship between the proportion of Hispanic students and the length of any type of disciplinary consequence, it is the case that schools with a greater proportion of Hispanic students generally give out shorter exclusionary type punishments.

There is a surprising result from the models that include a measure of the percent of students who are FRL-eligible as well as percent Black (columns 3 and 6 in Table 7). The coefficients on the school percent FRL in this model are negative, despite the fact that the coefficients on FRL are positive in the models that do not control for percent Black (columns 1 and 4 in Table 7). This is likely due to significant correlation between the percentage of students who are FRL-eligible in each school and the percentage of students who are Black in each school (r=0.499).

To further understand what is happening within schools in terms of both White/non-White breakdown and the general income level of the students served, we created indicators for four types of schools (Low-Income Mostly White, Low-Income Mostly non-White, Higher-Income Mostly White, and Higher-Income Mostly non-White). These four categories are based on whether a school is above or below the state average on two separate indicators (percent White and percent FRL). The state averages during the study period were about 65% White and about 60% FRL. The uneven distribution of observations across these groups, as in Table 8, reflects the relative presence of these types of schools in the state, in the sense that there are relatively few schools that are mostly-non-White and higher-income (8%), relative to the other three types.

---

[6] Tables are available from the authors upon request.

Table 8
Distribution of Four School Types

		Higher-Income <60% FRL	Low-Income ≥60%FRL
Mostly-Non-White	<65% White	585 School Year Combinations (8%)	2,185 School-Year Combinations (32%)
Mostly-White	≥65% White	2,237 School-Year Combinations (32%)	1,886 School-Year Combinations (27%)

According to the results in Table 9, it seems that the schools with more non-White students (regardless of whether those schools tend to be higher income or lower income), tend to administer harsher (longer) punishments than the baseline schools serving more white students and higher-income students. The first set of coefficients of interest indicates that schools serving more non-white students, who are also higher income, still receive an additional half a day (roughly) of exclusionary discipline, per infraction, relative to their peers in schools serving more white students. Similarly, the third set of coefficients of interest indicates that students in these relatively poor, relatively non-White schools, receive about 0.6 days of extra punishment, relative to students in the relatively wealthy, relatively white schools. Therefore, these two findings indicate that schools with more non-White students tend to give out longer punishments, regardless of the percentage of students receiving FRL in a school.

One of the most interesting results, however, is that there were no statistically significant differences between the length of punishments in wealthier and less wealthy schools, conditional on serving ≥65% white students. This suggests that racial factors appear more important than income factors for predicting the severity of disciplinary consequences. This seems consistent with our earlier models (Table 7); the magnitude and sign on the race variable is mostly unchanged by the inclusion of the poverty variable in the model. On the other hand, the poverty result is very sensitive to the inclusion of the race variable. In Table 9, open-enrollment charter schools still appear to use more severe consequences (about an extra 0.6 to 0.8 days of punishment, per infraction).

As an additional robustness check, we also created the SSI using days of any kind of punishment, rather than only the days of exclusionary discipline. The coefficients for each of the four school types in this additional model are nearly identical to the coefficients from the primary model reported in Table 9.

Table 9
School Characteristics Associated with Harsher Punishments (Dep Var = School Severity Index based on days of exclusionary discipline, units= number of days)

	(1)	(2)	(3)	(4)
Log (School Enrollment)	0.002	-0.004	0.004	0.000
	(0.019)	(0.018)	(0.019)	(0.018)
Schools serving <60%FRL, <65% White	0.479***	0.544***	0.471***	0.537***
	(0.041)	(0.040)	(0.041)	(0.040)
Schools serving ≥60%FRL, ≥65% White	0.033	0.009	0.042	0.019
	(0.028)	(0.028)	(0.028)	(0.029)
Schools serving ≥60%FRL, <65% White	0.624***	0.611***	0.619***	0.609***
	(0.028)	(0.026)	(0.028)	(0.026)
School % Gifted and Talented	2.235***		2.235***	
	(0.182)		(0.182)	
School % Special Education	-0.270		-0.301*	
	(0.165)		(0.165)	
School % LEP	-0.0615		-0.0184	
	(0.116)		(0.116)	
Open-Enrollment Charter School	0.782***	0.599***	0.785***	0.604***
	(0.088)	(0.085)	(0.088)	(0.085)
Middle School	-0.125***	0.015	-0.122***	0.0170
	(0.030)	(0.028)	(0.030)	(0.028)
High School	-0.111***	0.0128	-0.112***	0.0110
	(0.026)	(0.025)	(0.027)	(0.025)
Other School Type	0.563***	0.526***	0.575***	0.534***
	(0.096)	(0.095)	(0.096)	(0.095)
School Year Indicators	Y	Y	Y	Y
Constant	-0.396***	-0.25288	-0.404***	-0.274**
	(0.123)	(0.113)	(0.123)	(0.113)
Observations	6,891	6,892	6,891	6,892
R-squared	0.142	0.122	0.140	0.120

Note: Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Models 1 and 2 use SSI created with all 17 infraction categories in the first stage. Models 3 and 4 use SSI created with the 7 infraction groups in the first stage. In the first stage, SSI were created using school random effects. Baseline school type is Elementary. Baseline school type is more White and more non-FRL (<60% FRL, ≥65% White).

Discussion and Conclusions

There have been numerous studies over the past twenty years documenting the existence of racial disparities in disciplinary consequences. From this research base, we know, with some confidence based on multiple studies across many years, that Black students are both referred for discipline more often and receive exclusionary disciplinary far more often than other students (e.g. Skiba et al., 2002). Moreover, some more recent studies have suggested that Black students receive more severe and longer consequences than their peers, who have committed identical infractions (e.g. Skiba et al., 2002). Finally, based on a single recent study in a single state for a single year (Skiba et al., 2014), along with inferences we can draw from a 1998 survey and consequence-data from Chicago (Sartain et al., 2015), we believe these disparities are more likely driven by differences between schools than by differential treatment of students within a given school.

In this study, we aimed to build on this growing research base by analyzing all infraction-level disciplinary data for every public school in Arkansas over a seven year time span. Consistent with earlier evidence, we found disproportionate use of exclusionary discipline for Black students, and we also found that these disparities are primarily due to differences in discipline practices across schools, rather than within schools. This result supports the important work of Skiba et al. (2014), and builds upon that work by providing analysis of an entire state over seven school years, rather than just one school year.

When school fixed effects are not included, Black students are about 2.4 times as likely as their White peers in the state (in the same grade and with similar numbers of previous infractions) to receive exclusionary discipline for similar infraction types. Hispanic students are slightly less likely than their similar White peers in the state to receive exclusionary discipline. Importantly, the disparities are not only based on race. Depending on whether or not race was also controlled for, our results indicate that FRL students in the state are about 1.2 to 1.5 times as likely to receive exclusionary discipline as their non-FRL peers.

However, we conclude that most of the racial differences in rates of exclusionary discipline are across schools, because these racial disparities diminished greatly when school fixed effects were included. Within schools, Black students are only slightly more likely than White students to receive exclusionary discipline (relative risk ratio of 1.04, significant at the 99% confidence level). Interestingly, within schools, there still appear to be persistent gaps in the use of exclusionary discipline for FRL students and special education students (relative risk ratios of about 1.2). The implications of this for state or district-level policy are not exactly clear.

These results indicate that the large racial disparities tend to be across schools, and therefore a function of the types of schools that non-White students are likely to attend, whereas within schools, there may be larger concerns about disparities based on socio-economic status and special education status. Since the results indicate that the state-level racial disparities are likely a function of the school attended, we also tested which school level factors were associated with a measure of school disciplinary severity (SSI), and found that the percent of the school that is Black or the percent of the school that is of another non-White, non-Black, non-Hispanic group are both significant predictors of harsher (longer) consequences, which supports the idea that most of the racial disparities occur due to different disciplinary practices being used in districts/schools serving different racial compositions of students.

When schools were split into four categories based on the proportions of FRL students and White students in the school, we found that schools serving more non-White students (regardless of the proportion of FRL students) administered longer punishments than schools serving mostly White, non-FRL students. However, lower income, mostly White schools were actually quite similar

to the higher-income, mostly White schools, again indicating that differences in exclusionary practices across schools appear to be more driven by racial demographics than by income or poverty.

Overall, then, there seem to be two broad conclusions from this work. First, non-White students are far more likely to receive exclusionary discipline for a given infraction than are their White peers, and this disparity is driven by disciplinary practices employed at the schools non-White students attend. Second, the differences by race are far more impactful than the differences due to poverty. So, what are the implications for policy?

Based on the analyses presented here, and on our interactions with state level policymakers, we believe there are two broad lessons for policymakers and school leaders, the first related to data transparency and the second to targeted reforms.

First, we have uncovered patterns in disciplinary consequences that were previously not well known in the local education community. Thus, we believe that a critical first step in creating positive change with regard to student discipline is broadly sharing discipline data with education stakeholders including staff, administration, families, and communities. For example, school leaders and state policymakers would benefit from reports that allow for comparisons of disciplinary practices and statistics across schools. Given that the diaparities are primarily across schools and not within, school leaders may not be aware of a problem until they are able to compare their school to others in the state. When awareness of potential disparities in discipline is raised, school leaders may seek out more concrete programs or strategies to address such issues. It is possible that simply sharing data on school-level rates of exclusion may create awareness that serves as a catalyst for action within communities as well. Moreover, as Tatto et al. (2001) have shown, parental perceptions of unequal or overly strict disciplinary practices can undermine school culture. When policymakers and/or school leaders actively share discipline data, parents can be empowered to advocate for their children and work with school leaders to devise solutions.

Second, the primary conclusion policymakers should draw from our analyses is that the clearest evidence of racial disparities in discipline occur across schools. That is, schools serving predominantly Black students impose more severe (longer) exclusionary consequences on students, even after controlling for the type of infraction. Thus, to address these disparities statewide, policymakers can focus on these particular schools which serve mostly Black students and are engaging in particularly severe disciplinary practices.

One strategy that policymakers might adopt would be mandates to reduce suspensions – at least for minor nonviolent infractions – in targeted schools. This sort of change could be impactful as nearly half (49%) of the infractions that lead to exclusionary discipline are minor, non-violent, and perhaps subjective. These infractions include disorderly conduct (~26%), insubordination (~20%), and truancy (~3%). It surely seems possible to address such (mostly minor) infractions with preventative or restorative alternatives to exclusionary discipline. For example, there is some evidence that simply revising codes of conduct (or setting policies) to reduce the use of suspensions for minor offenses and limit the length of suspensions may be effective (Lacoe & Steinberg, 2016; Mader et al., 2016) and at little cost to school climate (Mader et al., 2016).

Further, there are school-based interventions, some of which have been rigorously evaluated, designed to improve school climate and disciplinary outcomes. For example, there are non-experimental studies that find reductions in referrals or suspensions and expulsions with programs such as Response to Intervention (RTI), which attempts to prevent recidivism by responding to behavioral issues as they arise (Fairbanks et al., 2007). Another strategy, commonly known as restorative justice, is viewed as a movement to "institutionalize peaceful and non-punitive approaches for addressing harm" that in a school setting can serve as an alternative to exclusionary

discipline (Fronius et al., 2016). Essentially, restorative justice is a non-punitive approach to handling conflicts, but these programs can take the form of whole-school interventions or as an "add-on" to respond to specific situations (Fronius et al., 2016). Finally, School-Wide Positive Behavioral Interventions and Supports (SWPBIS a.k.a. PBIS) may be the most well-known behavioral intervention and, fortunately, has been subject to some rigorous evaluation. Indeed, there is some experimental evidence that indicates implementation of the PBIS framework decreases office referrals (Flannery et al., 2014) and improves student perceptions of school safety and test scores (Horner et al., 2009). As of June 2016, there are currently 49 known schools implementing PBIS in Arkansas (Saarnio & Merten, 2016).

While mandated reductions in exclusionary discipline may be appealing and are certainly simple, we have two reservations about this approach. First, it is possible that school leaders may respond to mandates superficially, by changing reporting patterns without substantially improving their disciplinary practices. Second, mandated reductions without any other supports are unlikely to be effective. Front line educators in the schools will need some alternative to exclusionary discipline if such mandates are put in place; the school culture is unlikely to improve if educators do not have the necessary capacity to respond to behavioral infractions. Thus, if state policymakers are to mandate reductions, they should also consider providing schools with access to some of the more positive alternative disciplinary strategies described above.

While we advocate for data transparency, we understand that there are limitations to the conclusions that should be drawn and there is a real potential for unintended consequences. First of all, it is not obvious that high numbers of disciplinary referrals and consequences are bad – or good. For example, a school with very few reported infractions may either be one with a great school climate, or one where adminstrators fail to address real problems related to student discipline. Thus, school context matters. Finally, even without a policy mandate aimed at reduction, the greater public attention paid to disciplinary data may have the unitended effect of encouraging school personnel to simply under-report or code disciplinary infractions and consequences in vague categories (such as "other").

To be clear, we are not suggesting that policymakers ignore disciplinary data, but that these numbers be interpreted with appropriate caution and in context. In fact, while the discussion of disciplinary disparities has been ongoing for several decades, the practice of public reporting of school level discipline data is relatively new. Thus, while policymakers should certainly pay attention to these data, we would argue that it is premature to attach high-stakes consequences to disciplinary outcomes.

Ultimately, while the results presented here do not provide step-by-step solutions, they do provide further confidence in the early findings from the research literature that Black students face disciplinary disparities, even for relatively minor infractions. We have also provided some information about which schools in Arkansas are more likely to impose relatively severe consequences that remove students from classrooms. The first step in addressing a potential problem is identifying it. It is our hope that policymakers and researchers and school leaders collaborate on the next step: to implement potentially effective strategies and rigorously evaluate the results to improve the schooling experience for students in the future.

Acknowledgements

References

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. https://doi.org/10.1037/0003-066X.63.9.852

Anderson, K. P. & Ritter, G. W. (in press). *Do school discipline policies treat students fairly? A second look at school discipline rate disparities.* Manuscript submitted for publication.

Angrist, J. D., & Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

AR Code § 6-18-507 (2015). Elementary and Secondary Education, Subchapter 5: Discipline.

AR Code § 6-48-104 (2015). Special Education Programs, Chapter 48: Alternative Learning Environments.

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. https://doi.org/10.1002/yd.55

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. https://doi.org/10.1525/sop.2012.55.4.637

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

Cornwell, P. (September 23, 2015). Seattle School Board halts suspensions for elementary students. *The Seattle Times.* Retrieved from: http://www.seattletimes.com/seattle-news/education/seattle-school-board-halts-suspensions-for-elementary-students/

Costenbader, V., & Markson, S. (1998). School Suspension: A Study with Secondary School Students. *Journal of School Psychology, 36*(1), 59-82. https://doi.org/10.1016/S0022-4405(97)00050-2

Curran, F. C. (2016). Estimating the Effect of State Zero Tolerance Laws on Exclusionary Discipline, Racial Discipline Gaps, and Student Behavior. *Educational Evaluation and Policy Analysis, 38*(4). https://doi.org/10.3102/0162373716652728

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. https://doi.org/10.1177/001440290707300302

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. https://doi.org/10.1037/spq0000039

Fronius, T., Persson, H., Guckenberg, S., Hurley, N., & Petrosino, A. (2016). *Restorative justice in U.S. schools: A research review.* West Ed. Retrieved from: http://jprc.wested.org/wp-content/uploads/2016/02/RJ_Literature-Review_20160217.pdf

Horn, J., & Miron, G. (2000). *An evaluation of the Michigan charter school initiative: Performance, accountability, and impact.* The Evaluation Center, Western Michigan University. Retrieved from: http://homepages.wmich.edu/~miron/publics/michigan/Executive_Summary.pdf

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. https://doi.org/10.1177/1098300709332067

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.

Institute for Digital Research and Education. (2011). UCLA. Retrieved from http://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/

Institute for Digital Research and Education. (2017). UCLA. Retrieved from http://www.ats.ucla.edu/stat/stata/output/stata_logistic.htm

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. Retrieved from: http://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-ccrr-research/losen-gillespie-opportunity-suspended-2012.pdf

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.* Working paper.

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. https://doi.org/10.1111/j.1540-6237.2009.00674.x

O'Connor, J. (July 29, 2015). Miami-Dade Schools eliminating out-of-school suspensions. *State Impact, NPR.* Retrieved from: https://stateimpact.npr.org/florida/2015/07/29/miami-dade-schools-eliminating-out-of-school-suspensions/

Public Counsel. (September 27, 2014). California enacts first-in-the-nation law to eliminate student suspensions for minor misbehavior [Press Release]. Retrieved from: http://www.publiccounsel.org/press_releases?id=0088

Raffaele-Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development,* No. 99. https://doi.org/10.1002/yd.52

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. https://doi.org/10.1002/pits.10020

Rogers, W. H. (1993). Regression standard errors in clustered samples. *Stata Technical Bulletin 13: 19–23.* Reprinted in Stata Technical Bulletin Reprints, Vol. 3, 88–94. College Station, TX: Stata Press.

Saarnio, D. & Merten, A. (June 28, 2016). PBIS State Contacts for Arkansas. Retrieved February 12, 2017 from: http://cce.astate.edu/pbis/wp-content/uploads/2016/07/PBIS_Schools_in_Arkansas_-for-Natl-PBIS-Center_6-28-163.pdf

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.

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

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. https://doi.org/10.3102/0002831214541670

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. https://doi.org/10.1023/A:1021320817372

Skiba, R., & Peterson, R. (January 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. Center for Evaluation and Education Policy, Indiana University.

Steinberg, M. P., & Lacoe, J. (2016). What do we know about school discipline reform? *Education Next.* Retrieved from: http://educationnext.org/what-do-we-know-about-school-discipline-reform-suspensions-expulsions/

Sullivan, A. L., Klingbeil, D. A., & Van Norman, E. R. (2013). Beyond behavior: Multilevel analysis of the influence of sociodemographics and school characteristics on students' risk of suspension. *Children, Research, and Public Policy, 42*(1), 99-114.

Tatto, M. T., Rodriguez, A., Gonzalez-Lantz, D., Miller, C., Busscher, M., Trumble, D., Centeno, R., & Woo, A. The challengnes and tensions in reconstructing teacher-parent relations in the context of school reform: a case study. *Teachers and Teaching: Theory and Practice, 7*(3), 315-333.

Tobin, T., Sugai, G., & Colvin, G. (1996). Patterns in Middle School Discipline Records. *Journal of Emotional and Behavioral Disorders, 4*(2), 82-94. https://doi.org/10.1177/106342669600400203

U.S. Department of Education and U.S. 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

Welch, K., & Payne, A. N. (2010). Racial threat and punitive school discipline. *Social Problems, 57*(1), 25-48. https://doi.org/10.1525/sp.2010.57.1.25

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48,* 817–838. https://doi.org/10.2307/1912934

About the Authors

Kaitlin P. Anderson
University of Arkansas
kaitlina@uark.edu
Kaitlin P. Anderson is a Distinguished Doctoral Fellow in the Department of Education Reform at the University of Arkansas. Her research focuses on rigorously evaluating the effects of school policies designed for or affecting disadvantaged students. Her work covers issues such as school discipline, school choice, teacher quality, and quantitative methods.

Gary W. Ritter
University of Arkansas
garyr@uark.edu
Gary W. Ritter is a Professor of Education Policy and holder of the Endowed Chair in Education Policy in the Department of Education Reform at the University of Arkansas. His primary areas of interest are teacher quality, teacher evaluation, racial segregation in schools, and the implementation and evaluation of programs aimed at improving educational outcomes for traditionally under-served students. Ritter's work has been published in *Phi Delta Kappan, Review of Educational Research, Education Finance & Policy, Educational Policy, Education & Urban Society, Educational Evaluation and Policy Analysis, the Journal of Education Finance,* and *Education Next.*

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter feed** @epaa_aape.