Understanding the Racial Discipline Gap in Schools

Maithreyi Gopalan iD

The Pennsylvania State University

Ashlyn Aiko Nelson

Indiana University-Bloomington

We explore the discipline gap between Black and White students and between Hispanic and White students using a statewide student-level panel data set on Indiana public school students attending prekindergarten through 12th grade from 2008–2009 through 2013–2014. We demonstrate that the Black-White disciplinary gaps, defined in a variety of ways and robust to a series of specification tests, emerge as early as in prekindergarten and widen with grade progression. The magnitude of these disciplinary gaps attenuates by about half when we control for many student- and school-level characteristics, but it persists within districts and schools. In contrast, we find that Hispanic-White gaps are initially null and statistically insignificant at the prekindergarten/kindergarten level and attenuate substantially after adjustment for cross-school (district) variation and other covariates. We further disentangle the discipline gap using a decomposition technique that provides empirical support for the hypothesis that Black students nonrandomly sort into more punitive disciplinary environments.

THERE is ample evidence that students of color disproportionately experience adverse disciplinary actions in school, such as in- and out-of-school suspensions and expulsions. Nationally, 5% of White boys and 2% of White girls receive one or more out-of-school suspensions annually, as compared with 18% of Black boys and 10% of Black girls and 7% of Hispanic boys and 3% of Hispanic girls (U.S. Department of Education Office for Civil Rights, 2016).

Children who experience punitive disciplinary actions in school exhibit lower academic achievement on average (Arcia, 2006; Hwang, 2018; Lacoe & Steinberg, 2018). Because exclusionary disciplinary policies reduce children's opportunity to learn and reduce the time that children spend engaged in educationally meaningful activities (Losen, Hodson, Keith, Morrison, & Belway, 2015; Rausch & Skiba, 2006), there is reason to hypothesize that punitive disciplinary actions may have causal adverse effects on children's academic outcomes. Thus, racial disparities in adverse disciplinary outcomes likely exacerbate existing race-based achievement gaps (Morris & Perry, 2016).

However, several important aspects of race-based disciplinary gaps remain unexplored, in large part due to data limitations. For example, the federal Office of Civil Rights (OCR) only recently began (2013–2014 academic year) reporting biennial data on adverse disciplinary outcomes by student race/ethnicity for all schools and districts in the United States. Although the OCR data have national coverage, they cannot be disaggregated to the student or grade level or analyzed across a longer period, and they do not include an extensive set of covariates useful for explaining variation in student suspensions and expulsions. Other studies with disaggregated student-level data from state departments of education (Anderson & Ritter, 2017; Skiba et al., 2014) employed truncated data sets, including students who received only disciplinary infractions and thereby precluding an analysis that accurately predicts the probability of suspensions/expulsions with the full student sample.

In this article, we provide an extensive description of the racial disparities in school disciplinary outcomes that have eluded the literature thus far due to such data limitations. First, we highlight the early emergence of Black-White disciplinary gaps across multiple school disciplinary outcomes (likelihood, frequency, and length of suspension or expulsion) and the widening of these gaps across grades. Second, we answer questions regarding the source of these disparities: To what extent do these gaps arise within schools, across schools in the same district, or across districts? To what extent do disciplinary gaps attenuate when controlling for observable student-, school-, and district-level characteristics? Third, we document a key feature of segregation in Indiana—the substantial nonrandom sorting of Black

students across a few districts—and examine to what extent this cross-district sorting explains racial disciplinary gaps.

We apply these methods to statewide panel data on all public school students attending prekindergarten through 12th grade in Indiana during the 2008–2009 through 2013–2014 academic years ($N = 3,341,065$ students and 7,223,752 student-year observations). Our data set includes a large number of student and school characteristics that enable us to explore different research questions and account for a much more extensive set of covariates than what were available in prior studies of the discipline gap.

Results from our analyses indicate that the probability that Black students in Indiana will experience a suspension/expulsion in any school year is between 13 and 16 percentage points higher than for White students. The Black-White discipline gap widens with grade progression, attenuating slightly but remaining significant after accounting for school or district fixed effects. This finding suggests that, at least in Indiana, the discipline gap is driven by variation both across and within schools (districts). Furthermore, there is substantial heterogeneity in the magnitude of adverse disciplinary outcomes and race-based disciplinary gaps across school districts. While we cannot fully disentangle the unobserved sources of that heterogeneity, we find that, on average, Black students attend schools in districts with markedly different average demographic characteristics and that racial sorting across districts explains a significant portion of the racial discipline gaps in our fixed effects and decomposition analyses. In contrast, we find that the Hispanic-White gaps are lower than the Black-White gaps in Indiana (zero and statistically insignificant in lower grade levels) and are mostly explained by observed student-level differences and cross-district variation. Our detailed descriptive analyses follow recommendations advocated in a recent Institute of Education Sciences report by Loeb et al. (2017), which stated that analyses such as those in this article are essential for hypothesis generation of causal mechanisms that can be tested in the future, as well as for the design and evaluation of appropriate policy interventions.

In the sections that follow, we provide a brief review of relevant literature, describe our data and methods, discuss results from our analysis, and conclude with policy recommendations.

Literature Review

The Children's Defense Fund (1975) published one of the earliest reports documenting the overrepresentation of Black students in school suspensions. With national data obtained from the Office of Civil Rights, the report revealed that suspension rates among Black students exceeded those of White students in more than two-thirds of school districts. Since then, multiple studies have reported similar results. For example, Rocque (2010) found evidence of a Black-White

disciplinary gap in disciplinary referrals. Costenbalder and Markson (1998), Gregory and Weinstein (2008), and Skiba et al. (2014) reported racial disparities in in- and out-of-school suspensions, and Theriot (2009) documented a race-based disciplinary gap in school arrests. Yet, most of these studies suffered from data and methodological limitations that precluded estimation of race-based discipline gaps with more varied measures of disciplinary outcomes capturing multiple infractions and the duration of suspensions/expulsions. Many previous studies also did not account for students' prior disciplinary infractions—such as multiple suspensions and the length of suspensions/expulsion.

Furthermore, despite wide acknowledgment of the discipline gap, there is little consensus regarding its causes. One possible explanation is the individual deficit hypothesis, which may be grouped into two categories: differential behaviors and differential attributes. The differential behavior hypothesis posits that minority and White students exhibit different behaviors in school, which in turn contribute to racial disparities in adverse disciplinary outcomes. For example, Wright, Morgan, Coyne, Beaver, and Barnes (2014) used data from the Early Childhood Longitudinal Study–Kindergarten and found no evidence of racial differences in suspension rates among eighth graders after accounting for teacher assessments of early problem behaviors. Similarly, Rocque (2010) and Bradshaw, Mitchell, O'Brennan, and Leaf (2010) found that the magnitude of the race-based gap in disciplinary referrals decreased after accounting for teacher-reported measures of student behavior. Townsend (2000) raised the concern that Black students might use more emotive communication styles with louder voices, which may be misinterpreted as noncompliance, and that this behavioral difference may result in disproportionality. A major empirical issue in studies that accounted for teacher ratings of student behavior is that these measures may reflect racial bias. For example, if teachers' perceptions of student behavior exhibit racial bias (Lindsay & Hart, 2017; Okonofua & Eberhardt, 2015), then estimates of the remaining race-based discipline gap will be biased downward. Some studies (Losen & Skiba, 2010; Skiba et al., 2014) argued that race-based discipline gaps may be attributed to racial bias if differences in student behavior do not fully account for the gaps. As discussed here, however, measures of student behavior themselves may be subject to racial bias. Furthermore, objective measures of student behavior are rarely available from statewide or nationally representative data sets. While our data do not include measures of student behavior or teacher attributes or perceptions of student behavior, these mechanisms may reside in the "unexplained" components in our analysis. In our results, we explicitly disentangle and discuss this unexplained component. Future research should explore these potential mechanisms with a multitude of research designs.

The differential attributes hypothesis implies that the distribution of individual and school characteristics across minority students is different from that of White students and that these differences in endowments are associated with the race-based discipline gap. For example, minority students are disproportionately economically disadvantaged, and economically disadvantaged students are overrepresented in adverse disciplinary outcomes (Brantlinger, 1991; Wu, Pink, Crain, & Moles, 1982). Studies that include controls for student socioeconomic status (e.g., Skiba, Michael, Nardo, & Peterson, 2002) found that measures of student poverty explain some but not all of the race-based discipline gap.

Another issue is that the disproportionate sorting of minority students into more punitive school environments does not imply evidence of discrimination per se. Previous studies found that minority students are disproportionately exposed to punitive institutional environments (Gregory & Weinstein, 2008; Kinsler, 2013; Peguero & Shekarkhar, 2011; Skiba et al., 2014; Wu et al., 1982). Wu et al. (1982), Kinsler (2011), and Skiba et al. (2014) found that school-based disciplinary policies account for a substantial portion of the race-based discipline gap. Yet studies that attributed all cross-school variation in disciplinary outcomes to racial bias likely overestimated the role of discrimination in the race-based discipline gap. However, other studies (e.g., Kinsler, 2011) defined racial bias as the differential treatment of Black and White students within the same school. After accounting for cross-school variation in adverse disciplinary outcomes with school fixed effects, Kinsler (2011) and Anderson and Ritter (2018) found no evidence of racial disparities *within* schools. We argue, however, that this approach likely underestimates the role of discrimination in the race-based discipline gap because any time-invariant structural racial discrimination at the school level will be subsumed in the school fixed effects. For example, principal attitudes toward school discipline (Skiba & Edl, 2004), accountability pressures (Figlio, 2006), and entrenched racial segregation may affect the cross-school variation in race-based discipline gaps. Some of these time-invariant school-level characteristics are attributable to racial discrimination. Additionally, schools that are more racially homogeneous might "drop out" of fixed effects analyses given the absence of within-school racial variation, thereby biasing measures of race-based discipline gaps. Yet, the existence of several racially homogeneous schools, especially schools with high rates of minority student composition, are prevalent in the United States and exhibit high exclusionary disciplinary rates.

Since the 1954 *Brown vs. Board of Education* Supreme Court decision outlawing de jure school segregation, several studies have explored school segregation in the United States (e.g., Clotfelter, 2004; Coleman, Kelley, & Moore, 1975; Orfield & Lee, 2007). Nonetheless, there is little consensus about the mechanisms through which segregation influences student outcomes (Reardon & Owens, 2014). Furthermore, we are aware of only one study that examined the direct effect of segregation on school disciplinary outcomes. In that study, Billings, Deming, and Rockoff (2014) stated that they found no effect of segregation on school disciplinary outcomes, such as out-of-school suspensions, although they did not report results from that analysis. Even though several studies (e.g., Rocha & Hawes, 2009; Skiba et al., 2014; Welch & Payne, 2010) found that the racial composition of schools and school districts—as measured by the share of Black students—predicted suspension/expulsion rates, we are aware of no other study that examined the extent to which the nonrandom sorting of minority students across districts (schools) was associated with the race-based disciplinary gap.

Far less attention has been paid to racial disparities in discipline for non-Black minority students. For Hispanic students, the results are inconsistent (see Welsh & Little, 2018). While some studies found significant overrepresentation of Hispanic students in exclusionary discipline (Gregory, Skiba, & Noguera, 2010; Peguero & Shekarkhar, 2011), others found statistically insignificant differences (Ramey, 2018), especially in lower grade levels (Skiba et al., 2011). Furthermore, there is much less evidence on the sources of and mechanisms causing these disparities. The mechanisms through which racial discrimination may cause disparities across students from different minority races could be vastly different. For instance, Rueda (2015) found that teachers interpret Hispanic students' behavior more positively as compared with Black students. She found that schools often do not punish Hispanic students, even when teachers catch them breaking rules. Furthermore, past studies showed that the "racialized legacy" (Ramey, 2018) of incarceration does not affect Hispanics in the same manner that it does Blacks; Hispanic children may thus not experience criminalization in the same ways as Black children (Irwin, Davidson, & Hall-Sanchez, 2013; Western, 2006).

Taken together, previous research suggests that poverty and school-level factors account for a significant share of the racial discipline gap. Prior research is mixed regarding Hispanic-White gaps and the persistence of within-school (district) racial discipline gaps, and their causes remain unclear. This study advances the extant literature in several important ways. First, we describe Black-White and Hispanic-White gaps across all grades using better measures of disciplinary outcomes than what are available in previous literature. Second, we depart from past studies to analyze within-district gaps more extensively and estimate the extent to which differences in the districts (schools) attended by students of different races may contribute to variation in racial discipline gaps, using fixed effects and decomposition analyses. Specifically, our decomposition and fixed effects analyses disentangle the sources of racial discipline gaps into within- and across-school (district) factors, which prior literature did not do.

Data and Measures

This study employs statewide student-level panel data obtained from the Indiana Department of Education (IDOE) on all prekindergarten–12th grade Indiana public school students in the 2008–2009 through 2013–2014 academic years. For each year, the data include information on student race, gender, poverty status (indicated by enrollment in a free or reduced-price meal program), special education status, attendance, grade, and school and district indicators.

We categorize student race into four main categories for ease of exposition: White, Black, Hispanic, and other. The "other" race category constitutes between 5% and 7% of the sample across years and grade levels and is composed of multiracial students (4%–5%), Asians (1%–2%), American Indians (<0.5%), and Native or other Pacific Islanders (<0.5%).[1]

We link each student-year record to school-level data on the share of minority students, the share of students in poverty, the share of special education students, and attendance at the school level. IDOE requires all public schools in Indiana to provide this student-level information every year, including charter schools that receive some exemptions from reporting on bullying, gang, and arrest-related information. IDOE provides schools with detailed standardized data reporting guidelines, which include codes for classifying suspensions and expulsions.[2]

For each student in each year, we also observe all suspensions and expulsions, including the length of the suspension/expulsion period. We use these data to construct a series of outcome measures, which include the following:

- Dichotomous variables indicating, in each year, whether a student experienced any adverse disciplinary event (suspension or expulsion)
- Duration of adverse disciplinary outcomes, in days, of all suspensions and expulsions[3]

In addition to employing simple racial composition measures at the school and district levels in our analyses, we calculate district-level measures of *exposure* and *unevenness* (see Massey & Denton, 1988, for an overview of the various dimensions of segregation that can be broadly classified into these two categories).[4] Following Reardon (2016), we first calculate the within-district exposure of Black students to other Black students and to poor students (as measured by the share of students enrolled in free or reduced-price meal programs) by weighting the share of Black or poor students in each school by the proportional representation of that school's Black students among all Black students within the district and then averaging that share across Black students in the district (see Table A10). This measure provides a district-level average share of Black or poor students in the schools attended by an average Black student within the

district and is computationally equivalent to calculating the average share of Black or poor students within schools, conditional on being Black. In Indiana, the average district-level racial exposure of Black students to other Black students is 0.46; that is, for an average Black student in Indiana, across all schools in the district, 46 of every 100 students in the school also are Black. By comparison, the average racial exposure of White students to Black students in the district is 0.06. Similarly, the average share of free/reduced-price lunch recipients in an average Black student's school, across all schools in the district, is a measure of Black students' exposure to poverty in the school. This measure is on average 0.64 for Black students and 0.40 for White students.

Second, unevenness measures describe the *difference* in average school, district, or neighborhood racial or socioeconomic composition among children of different races (Reardon, 2016). For example, the aggregate district-level measure of racial unevenness between Black and White students in Indiana, on average, is 0.40. We calculate district-level unevenness measures by taking simple district-level differences between the exposures of Black and White students to Black and poor students within districts and across schools. Table A10 provides the calculations for these segregation measures. Each measure takes on a value ranging from 0 to 1, with values close to 0 representing low segregation levels and values close to 1 representing high segregation levels.

Descriptive Analyses

We first present descriptive statistics and results from basic regression analyses to characterize the discipline gap in our data set. In the 2008–2009 through 2013–2014 academic years in Indiana, 22% of Black students experienced a suspension or expulsion, as opposed to 8% of White students. Figure 1 depicts suspension/expulsion rates by race across the years in our data set. While the suspension/expulsion rates decline over time, racial disparities persist. We also observe substantial variation in adverse disciplinary outcomes across grade levels, both in terms of suspension/expulsion rate levels and in terms of the magnitude of the Black-White and Hispanic-White discipline gaps. On average, around 1% of White prekindergarten/kindergarten students receive an expulsion or suspension, as compared with 5% of Black and 1.2% of Hispanic prekindergarten/kindergarten students. Across the 2008–2009 through 2013–2014 academic years, 9% to 13% of White high school students received an expulsion or suspension versus 25% to 34% of Black high school students and 15% to 23% of Hispanic high school students (see Figure 2). Therefore, we estimate models separately for students at the prekindergarten/kindergarten, elementary and middle school, and high school levels.[5]
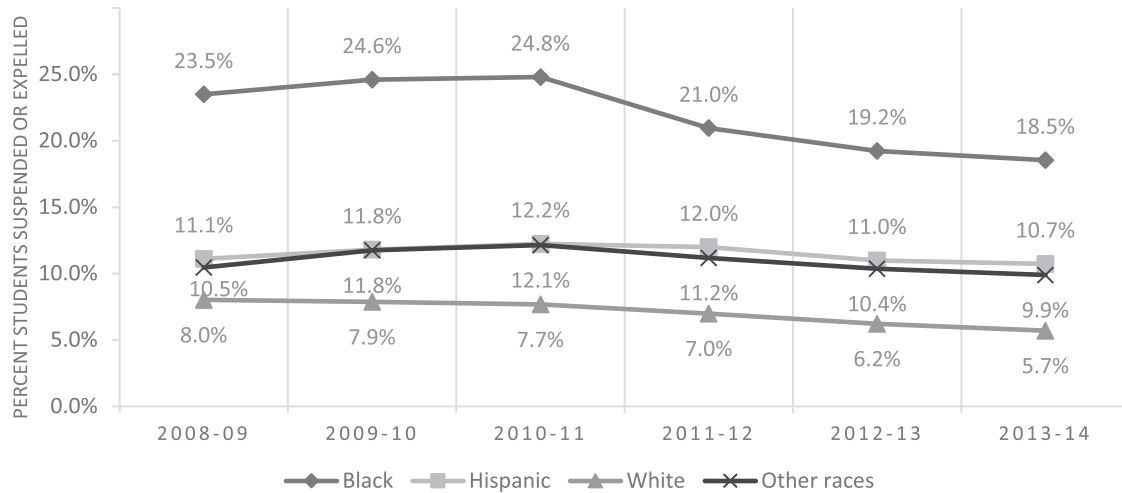
FIGURE 1. *Raw suspensions and expulsions rates, by race across Indiana. The analytic sample includes statewide panel data on all public school students attending prekindergarten through 12th grade in Indiana during the 2008–2009 through 2013–2014 academic years (N = 3,341,065 students).*

Table 1 further illustrates these descriptive statistics and reports summary statistics for the student- and school-level characteristics that we use in our core model specifications, by race and grade level.

We observe substantial cross-race differences in student and school characteristics—for example, White students, on average, attend schools with a lower share of students receiving free or reduced-price meals (29%–40%, depending on grade level) than do Black students (51%–70%) or Hispanic students (42%–60%). We also observe racial differences in school racial composition. An average White student in our sample attends a school in which 6% to 7% of students are Black and 5% to 6% of students are Hispanic, as opposed to Black students—who, on average, attend schools in which 45% to 50% of students are Black and 18% to 20% are Hispanic. On average, Hispanic students attend schools in which 15% to 25% of students are Hispanic and 18% to 20% are Black. In all, as compared with White students, Black students are more likely to be poor (as measured by their eligibility for the free lunch program), attend larger schools with higher minority student composition, and are more likely to receive suspensions and expulsions (with longer days of suspension/expulsion).

Results and Discussion

Magnitudes of Racial Disciplinary Gaps

To estimate raw racial disciplinary gaps, we first regress our outcome measures on race indicators and year fixed effects only, with White students serving as the omitted category. Therefore, the coefficients on the race indicators capture the mean differences in the outcomes between students of the corresponding race and White students.[6] We specify our initial regression model as follows:

$$Y_{it} = \beta \mathbf{X}_{it} + \eta_t + \varepsilon_{it} \qquad (1)$$

In separate specifications, the dependent variable *Y* denotes, in each year, (1) an indicator for whether a student experienced any adverse disciplinary event (suspension or expulsion) and (2) the duration of adverse disciplinary outcomes, in days, of all suspensions and expulsions. We also explore alternative models with dependent variables measured in a variety of ways—count of suspensions and expulsions and likelihood of multiple suspensions/expulsions. The vector $\mathbf{X}_{it}$ denotes the race indicators; η, year indicators; and ε, a random error term. We estimate all models clustering robust standard errors at the district level to account for serial correlation of the errors within districts; robust standard errors also address heteroskedasticity, which may occur when ordinary least squares (OLS) are applied in analyses with dichotomous outcome variables. For models with dichotomous outcome variables, the application of OLS yields coefficient estimates that may be interpreted as linear probabilities. (See Table A5, which includes a robustness check with the use of a logistic regression model instead of OLS and yields qualitatively similar results; we include the OLS results in the main text for ease of interpretation and use cluster-robust standard errors at the district level to correct for heteroscedasticity.)

We present results from this initial model specification in columns 1 and 5 of Table 2. We find that the Black-White discipline gap in the likelihood of suspension/expulsion is 3.8 percentage points in prekindergarten/kindergarten and widens with grade progression to 19.1 percentage points in high school. Black students also receive suspensions/expulsions that are, on average, 0.1 to 1.4 days longer than those of White students. Consistent with past literature, we also find significant gender-by-race interactions for Black

5

PRE-K AND KINDERGARTEN

GRADES 1-8

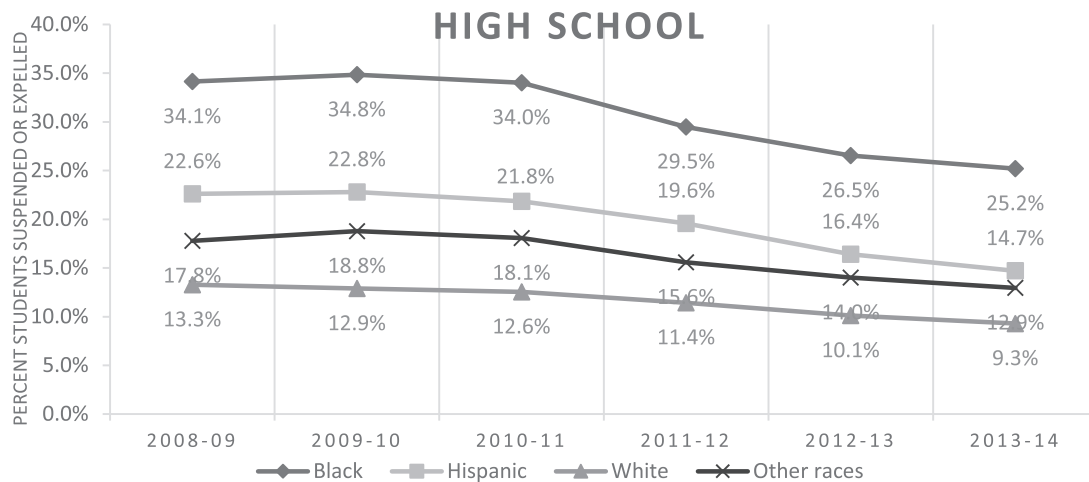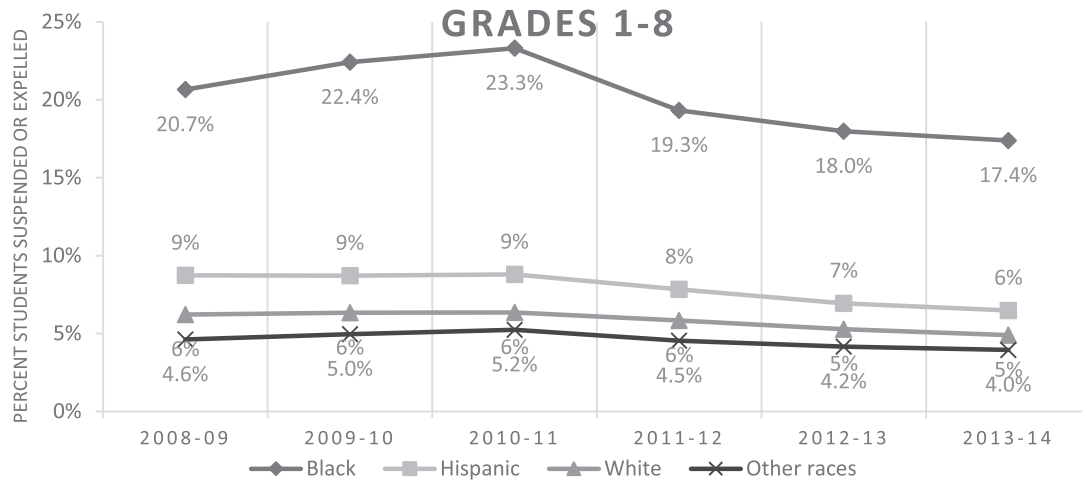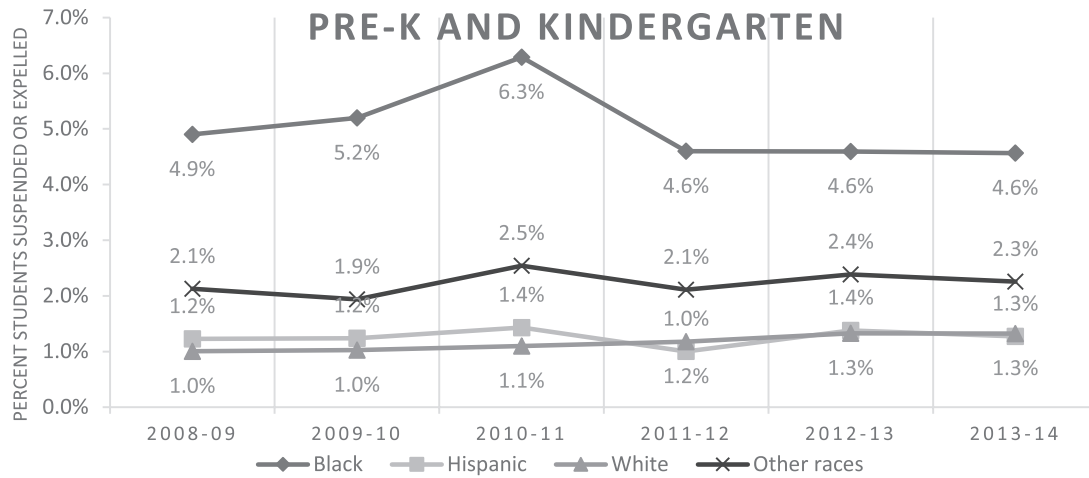## HIGH SCHOOL

FIGURE 2. *Suspension and expulsion rates, by race and partitioned by grade level—prekindergarten and kindergarten, elementary and middle school, and high school. The analytic sample includes statewide panel data on all public school students attending prekindergarten through 12th grade in Indiana during the 2008–2009 through 2013–2014 academic years (N = 3,341,065 students).*

TABLE 1

Descriptive Statistics by Race and Grade Level

	Prekindergarten-kindergarten				Grades 1–8				High school			
	Full sample	Black	White	Hispanic	Full sample	Black	White	Hispanic	Full sample	Black	White	Hispanic
Outcome variables												
Ever expelled/suspended	0.016 (0.132)	0.049 (0.218)	0.010 (0.107)	0.012 (0.115)	0.085 (0.272)	0.207 (0.401)	0.062 (0.234)	0.087 (0.269)	0.167 (0.355)	0.341 (0.461)	0.133 (0.320)	0.226 (0.394)
Ever received an out-of-school suspension	0.012 (0.116)	0.040 (0.200)	0.007 (0.091)	0.009 (0.099)	0.057 (0.227)	0.168 (0.370)	0.037 (0.183)	0.057 (0.218)	0.111 (0.293)	0.274 (0.426)	0.081 (0.253)	0.147 (0.320)
Number of suspensions/expulsions	0.031 (0.378)	0.106 (0.710)	0.026 (0.284)	0.021 (0.281)	0.201 (0.991)	0.530 (1.743)	0.141 (0.763)	0.194 (0.968)	0.416 (1.355)	0.958 (2.279)	0.311 (1.078)	0.590 (1.489)
Length of suspension/expulsion (uncensored)	0.046 (0.776)	0.166 (1.585)	0.026 (0.545)	0.030 (0.460)	0.498 (4.098)	1.264 (6.444)	0.355 (3.510)	0.545 (4.071)	1.529 (8.627)	3.222 (11.64)	1.208 (7.903)	1.976 (9.146)
Length of suspension/expulsion (censored to 20 days)	0.045 (0.609)	0.165 (1.172)	0.025 (0.440)	0.030 (0.035)	0.394 (1.903)	1.045 (3.156)	0.276 (1.542)	0.399 (1.837)	0.932 (2.886)	2.114 (4.289)	0.710 (2.494)	1.212 (3.097)
Race variables												
Black	0.120 (0.337)	1 (0)			0.129 (0.334)	1 (0)			0.127 (0.331)	1 (0)		
White	0.710 (0.465)		1 (0)		0.728 (0.453)		1 (0)		0.769 (0.433)		1 (0)	
Hispanic	0.099 (0.316)			1 (0)	0.081 (0.290)			1 (0)	0.058 (0.255)			1 (0)
Other race	0.071 (0.262)				0.063 (0.250)				0.046 (0.227)			
Other covariates												
Female	0.479 (0.500)	0.474 (0.499)	0.477 (0.499)	0.481 (0.500)	0.487 (0.500)	0.485 (0.500)	0.486 (0.500)	0.493 (0.500)	0.488 (0.500)	0.488 (0.500)	0.488 (0.500)	0.480 (0.500)
Special education status	0.113 (0.320)	0.093 (0.301)	0.124 (0.330)	0.076 (0.285)	0.151 (0.355)	0.161 (0.367)	0.155 (0.358)	0.107 (0.316)	0.136 (0.342)	0.184 (0.382)	0.132 (0.339)	0.096 (0.298)
FRPL status	0.438 (0.500)	0.703 (0.452)	0.350 (0.489)	0.704 (0.450)	0.442 (0.500)	0.740 (0.428)	0.351 (0.487)	0.729 (0.434)	0.326 (0.483)	0.621 (0.476)	0.252 (0.456)	0.599 (0.479)
School enrollment	522.1 (297.3)	570.4 (310.2)	505.9 (298.0)	567.5 (269.8)	604.2 (375.7)	675.6 (369.6)	583.7 (380.6)	657.6 (336.3)	1491.5 (989.1)	1907.6 (1034.5)	1391.5 (959.0)	1712.7 (911.0)
School attendance	152.1 (13.09)	142.3 (16.65)	154.5 (10.87)	147.7 (13.23)	155.1 (12.19)	144.8 (16.39)	157.5 (10.04)	151.1 (12.12)	154.8 (16.11)	142.9 (23.67)	157.1 (13.24)	151.8 (16.46)
Proportion FRPL students in school	0.463 (0.250)	0.686 (0.209)	0.402 (0.231)	0.606 (0.234)	0.437 (0.238)	0.652 (0.199)	0.382 (0.216)	0.566 (0.233)	0.329 (0.188)	0.510 (0.181)	0.292 (0.165)	0.422 (0.205)
Proportion Black students in school	0.129 (0.217)	0.495 (0.292)	0.055 (0.110)	0.189 (0.203)	0.128 (0.208)	0.478 (0.289)	0.057 (0.103)	0.183 (0.201)	0.127 (0.195)	0.448 (0.270)	0.066 (0.112)	0.196 (0.205)
Proportion Hispanic students in school	0.091 (0.129)	0.126 (0.137)	0.063 (0.089)	0.250 (0.196)	0.082 (0.116)	0.118 (0.128)	0.059 (0.082)	0.224 (0.185)	0.058 (0.085)	0.089 (0.097)	0.045 (0.064)	0.155 (0.149)
Observations, n	595,071	77,849	406,306	66,956	4,471,416	571,572	3,185,770	415,309	2,157,265	270,850	1,617,157	151,397

Note. The entries are means (that have been time de-meaned) and standard deviations (overall across the panel in parentheses) of student-level data. The unit of observation is a student-year. FRPL = free/reduced-price lunch.

7

male students across most models (see Table A6). However, the magnitude of the Hispanic-White discipline gap in terms of likelihood of suspension or expulsion is smaller across grade levels and ranges between 0 (statistically insignificant) in prekindergarten/kindergarten and 8 percentage points in high school.

Do Student- and School-Level Factors Attenuate Racial Disciplinary Gaps?

To understand the importance of factors other than race that may influence the outcomes, we next expand our specification in Equation 1 to include a comprehensive set of student and school characteristics. Thus, the race coefficients in these models capture the mean differences in the outcomes between students of the corresponding race and White students, controlling for other factors in the equations. In other words, they provide estimates of regression-adjusted racial discipline gaps. Student characteristics in these models, in addition to race, include indicators for gender, poverty status, grade, special education status, indicators for students' adverse disciplinary events in prior years (a lagged dichotomous measure of whether a student received a suspension/expulsion: $1 = $ yes, $0 = $ no),[7] and attendance. School characteristics include the share of minority students, the share of students receiving free or reduced-price lunch, the share of special education students, and the log of school size. Finally, we include district-level measures of exposure and unevenness to Black and poor students, as well as attendance. We adapt Equation 1 as follows:

$$Y_{it} = \partial \gamma_{st} + \beta \mathbf{X}_{it} + \eta_t + \varepsilon_{it}$$

In this specification, $\mathbf{X}_{it}$ denotes the vector of expanded student characteristics, and γ_{st} denotes school and district characteristics. We report results from this specification in Table 2, columns 2 and 6. After controlling for additional student and school characteristics, the magnitude of the Black coefficient reported in column 1 decreases by about half. Coefficients on the student- and school-level covariates exhibit the expected signs; the coefficient on the indicator for student adverse disciplinary events in prior years is the largest in magnitude and highly correlated with the dependent variables of interest.[8]

Cross-District (School) Versus Within-District (School) Magnitudes of Racial Disciplinary Gaps

Given the mixed evidence in past literature regarding the source of racial discipline disparities within versus across schools, we next employ a series of fixed effects regressions to estimate within-school (district) racial disciplinary gaps. In other words, the race coefficients now capture the mean regression-adjusted within-school (district)

differences between the students of the indicated race group and White students. We adapt Equation 2 to accommodate these fixed effects:

$$Y_{it} = \partial \gamma_{st} + \beta \mathbf{X}_{it} + \theta_d + \eta_t + \varepsilon_{it} \qquad (3)$$

where θ_d alternately denotes time-invariant school and district characteristics. Coefficients on the school (district) fixed effects indicate how the outcome measures vary with systematic features of particular schools (districts) that are constant over time, and they capture variation in disciplinary outcomes across schools (districts). A coefficient associated with a particular school (district) fixed effect should be interpreted as the average baseline level of the outcome variable across periods in that particular school (district). Because school (district) fixed effects account for cross-school (district) variation in the outcomes, the coefficients on school characteristics in γ and on student characteristics in X are estimated with within-school (district) variation; in the school fixed effects specifications, any school characteristics in γ that do not vary over time are subsumed in the school fixed effect and drop out of the model. We report results from our school (district) fixed effects specifications in columns 3 and 7 (columns 4 and 8) of Table 2.

A comparison of results reported in columns 3, 4, 7, and 8 reveals that the coefficients on the Black race indicators attenuate very little after accounting for school (district) fixed effects. Alternatively, to explore the extent to which the unadjusted race-based discipline gaps are driven by within- versus between-district variation, we add district fixed effects prior to adding all the other student- and school-level characteristics. We find that the Black-White disciplinary gap attenuates by about a third but remains significant when we account for district fixed effects (see Figure A1), demonstrating that Black-White disciplinary gaps persist within schools (districts). In contrast, the Hispanic-White disciplinary gap is very small at the lower grade levels, reverses in direction after inclusion of fixed effects in the elementary and middle school models, and attenuates the high school Hispanic-White gap by about half. These findings provide evidence that cross-district variation forms a substantial portion of the Hispanic-White gap in our sample. Taken together, this finding suggests that, at least in Indiana, cross-school (district) differences in disciplinary practices reduce but do not eliminate the race-based discipline gap; within-district (school) racial discipline gaps persist.

Within-District Racial Disciplinary Gaps

We turn next to a deeper exploration of the within-district, regression-adjusted Black-White discipline gaps that

TABLE 2
Estimates of Race-Based Discipline Gaps

	Ever suspended/expelled				Length of suspension/expulsion, days			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prekindergarten-kindergarten								
Black	0.038*** (0.005)	0.024*** (0.002)	0.024*** (0.001)	0.023*** (0.003)	0.149*** (0.021)	0.099*** (0.013)	0.096*** (0.011)	0.096*** (0.013)
Hispanic	0.002 (0.001)	−0.008*** (0.001)	−0.008*** (0.001)	−0.008*** (0.001)	0.002 (0.003)	−0.027*** (0.005)	−0.029*** (0.004)	−0.028*** (0.005)
Other race	0.011*** (0.002)	0.006*** (0.001)	0.004*** (0.001)	0.005*** (0.001)	0.040*** (0.007)	0.024*** (0.005)	0.016*** (0.004)	0.019*** (0.005)
Included other covariates		×	×	×		×	×	×
Year fixed effects	×	×	×	×	×	×	×	×
School fixed effects			×				×	
School district fixed effects				×				×
Observations, *n*				595,071				
Grades 1–8								
Black	0.144*** (0.010)	0.077*** (0.006)	0.078*** (0.003)	0.077*** (0.005)	0.815*** (0.063)	0.428*** (0.049)	0.441*** (0.029)	0.432*** (0.038)
Hispanic	0.021*** (0.005)	−0.011*** (0.002)	−0.012*** (0.001)	−0.011*** (0.002)	0.105*** (0.033)	−0.055*** (0.010)	−0.063*** (0.009)	−0.056*** (0.010)
Other race	0.031*** (0.005)	0.016*** (0.003)	0.015*** (0.001)	0.015*** (0.002)	0.168*** (0.028)	0.079*** (0.018)	0.083*** (0.008)	0.080*** (0.013)
Included other covariates		×	×	×		×	×	×
Year fixed effects	×	×	×	×	×	×	×	×
School fixed effects			×				×	
School district fixed effects				×				×
Observations, *n*				4,471,416				
High school								
Black	0.191*** (0.018)	0.103*** (0.008)	0.105*** (0.005)	0.104*** (0.008)	1.316*** (0.160)	0.731*** (0.075)	0.745*** (0.050)	0.743*** (0.075)
Hispanic	0.079*** (0.013)	0.009** (0.003)	0.009** (0.003)	0.009** (0.003)	0.403*** (0.075)	−0.017 (0.025)	−0.015 (0.023)	−0.015 (0.024)
Other race	0.046*** (0.008)	0.019*** (0.004)	0.017*** (0.003)	0.017*** (0.004)	0.298*** (0.052)	0.115** (0.031)	0.120*** (0.022)	0.126*** (0.028)
Included other covariates		×	×	×		×	×	×
Year fixed effects	×	×	×	×	×	×	×	×
School fixed effects			×				×	
School district fixed effects				×				×
Observations, *n*				2,157,265				

Note. Standard errors in parentheses are clustered at the district level (except in columns 3 and 7, where they are clustered at the school level) to correct for heteroskedasticity and autocorrelation. The unit of observation is a student-year. Length of suspension or expulsion censored to 20 days to limit influence of potential outliers. To economize on space, we report race coefficients only. Alternate specifications include an extensive set of student- and school-level covariates as specified in the main text, school (or district) fixed effects, and year fixed effects.
**$p < .01$. ***$p < .001$.

persist across grade levels. First, we estimate the magnitude of the Black-White gaps based on the likelihood of receiving multiple suspensions/expulsions in a year to better understand if racial disparities persist across multiple disciplinary infractions (see Table A6 for estimates of race-based discipline gaps with the number of suspensions/expulsions as a dependent variable). We also examine the robustness of the within-district race gaps using multiple strategies. For example, we examine if racial gaps in the length of suspension/expulsion are sensitive to extreme values by limiting the analysis to include only students who received at least 1 day of suspension. Furthermore, we explore three interrelated hypotheses. Do within-district Black-White gaps in the length of suspension/expulsion arise (1) from a reduced likelihood of White children ever being suspended and/or (2) from a reduced likelihood of White children receiving fewer suspensions and/or (3) from White children being suspended for fewer days conditional on being suspended? We report results from these additional robustness checks in Table 3.

Column 1 of Table 3 shows that the Black-White within-district gap in the likelihood of receiving more than one suspension/expulsion is statistically significant yet smaller in magnitude as compared with the likelihood of ever receiving a suspension/expulsion across grade levels (also see Table A5, which shows that Black students receive, on average, 0.05 to 0.3 more suspensions in a year as compared with White students). On average, the likelihood that Black students receive multiple suspensions/expulsions in a year is between 1 and 7 percentage points higher than that of White students.

Next, given the statistically significant and large coefficient on the disciplinary history variable in the previous regressions, we explore the sensitivity of the race-based discipline gaps to the inclusion/exclusion of this variable. We also test the robustness of our primary results by including a richer specification of prior disciplinary actions; specifically, we include as controls the number of times that a student was suspended/expelled in prior years and the length of past suspensions/expulsions in prior years within the same school. Columns 2, 3, and 4 of Table 3 reveal these results. We find that while the disciplinary history measure is a predictive student-level characteristic, it does not explain a very large portion of the observed Black-White discipline gaps independently. In other words, students' disciplinary histories do not appear to affect students differentially across race. We substantiate these results in our decomposition analysis in the next section.

Column 5 includes race-specific time trends for Black students given that our descriptive results showed an overall downward time trend in adverse disciplinary rates. The Black-White gap attenuates by a small extent at the prekindergarten/kindergarten, elementary, and middle school levels. However, the magnitude of the Black-White gap is higher by 2 percentage points at the high school level. Several states have adopted policies to limit the use of suspensions and expulsions only to serious infractions at lower grade levels. While we are not aware of any such official policy adoption in Indiana, we do observe downward time trends affecting Black students favorably to some minor extent in the lower grade levels, potentially due to greater awareness of such policies and the negative effects of exclusionary discipline among educators in recent years. Finally, in column 6, we limit the analysis to ninth graders in the high school models to explore the sensitivity of race-based discipline gaps to students dropping out of school in subsequent grades. The magnitudes of the racial disciplinary gaps are marginally higher, potentially indicating that differential dropout rates by race may attenuate racial disciplinary gaps at the high school level.[9]

Column 7 reports results from models that limit the analysis to only those students who receive at least 1 day of suspension. As expected, the magnitudes of these race gaps are larger than those reported in Table 2, because a large number of students have 0 days of suspension. Column 8 controls for the total number of prior disciplinary infractions. Together, these models show that Black-White gaps in the length of suspension/expulsion arise from multiple sources: the reduced likelihood of White children to ever be suspended, to receive multiple suspensions, and to be suspended for longer durations conditional on being suspended.

We also examine the robustness of our results to the exclusion of charter schools[10] and by exploring alternative fixed effects specifications. We ran models including Grade × Year fixed effects, School (or District) × Year fixed effects, and School (or District) × by Grade fixed effects (results not shown). School (or District) × Grade fixed effects account for variation in school disciplinary outcomes that vary by grade within schools or districts (e.g., due to policies that vary behavioral standards by grade). Results from these alternative specifications were qualitatively similar to those reported in Table 2 and are available upon request.

How Does Racial Sorting Across Districts Explain Racial Disciplinary Gaps?

The fixed effects in Equation 3 capture systematic differences across schools (districts) in average suspension/expulsion outcomes that are constant over time and therefore subsume any time-invariant discriminatory behavior at the school (district) levels. Thus, the effect of discriminatory school (district) actions that do not vary over time cannot be identified separately from other time-invariant school (district) features that also explain variation in adverse disciplinary events. For example, a positive coefficient on a school fixed effect captures not only potentially discriminatory attitudes among school administrators that do not vary over time but

TABLE 3

Within-District Race-Based Gaps: Additional Robustness Checks

	Ever suspended / expelled						Length of suspension / expulsion, days	
	(1) Greater than 1 suspension or expulsion	(2) No disciplinary history controls	(3) Count of past disciplinary infractions controlled	(4) Length of past suspensions / expulsion controlled	(5) Race-specific time trend included for Black students	(6) Only ninth-grade students in the high school sample	(7) Length of suspension >0 days	(8) Length of suspension >0 days + count of suspensions / expulsions controlled
Prekindergarten-kindergarten								
Black	0.012*** (0.002)	0.024*** (0.003)	0.023*** (0.003)	0.023*** (0.003)	0.022*** (0.005)		0.513*** (0.098)	0.254*** (0.052)
Hispanic	−0.003 (0.001)	−0.008*** (0.001)	−0.008*** (0.001)	−0.008*** (0.001)	−0.008*** (0.001)		−0.321*** (0.127)	0.016 (0.067)
Other race	0.003*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.005*** (0.001)		0.265*** (0.089)	0.223*** (0.061)
Included other covariates		×	×	×	×		×	×
Year fixed effects	×	×	×	×	×		×	×
School fixed effects								
School district fixed effects	×	×	×	×	×		×	×
Observations, *n*			595,071				10,485	
Grades 1–8								
Black	0.047*** (0.005)	0.089*** (0.006)	0.079*** (0.005)	0.085*** (0.005)	0.092*** (0.011)		0.612*** (0.059)	0.097* (0.038)
Hispanic	−0.008*** (0.001)	−0.014*** (0.002)	−0.012*** (0.002)	−0.013*** (0.002)	−0.015*** (0.002)		−0.146*** (0.041)	0.022 (0.037)
Other race	0.001*** (0.001)	0.017*** (0.003)	0.015*** (0.002)	0.016*** (0.003)	0.017*** (0.003)		0.262*** (0.050)	0.074** (0.031)
Included other covariates		×	×	×	×		×	×
Year fixed effects	×	×	×	×	×		×	×
School fixed effects								
School district fixed effects	×	×	×	×	×		×	×
Observations, *n*			4,471,416				360,318	
High school								
Black	0.071*** (0.008)	0.120*** (0.009)	0.106*** (0.007)	0.116*** (0.008)	0.137*** (0.016)	0.126*** (0.010)	0.887*** (0.098)	0.283*** (0.052)
Hispanic	0.004 (0.003)	0.012** (0.004)	0.011*** (0.003)	0.012** (0.004)	0.012** (0.004)	0.014** (0.005)	−0.050 (0.059)	−0.067 (0.050)
Other race	0.012*** (0.003)	0.019*** (0.005)	0.017*** (0.004)	0.018*** (0.005)	0.012** (0.005)	0.018* (0.006)	0.306*** (0.062)	0.116** (0.037)
Included other covariates		×	×	×	×	×	×	×
Year fixed effects	×	×	×	×	×	×	×	×
School fixed effects								
School district fixed effects	×	×	×	×	×	×	×	×
Observations, *n*			2,157,265			575,367	319,128	

Note. Standard errors in parentheses are clustered at the district level (except in columns 3 and 7, where they are clustered at the school level) to correct for heteroskedasticity and autocorrelation. The unit of observation is a student-year. Length of suspension or expulsion censored to 20 days to limit influence of potential outliers. To economize on space, we report race coefficients only. Alternate specifications include an extensive set of student- and school-level covariates as specified in the main text, school (or district) fixed effects, and year fixed effects.
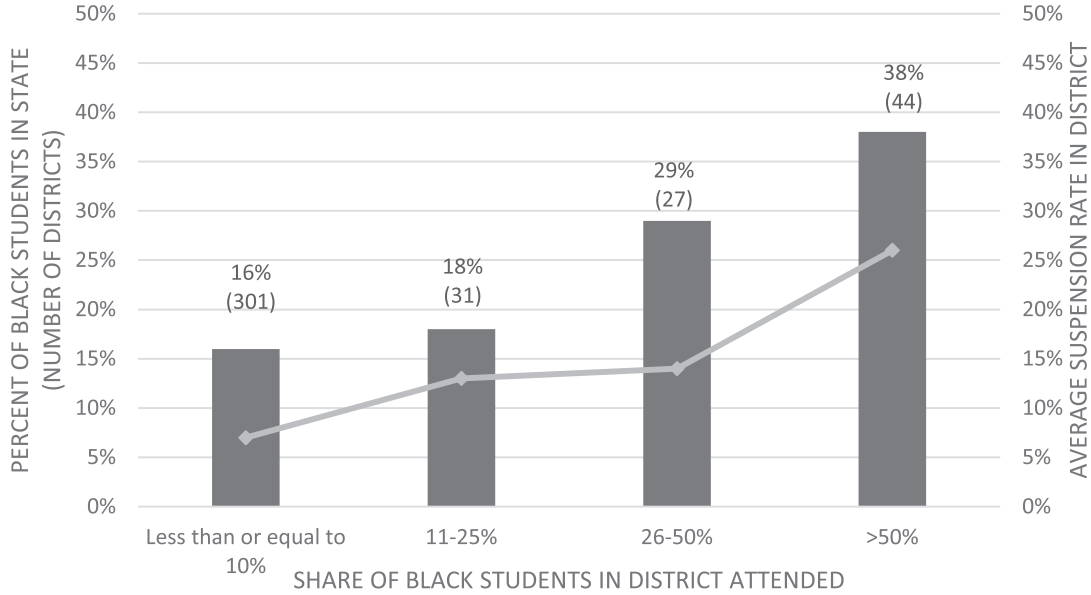*p < .05. **p < .01. ***p < .001.

11

FIGURE 3. *Distribution of Black students across districts with varying shares of Black students. The number of districts is shown in parentheses. For example, the leftmost column indicates the share of Black students in Indiana (16%) who attended school districts in which <10% of enrolled students were Black. The secondary Y-axis and the line graph show the average suspension/expulsion rate in those districts.*

also other structural features of the school that systematically contribute to higher suspensions/expulsions on average (e.g., the location of the school in an area with high crime rates or a zero-tolerance disciplinary policy).

Approximately 82% of all Black students in Indiana attend a school in one of 35 districts (out of nearly 400 districts), and those 35 districts account for 51% of all suspensions or expulsions in the state but only 38% of total enrollment. Figure 3 depicts the distribution of Black students across districts by the share of Black students in the district, averaged across the 2008–2009 and 2013–2014 academic years. The columns indicate the share of Black students in Indiana who attended school districts in which the average share of enrolled Black students in the district is equal to the column label. For example, the leftmost column indicates the share of Black students in Indiana (16%) who attended one of the 301 school districts in Indiana in which <10% of enrolled students were Black. The secondary *Y*-axis depicts the average suspension/expulsion rate in those districts.

The racial composition of districts in Indiana is uneven. Sixteen percent of Black students in Indiana attend schools in districts in which the share of Black students is less than10%, while 38% of Black students attend schools within majority-Black districts. Furthermore, there is little variation over time in the share of Black students within school districts, indicating the entrenched nature of segregation in our sample.

In this final section, we examine, using decomposition techniques, the extent to which the cross-district racial sorting of students explains race-based discipline gaps. We carry out a Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973) to determine how much of the Black disadvantage in exclusionary discipline can be attributed to differences in the districts in which Black and White students attend schools. In the counterfactual decomposition, we simulate the level of adverse disciplinary outcomes that we would expect Black students to experience if they exhibited sensitivities to the underlying covariates that are identical to those of White students. We decompose the discipline gap into two components: the component due to differences in the distribution of observed covariates by race (the "endowment effect") and the component due to differential race-based sensitivities to the underlying covariates and other unobserved heterogeneity (the "coefficient effect").

In Equation 4, *Y* indicates whether a student experiences any suspension/expulsion in a particular year. For the subsamples of Black students, denoted by subscript *A*, and White students, denoted by *B*, we obtain coefficient estimates $\widehat{\beta_A}$, $\widehat{\beta_B}$, $\widehat{\theta_A}$, $\widehat{\theta_B}$, corresponding to the linear probability model defined in Equation 3 separately for the subsamples of Black and White students. The difference in expected suspension/expulsion rates between Black and White students may be defined as follows:

$$\overline{Y_A} - \overline{Y_B} = \left(\overline{X_A} - \overline{X_B}\right)' \widehat{\beta_B} + \overline{(D_A} - \overline{D_B})'\widehat{\theta_B} + \left(\overline{X_A}\right)'\left(\widehat{\beta_A} - \widehat{\beta_B}\right) + \left(\overline{D_A}\right)'\left(\widehat{\theta_A} - \widehat{\theta_B}\right). \tag{4}$$

TABLE 4

Black-White Discipline Gap Decomposition Analysis (Based on Likelihood of Suspension/Expulsion as Dependent Variable)

Ever suspended / expelled	Discipline gap	No school district fixed effects		School district fixed effects included		
		Endowment effect	Coefficient effect	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp
Prekindergarten-kindergarten	0.038*** (100)	0.014* (37)	0.025*** (63)	0.024*** (63)	0.015*** (37)	25
Grades 1–8	0.143*** (100)	0.077*** (54)	0.066*** (46)	0.092*** (65)	0.051*** (35)	11
High school	0.191*** (100)	0.100*** (53)	0.091*** (47)	0.126*** (66)	0.065*** (34)	13

Note. Discipline gap expressed as percentages in parentheses. Standard errors (not reported) are clustered at the district level to correct for heteroskedasticity and autocorrelation. pp = percentage points.
*$p < .05$. ***$p < .001$.

TABLE 5

Hispanic-White Discipline Gap Decomposition Analysis (Based on Likelihood of Suspension/Expulsion as Dependent Variable)

Ever suspended / expelled	Discipline gap	No school district fixed effects		School district fixed effects included		
		Endowment effect	Coefficient effect	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp
Prekindergarten-kindergarten	0.002 (100)					
Grades 1–8	0.020*** (100)	0.017*** (85)	0.003 (15)	0.019*** (95)	0.001 (5)	10
High school	0.077*** (100)	0.059*** (77)	0.017*** (23)	0.064*** (83)	0.012*** (17)	5

Note. Discipline gap expressed as percentages in parentheses. Standard errors (not reported) are clustered at the district level to correct for heteroskedasticity and autocorrelation. The raw discipline gap is close to zero for the prekindergarten-kindergarten level; therefore, we do not decompose the gap further into the endowment and coefficient effects. pp = percentage points.
***$p < .001$.

To determine the share of the racial discipline gap attributable to the nonrandom sorting of Black and White students across districts, we decompose each race gap into the portion that can be explained by differences in the covariates. We report the portion of each gap incrementally attributable to district fixed effects $(\overline{D}_A - \overline{D}_B)' \hat{\theta}_B$ in column 6 of Tables 4 and 5. In our analysis, we use the White student subsample to estimate the average penalties associated with exclusionary discipline. In other words, differences in the districts attended by Black and White students explain 11% to 25% of the Black-White discipline gaps across grade levels, even after controlling for an extensive set of student- and school-level characteristics. That is, Black students appear to be disadvantaged by nonrandom sorting across districts, above and beyond the disadvantage associated with racial differences in observable student- and school-level characteristics. In contrast, a significant portion of the Hispanic-White gap is explained by observable differences in student- and school-level factors even though the overall magnitude of these gaps is considerably lower.

We argue that there are two possible explanations for the contribution of the district fixed effects to the Black-White discipline gaps. The first explanation is that the gap is due to the disproportionate enrollment of Black students into districts with punitive disciplinary environments. The second possibility is that these estimated fixed effects subsume the influence of omitted, unobserved student and/or neighborhood attributes. Our models control for several characteristics. But if, for example, Black students exhibit higher behavioral problems than White students with similar disciplinary histories in ways that we have not observed but select into districts that place a high emphasis on exclusionary discipline, then the causal role of the district will be overestimated. We suspect that the true interpretation lies somewhere in between these two extremes. We also include supplemental district-level analyses to investigate how the effects of de facto segregation across (and within) districts may be associated with school disciplinary outcomes through mechanisms beyond simple racial and socioeconomic composition effects (see appendix for details). Note

that we are unable to attribute directly the sources of racial gaps to the districts per se.

Additionally, our decompositions reveal that about a third of the Black-White discipline gap remains unexplained due to unobserved heterogeneity, even after controlling for multiple student-level characteristics. In other words, if Black students exhibited the same average characteristics as White students (i.e., if Black students were on average as poor and had comparable past disciplinary histories) and were exposed to schools (districts) exhibiting similar average characteristics, the observed Black-White gap would attenuate but would not be eliminated. We argue that this finding provides suggestive evidence of differential treatment discrimination, which may be subsumed in omitted variables not captured in the model. Relatedly, we argue that our specifications that account for multiple student-level characteristics provide an upper bound estimate of the endowment effect because the fixed effects also capture some time-invariant unobserved heterogeneity (also see Elder, Goddeeris, & Haider, 2010).

Conclusion and Policy Implications

Our analysis of statewide, administrative data in Indiana adds to the body of research on racial discipline gaps by focusing on six large cohorts of students across all grades in a large state. We precisely estimate disciplinary gaps for students in two minority groups and look beyond dichotomous measures of exclusionary discipline to uncover a more complicated pattern. Data limitations restrict the ability of many prior studies to investigate disparities across multiple measures of school discipline outcomes and estimate precise race-based discipline gaps.

We demonstrate that Black-White disciplinary gaps emerge as early as in prekindergarten, widen with grade progression, and persist even after controlling for an extensive set of student- and school-level characteristics. Our primary descriptive analysis reveals that between 2008–2009 and 2013–2014, more than 5% of Black children enrolled in Indiana prekindergarten or kindergarten programs experienced a suspension or expulsion, as opposed to about 1% of White children. This Black-White disparity continues to increase through high school, with approximately 34% of Black high school students experiencing a suspension/expulsion in any year, in contrast to 13% of White high school students. We find that the estimated Black-White discipline gaps attenuate by about half after including a comprehensive set of control variables and school or district fixed effects. Additionally, we find that while a student's disciplinary history is an important determinant of the likelihood of future exclusionary discipline, it does not explain a large portion of the Black-White disciplinary gap.

Hispanic-White disciplinary gaps, while lower in magnitude across grade levels when compared with Black-White gaps, show minor increases across grade levels as well; however, almost all of the gaps can be explained by school or district fixed effects and student-level control variables. We view these findings as a substantive contribution to the literature given the conflicting data on Hispanic-White disciplinary gaps in prior studies.

We also identify a key mechanism underlying this disparity: the nonrandom sorting of Black students into school districts with higher suspension/expulsion rates. Is it then possible to achieve some reduction in Black-White discipline gaps by redistributing children across schools? Our estimates suggest that if Black students re-sorted into districts predominantly attended by White students in the state and received disciplinary sanctions at the average rates of those districts, the Black-White disciplinary gap would decline by 11% to 25% across grade levels. The uneven distribution of Black students across districts likely corresponds to the uneven distribution of resources, teacher quality, and other factors that are associated with adverse disciplinary outcomes. Furthermore, it is possible that minority students who attend schools with more minority (or poor) students exhibit similar behaviors but are treated differently. For example, schools with more minority and poor students may experience a larger presence of school resource officers, which may increase the likelihood of punitive disciplinary action conditional on behavior. Future research should also seek to understand the impact of peer effects on student misbehavior and subsequent disciplinary outcomes and whether those effects differ across race.

Nevertheless, as emphasized earlier, even if we were able to achieve such a re-sorting across schools, significant Black-White discipline gaps would persist *within* districts (schools). Our estimates suggest that even after controlling for individual characteristics, disciplinary history, school-level factors, and the district in which the student attends school, within-district Black-White discipline gaps remain across a range of outcomes. Although the magnitudes of these discipline gaps are larger for boys across grade levels, we observe similar patterns in covariate-adjusted and within-district gaps across gender. We also observe the pervasive nature of the Black-White discipline gaps within districts. For example, Black students are more likely than White students to ever be suspended/expelled or to receive multiple suspensions. Furthermore, Black students are likely to receive longer suspensions and expulsions unconditionally as well as conditional on being suspended; this finding persists across models with and without additional controls for the number of prior disciplinary infractions.

Despite extensive covariate adjustment, about a third of the observed Black-White discipline gap remains unexplained

due to unobserved heterogeneity. One potentially important source of unobserved heterogeneity is student behavior. Unfortunately, our data do not capture information on student behavior, and so we are unable to examine whether disciplinary gaps may be attributed to actual differences in student behavior. Such data on objective measures of student behavior are rarely available for research. Thus, we recommend that schools and districts implement improved, innovative data collection techniques that will allow administrators and researchers to examine how differences in student behavior contribute to observed discipline gaps. For example, researchers could access the corpus of classroom video collections available from the Measures of Effective Teaching project to develop objective measures of student behavior as well as the ways in which teachers and administrators translate those behaviors to infractions that result in exclusionary discipline.

What policy prescriptions can we recommend to schools, districts, and states given our empirical analysis? First, our analysis reveals considerable heterogeneity across grade levels in the magnitude of race-based disciplinary gaps—a finding that we hope will inform the design and focus of future interventions and policy reforms aimed at reducing punitive disciplinary practices. Given the emergence of Black-White gaps in exclusionary discipline as early as in prekindergarten—a significant portion of which can be attributed to racial sorting across districts in early grade levels—standardization of disciplinary policies that limit the use of exclusionary discipline at lower grade levels across the state could lead to meaningful reductions in exclusionary disciplinary rates for all students and especially Black students attending school in highly punitive districts. For example, several states and school districts (e.g., Texas, District of Columbia, Denver, Michigan, and Chicago Public Schools) have introduced policies to eliminate exclusionary discipline at lower grade levels; such policies could be adopted more widely.

Second, at higher grade levels, we recommend reenvisioning school discipline. For example, it may be beneficial to enact policy-level changes that limit the use of exclusionary discipline to serious violations only, in addition to providing targeted supports to build alternative restorative justice programs within schools and districts exhibiting high exclusionary discipline rates. We draw on our empirical results regarding cross-district variation in exclusionary discipline as well as the growing literature on these topics to motivate these policy implications. For example, when the School District of Philadelphia banned the use of exclusionary discipline district-wide, researchers observed poor fidelity in implementing this policy, especially in disadvantaged schools with high minority student compositions. As a result, these policies resulted in minimal impacts on racial discipline gaps (Steinberg & Lacoe, 2018). Thus, policy reforms must be accompanied by robust monitoring of implementation

fidelity as well as supports to ensure that teachers and school administrators use alternatives to punitive disciplinary practices. Professional development for teachers and administrators must strive to include evidence-based, restorative justice programs that are theoretically informed and appreciate the inherent complexity of this issue (Welsh & Little, 2018).

Third, none of our models provide evidence against the potential presence of differential treatment discrimination, which may arise due to the differential response of teachers and school administrators to behaviors of Black students. Given that about a third of the Black-White discipline gap remains unexplained even after extensive covariate adjustment, policy makers must continue investigating potential cases of differential treatment discrimination within schools and districts. Future research should also strive to disentangle the potentially differential mechanisms that may drive racial disparities in school discipline across students of different minority races differentially, given our findings regarding the low magnitudes of Hispanic-White discipline gaps.

Finally, continued collection and dissemination of robust, disaggregated, student-level data on disciplinary outcomes is paramount to move research forward in this area. In all, policy prescriptions must embrace the complexity of cross- and within-school (district) Black-White disciplinary gaps. As we point out throughout our article, potential sources of racial discrimination exist within as well as across schools and districts. Therefore, a more pragmatic approach would be to understand the sources of these observed racial disparities and design policy prescriptions and interventions that go beyond a one-size-fits-all approach.

Appendix

Segregation and School Discipline: District-Level Analysis

In this section, we link each district-year record from the IDOE data to district-level data from the Stanford Education Data Archive (SEDA; Reardon et al., 2017). The SEDA includes a comprehensive set of district-level covariates, such as parental education levels, income, and unemployment levels compiled from the Common Core of Data and the School District Demographic System, which is a special tabulation of the 2006–2010 American Community Survey data.

Following Reardon (2016), we calculate several segregation measures. We then examine the associations between those measures and school discipline outcomes. Although the exposure and unevenness measures are highly related, they reflect distinct mechanisms through which segregation may influence suspensions and expulsions. For example, if school or district racial and socioeconomic composition affects the likelihood of suspension/expulsion for all students equally, then the exposure measures of segregation

will be associated with suspension/expulsion rates only to the extent that students' schools or neighborhoods differ by race in their racial and socioeconomic compositions. In other words, if adverse disciplinary outcomes among Black and White students demonstrate similar sensitivities to school and district composition, then unevenness in the racial distribution of students across schools will be more strongly associated with discipline gaps. Alternatively, if the concentrations of minority or poor students within schools and districts differentially affect the probability of a suspension or expulsion for Black and White students, then exposure measures may be more strongly associated with suspension/expulsion rates. These two segregation mechanisms imply different policy strategies for remedying the discipline gap.

Table A9 reports associations between suspension and expulsion rates and our segregation measures for all district-years in our sample. We first regress district-level suspension/expulsion rates on various measures of the share of Black students in the district (see columns 1–3 in Table A9). We find that a percentage point increase in the share of Black students is associated with a 0.23–percentage point increase in suspension and expulsion rates. Using a less parametric measure of district racial composition (column 2), we find that districts in the top quartile of Black student enrollment exhibit suspension/expulsion rates approximately 8 percentage points higher than those of districts in the bottom quartile. In another specification (not shown), we find a negative and statistically significant coefficient on the square of Black students, indicating that the estimated effect on suspensions and expulsions decreases as the share of Black students in the district increases. The effect of Black enrollment on suspensions and expulsions attenuates substantially after controlling for other district- and neighborhood-level covariates (see column 3), indicating the uneven distribution of student and school attributes across Black and White students. After inclusion of additional covariates, the coefficient on the share of free/reduced-price lunch students in the district is considerably larger than the coefficient on the share of Black students in the district.

Because segregation may be associated with adverse disciplinary outcomes through mechanisms beyond simple racial composition measures, we next explore the relationship between exposure and unevenness measures and adverse disciplinary outcomes. Column 4 of Table A9 reports bivariate correlations between suspension/expulsion rates and measures of racial and socioeconomic composition, exposure, and unevenness. The correlations range from –0.04 to 0.57. First, with the exception of the unevenness in exposure of Black and White students to poverty, all segregation measures are positively associated with adverse disciplinary outcomes. Second, each measure of exposure is more highly correlated with suspension and expulsion rates than the corresponding unevenness measure. This finding underscores our earlier analysis in Table 3, demonstrating that the probability of a suspension/ expulsion for Black students is more sensitive to peer composition than the same probability for White students. First, Black students are more unevenly distributed across districts, with some districts having a larger share of minority and poor students; those same districts also exhibit higher suspension and expulsion rates. While the Black-White unevenness in exposure to poor students is negatively correlated (not significant) with suspension and expulsion outcomes, Black-White unevenness in exposure to Black students is significantly and positively correlated with adverse disciplinary outcomes.

Columns 5 and 6 of Table A9 report results from multivariate regression models that include all exposure and unevenness measures, respectively. Column 7 reports results from a multivariate model that includes all measures, and column 8 adds a comprehensive set of district-level covariates, including per-pupil expenditures, the log of district enrollment, and a composite measure of socioeconomic status obtained from the SEDA (calculated from Common Core of Data and School District Demographic System data sources and with shares of unemployed, low-income, SNAP beneficiary [Supplemental Nutrition Assistance Program], and single-mother households). Interpreting results from these multivariate models is not straightforward due to high levels of correlation between the segregation measures (see Table A2). Overall, we find that similar patterns persist in the multivariate models. In particular, within-district measures of Black students' exposure to other Black and poor students is strongly associated with suspension/expulsion rates. In short, policy proposals that aim to reduce segregation also may reduce racial disparities in suspension and expulsion outcomes.
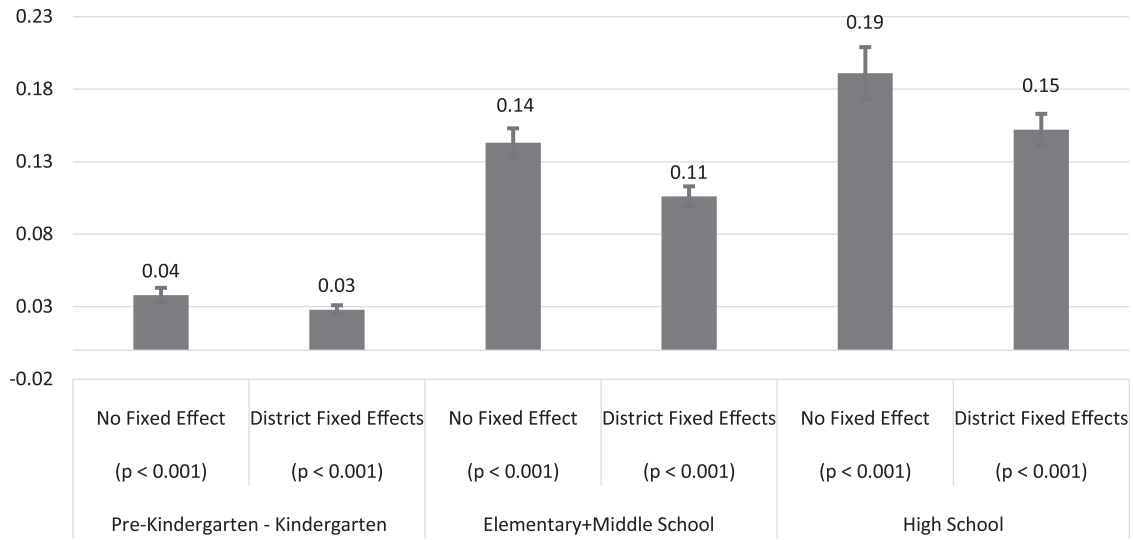
FIGURE A1. *The Black-White discipline gap attenuates little with the inclusion of district fixed effects. Coefficient on Black indicator from models including only race indicators and year fixed effects. Here, we report coefficients estimated with and without school district fixed effects for models estimated separately by school grade level. Error bars represent standard errors that are clustered at the school district level to correct for heteroskedasticity and autocorrelation.*
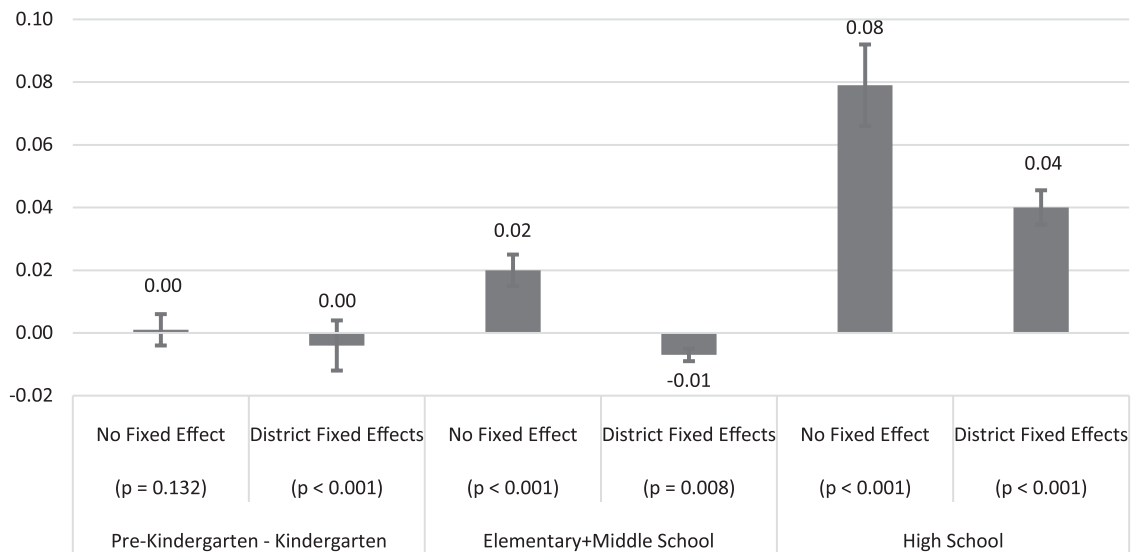
FIGURE A2. *The Hispanic-White discipline gap attenuates to a large extent with the inclusion of district fixed effects. Coefficient on Hispanic indicator from models including only race indicators and year fixed effects. Here, we report coefficients estimated with and without school district fixed effects for models estimated separately by school grade level. Error bars represent standard errors that are clustered at the school district level to correct for heteroskedasticity and autocorrelation.*

TABLE A1

Descriptive Statistics of District-Level Segregation Measures and Covariates

	Mean (*SD*)
Within district	
Black exposure to Black students	0.050 (0.115)
Black exposure to poor students	0.444 (0.144)
Black-White exposure to Black students	0.012 (0.030)
Black-White exposure to poor students	0.015 (0.042)
Per-pupil expenditures, $	10,349 (1,480)
District enrollment (logged)	7.797 (0.880)
Socioeconomic status composite	0.073 (0.710)
Observations (district-years), *n*	1,436

Note. The entries are means (that have been time de-meaned) and standard deviations (overall across the panel) of district-level panel data. The unit of observation is a district-year.

TABLE A2

Measures of Segregation: Correlation Matrix

	Within district		
Within district	Black exposure to Black students	Black exposure to poor students	Black-White exposure to Black students
Black exposure to Black students	—		
Black exposure to poor students	.593	—	
Black-White exposure to Black students	.254	.097	—
Black-White exposure to poor students	.003	.080	.419

TABLE A3

Sensitivity of Adverse Disciplinary Outcomes (Ever Suspended/Expelled) to Covariates, Partitioning Sample by Race

	Full sample	Black	White	Hispanic	Other race
		Prekindergarten-kindergarten			
Black	0.024*** (0.003)				
Hispanic	−0.008*** (0.001)				
Other race	0.006*** (0.001)				
Special education status	0.008*** (0.001)	0.007* (0.003)	0.007*** (0.001)	0.011*** (0.003)	0.013*** (0.003)
FRPL status	0.009*** (0.001)	0.013*** (0.003)	0.009*** (0.001)	0.002* (0.001)	0.013*** (0.002)
Chow test statistic for FRPL status		2.92* (.088)	0 (.969)	46.64*** (0)	6.42* (.011)
Proportion of FRPL students in school	0.026*** (0.005)	0.076*** (0.014)	0.015*** (0.004)	0.01 (0.92)	0.029*** (0.005)
Chow test statistic for proportion of FRPL students in school		21.70*** (0)	34.16*** (0)	0.01 (.92)	0.35 (.557)
Included other covariates	×	×	×	×	×
Year fixed effects	×	×	×	×	×
Observations, *n*	595,071	77,849	406,306	66,956	43,960
		Grades 1–8			
Black	0.077*** (0.006)				
Hispanic	−0.011*** (0.002)				
Other race	0.016*** (0.003)				
Special education status	0.031*** (0.001)	0.027*** (0.004)	0.030*** (0.001)	0.024*** (0.002)	0.036*** (0.003)
FRPL status	0.042*** (0.001)	0.054*** (0.005)	0.041*** (0.001)	0.015*** (0.002)	0.042*** (0.002)
Chow test statistic for FRPL status		11.41*** (0)	1.31 (.253)	176.38*** (0)	0.93 (.335)
Proportion of FRPL students in school	0.062*** (0.011)	0.190*** (0.036)	0.054*** (0.008)	0.057*** (0.012)	0.082*** (0.015)
Chow test statistic for proportion of FRPL students in school		17.95*** (0)	9.22** (.002)	0.01 (.92)	5.74* (.017)

(continued)

	Full sample	Black	White	Hispanic	Other race
Included other covariates	×	×	×	×	×
Year fixed effects	×	×	×	×	×
Observations, *n*	4,471,416	571,572	3,185,770	415,309	298,765
		High school			
Black	0.103*** (0.008)				
Hispanic	0.009** (0.003)				
Other race	0.019*** (0.004)				
Special education status	0.053*** (0.002)	0.053*** (0.008)	0.051*** (0.002)	0.052*** (0.005)	0.072*** (0.006)
FRPL status	0.072*** (0.003)	0.082*** (0.007)	0.070*** (0.002)	0.041*** (0.005)	0.073*** (0.004)
Chow test statistic for FRPL status		3.58* (.061)	0.15 (.69)	68.30*** (0)	2.81* (.09)
Proportion of FRPL students in school	0.217*** (0.029)	0.334*** (0.070)	0.183*** (0.026)	0.276*** (0.041)	0.225*** (0.041)
Chow test statistic for proportion of FRPL students in school		4.91* (.027)	9.81** (.002)	4.88* (.027)	0.36 (.54)
Included other covariates	×	×	×	×	×
Year fixed effects	×	×	×	×	×
Observations, *n*	2,157,265	270,850	1,617,157	151,397	117,861

Note. The first column replicates the results reported in column 2 of Table 2. The other columns report coefficient estimates from identical specifications but estimated separately for subsamples partitioned by race. The unit of observation is a student-year. To economize on space, we report relevant coefficients only. All specifications include the full set of student- and school-level covariates and year fixed effects. Standard errors in parentheses are clustered at the district level to correct for heteroskedasticity and autocorrelation. The *p* values are provided in parentheses under the Chow test statistic. FRPL = free/reduced-price lunch.
*p < .05. **p < .01. ***p < .001.

TABLE A4

Sensitivity of Adverse Disciplinary Outcomes (Ever Suspended/Expelled) to Covariates, Partitioning Sample by Race, Inclusion of Student Fixed Effects

Ever suspended/expelled	Full sample	Black	White	Hispanic	Other race
		Prekindergarten-kindergarten			
Special education status	0.005 (0.004)	0.007* (0.003)	0.008*** (0.001)	0.011*** (0.003)	0.015*** (0.003)
FRPL status	0.003 (0.009)	0.009*** (0.003)	0.009*** (0.001)	0.002* (0.001)	0.013*** (0.002)
Proportion of FRPL students in school	−0.027* (0.009)	0.011 (0.017)	0.016*** (0.005)	0.007 (0.005)	0.013* (0.006)
Included other covariates	×	×	×	×	×
Fixed effects (year, district, and student)	×	×	×	×	×
Observations, *n*	595,071	77,849	406,306	66,956	43,960
		Grades 1–8			
Special education status	0.012*** (0.001)	0.010* (0.004)	0.010*** (0.001)	0.010* (0.001)	0.014*** (0.005)
FRPL status	0.009*** (0.001)	0.010*** (0.005)	0.001*** (0.001)	0.005** (0.002)	0.004 (0.001)
Proportion of FRPL students in school	0.026* (0.005)	0.007 (0.047)	0.046*** (0.009)	0.013 (0.009)	0.022 (0.017)
Included other covariates	×	×	×	×	×
Fixed effects (year, district, and student)	×	×	×	×	×
Observations, *n*	4,471,416	571,572	3,185,770	415,309	298,765
		High school			
Special education status	0.063*** (0.004)	0.082*** (0.009)	0.061*** (0.005)	0.062*** (0.013)	0.037* (0.018)
FRPL status	0.026*** (0.003)	0.033*** (0.006)	0.023*** (0.002)	0.027*** (0.005)	0.028*** (0.005)
Proportion of FRPL students in school	0.031 (0.028)	0.004 (0.081)	0.028 (0.025)	0.039 (0.065)	0.035 (0.041)
Included other covariates	×	×	×	×	×
Fixed effects (year, district, and student)	×	×	×	×	×
Observations, *n*	2,157,265	270,850	1,617,157	151,397	117,861

Note. The first column reports results for the full sample. The other columns report coefficient estimates from identical specifications but estimated separately for subsamples partitioned by race. The unit of observation is a student-year. To economize on space, we report relevant coefficients only. All specifications include the full set of student- and school-level covariates and year, school district, and student fixed effects. Standard errors in parentheses are clustered at the school district level to correct for heteroskedasticity and autocorrelation.
*p < .05. **p < .01. ***p < .001.

TABLE A5

Race-Based Gaps: Additional Robustness Checks

	Ever suspended/expelled							
	Logistic regression relative marginal effects (odds ratio)	Logistic regression relative marginal effects marginal effects (odds ratio)	Number of years of free/reduced-price lunch included	Grade fixed effects included	10th-grade students only in high schools	11th-grade students only in high schools	12th-grade students only in high school	Count of suspensions / expulsions
	Prekindergarten–kindergarten							
Black	0.041*** (4.48)	0.009*** (2.25)	0.023*** (0.003)	0.039*** (0.001)				0.056*** (0.008)
Hispanic	0.002 (1.15)	−0.003*** (0.66)	−0.008*** (0.001)	0.002*** (0.001)				−0.020*** (0.003)
Other race	0.013*** (1.94)	0.004*** (1.45)	0.005*** (0.001)	0.011*** (0.001)				0.010*** (0.003)
Included other covariates		×	×	×				×
Year fixed effects	×	×	×	×				×
School district fixed effects			×					×
Grade fixed effects				×				
Observations, *n*	595,071	595,071	595,071	595,071				595,071
	Grades 1–8							
Black	0.150*** (4.09)	0.039*** (2.19)	0.087*** (0.006)	0.145*** (<0.001)				0.239*** (0.029)
Hispanic	0.025*** (1.39)	−0.005*** (0.862)	0.002 (0.002)	−0.027*** (<0.001)				−0.034** (0.006)
Other race	0.037*** (1.58)	0.011*** (1.31)	0.020*** (0.003)	0.036*** (<0.001)				0.043*** (0.008)
Included other covariates		×	×	×				×
Year fixed effects	×	×	×	×				×
School district fixed effects			×					×
Grade fixed effects				×				
Observations, *n*	4,471,416	4,471,416	4,471,416	4,471,416				4,471,416
	High school							
Black	0.198*** (3.40)	0.078*** (1.97)	0.123*** (0.008)	0.190*** (0.001)	0.104*** (0.007)	0.098*** (0.007)	0.081*** (0.007)	0.379*** (0.059)
Hispanic	0.089*** (1.86)	0.010*** (1.11)	0.031*** (0.004)	0.077*** (0.001)	0.010* (0.004)	0.008* (0.004)	0.003 (0.003)	0.018 (0.014)
Other race	0.052** (1.48)	0.018*** (1.20)	0.027*** (0.004)	0.044*** (0.001)	0.020** (0.005)	0.015** (0.004)	0.014** (0.003)	0.058*** (0.016)
Included other covariates		×	×	×	×	×	×	×
Year fixed effects	×	×	×	×	×	×	×	×
School district fixed effects			×		×	×	×	×
Grade fixed effects				×				
Observations, *n*	2,157,265	2,157,265	2,157,265	2,157,265	558,110	528,247	494,430	2,157,265

Note. The unit of observation is a student-year. To economize on space, we report relevant coefficients only. The first two columns report results from logistic regressions. We provide the relative marginal effects (i.e., relative to the White students; marginal effects of the coefficients on each race category evaluated at the mean values across all other covariates) and the corresponding odds ratios in parentheses. In columns 3 to 8, we report coefficient estimates from ordinary least squares regressions. Standard errors shown in parentheses are clustered at the school district level to correct for heteroskedasticity and autocorrelation.

*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE A6

Racial Disciplinary Gaps Decomposition Analysis: Males-Only Sample

	Black-White gap				Hispanic-White gap			
Ever suspended / expelled	Discipline gap	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp	Discipline gap	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp
Prekindergarten-kindergarten	0.055*** (100)	0.040*** (73)	0.015*** (27)	44	0.003 (100)			
Grades 1–8	0.169*** (100)	0.094*** (56)	0.074*** (44)	6	0.023*** (100)	0.023*** (100)	0*** (0)	14
High school	0.197*** (100)	0.129*** (65)	0.068*** (34)	15	0.088*** (100)	0.072*** (83)	0.015*** (17)	7

Note. Discipline gap expressed as percentages in parentheses. Standard errors (not reported) are clustered at the district level to correct for heteroskedasticity and autocorrelation. pp = percentage points.
***$p < .001$.

TABLE A7

Racial Disciplinary Gaps Decomposition Analysis: No Disciplinary History (District Fixed Effects Included)

	Black-White gap				Hispanic-White gap			
Ever suspended / expelled	Discipline gap	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp	Discipline gap	Endowment effect	Coefficient effect	Incremental share explained by district fixed effects, pp
Prekindergarten-kindergarten	0.038*** (100)	0.024* (62)	0.015*** (38)	26	0.002 (100)			
Grades 1–8	0.143*** (100)	0.088*** (61)	0.056*** (39)	13	0.020*** (100)	0.020*** (100)	0*** (0)	11
High school	0.191*** (100)	0.120*** (63)	0.071*** (37)	16	0.077*** (100)	0.063*** (83)	0.013*** (17)	8

Note. Discipline gap expressed as percentages in parentheses. Standard errors (not reported) are clustered at the district level to correct for heteroskedasticity and autocorrelation. The raw Hispanic-White discipline gap is close to zero for the prekindergarten-kindergarten level; therefore, we do not decompose the gap further into the endowment and coefficient effects. pp = percentage points.
*$p < .05$. ***$p < .001$.

TABLE A8

Intersectionality: Race-Gender Interactions

	Ever suspended/expelled		Length of suspension/expulsion	
	(1)	(2)	(3)	(4)
	Prekindergarten-kindergarten			
Black male	0.036*** (0.004)	0.035*** (0.004)	0.154*** (0.021)	0.152*** (0.021)
Black	0.020*** (0.003)	0.005* (0.002)	0.068*** (0.010)	0.017* (0.001)
Hispanic male	0.002 (0.001)	0.002 (0.001)	0.002 (0.005)	0.004 (0.005)
Hispanic	0.001 (0.001)	−0.009*** (0.001)	0.002 (0.001)	−0.030*** (0.007)
Included other covariates		×		×
Year fixed effects	×	×	×	×
School district fixed effects		×		×
Observations, n	595,071	595,071	595,071	595,071

(continued)

	Ever suspended/expelled		Length of suspension/expulsion	
	(1)	(2)	(3)	(4)
		Grades 1–8		
Black male	0.052*** (0.006)	0.044*** (0.005)	0.409*** (0.041)	0.350*** (0.037)
Black	0.117*** (0.008)	0.054*** (0.005)	0.604*** (0.047)	0.252*** (0.030)
Hispanic male	0.005 (0.003)	0.004 (0.002)	0.039*** (0.020)	0.031 (0.005)
Hispanic	0.019*** (0.005)	−0.013*** (0.002)	0.087*** (0.025)	−0.073*** (0.014)
Included other covariates		×		×
Year fixed effects	×	×	×	×
School district fixed effects		×		×
Observations, n	4,471,416	4,471,416	4,471,416	4,471,416
		High school		
Black male	0.012 (0.009)	0.016*** (0.006)	0.382*** (0.073)	0.382*** (0.063)
Black	0.185*** (0.017)	0.096*** (0.008)	1.119*** (0.135)	0.546*** (0.061)
Hispanic male	0.023*** (0.004)	0.020*** (0.003)	0.223*** (0.047)	0.204*** (0.041)
Hispanic	0.068*** (0.007)	−0.001 (0.003)	0.288*** (0.054)	−0.120*** (0.034)
Included other covariates		×		×
Year fixed effects	×	×	×	×
School district fixed effects		×		×
Observations, n	2,157,265	2,157,265	2,157,265	2,157,265

Note. Standard errors in parentheses are clustered at the district level to correct for heteroskedasticity and autocorrelation. The unit of observation is a student-year. Length of suspension or expulsion censored to 20 days to limit influence of potential outliers. To economize on space, we report relevant race and interaction coefficients only. Alternate specifications include an extensive set of student- and school-level covariates as specified in the main text, school (or district) fixed effects, and year fixed effects.
*$p < .05$. ***$p < .001$.

TABLE A9

District-Level Analysis: Association Between Suspension/Expulsion Rates and Black-White Segregation Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Simple racial and socioeconomic composition							
Proportion Black students in district	0.233*** (0.033)		0.087*** (0.025)	0.578***				
District in second quartile of proportion Black students (yes = 1, no = 0)		0.001 (0.004)						
District in third quartile of proportion Black students (yes = 1, no = 0)		0.015* (0.006)						
District in upper quartile of proportion Black students (yes = 1, no = 0)		0.082*** (0.009)						
Proportion free/reduced-price lunch students in district			0.129*** (0.026)	0.521***				
Proportion Hispanic students in district			0.086* (0.034)	0.198***				
Year fixed effects	×	×	×		×	×	×	×
Included other covariates[a]			×					×
	Exposure measures							
Black exposure to Black students (within district)					0.573***	0.127*** (0.022)	0.135*** (0.026)	0.101*** (0.027)
Black exposure to poor students (within district)					0.540***	0.165*** (0.018)	0.177*** (0.017)	0.164*** (0.029)
	Unevenness measures							
Black-White exposure to Black students (within district)					0.128***	0.662*** (0.213)	0.001 (0.203)	−0.022 (0.124)
Black-White exposure to poor students (within district)					−0.039	−0.156 (0.115)	−0.187* (0.115)	−0.170* (0.066)
Observations, *n*	1,436	1,436	1,436	1,436	1,436	1,436	1,436	1,436

Note. Standard errors in parentheses are clustered at the school district level to correct for heteroskedasticity and autocorrelation. The unit of observation is a district-year.
[a]Additional covariates from the Common Core of Data and American Community Survey include the per-pupil expenditures, log of district enrollment, and a composite of socioeconomic status constructed at the district level (see text for details).
*p < .05. ***p < .001.

23

TABLE A10
Measures of Segregation

Segregation measures (within district)	Formulae
Exposure	
Black exposure to Black students	$$\dfrac{\sum_1^s (\dfrac{b_{st}}{t_{st}} * b_{st})}{b_{dt}}$$
Black exposure to poor students	$$\dfrac{\sum_1^s (\dfrac{p_{st}}{t_{st}} * b_{st})}{b_{dt}}$$
Unevenness	
Black-White exposure to Black students	$$\dfrac{\sum_1^s (\dfrac{b_{st}}{t_{st}} * b_{st})}{b_{dt}} - \dfrac{\sum_1^s (\dfrac{b_{st}}{t_{st}} * w_{st})}{w_{dt}}$$
Black-White exposure to poor students	$$\dfrac{\sum_1^s (\dfrac{p_{st}}{t_{st}} * b_{st})}{b_{dt}} - \dfrac{\sum_1^s (\dfrac{p_{st}}{t_{st}} * w_{st})}{w_{dt}}$$

Note. b_{dt} = number of Black students in a district each year; b_{st} = number of Black students in a school each year; p_{dt} = number of students who receive free or reduced-price lunch in a district each year; p_{st} = number of students who receive free or reduced-price lunch in a school each year; S = total number of schools in a district each year; t_{dt} = total number of students in a district each year; t_{st} = total number of students in a school each year; w_{dt} = number of White students in a district each year; w_{st} = number of White students in a school each year.

ORCID iD

Maithreyi Gopalan iD https://orcid.org/0000-0002-1013-0672

Notes

1. The race variable included in the administrative data contains a nonmissing value for nearly all student observations. However, we cannot identify the multiple races/ethnicities of multiracial students, so we classify multiracial students into the "other" race category.

2. The latest school guidelines for reporting data on exclusionary discipline can be found at the following IDOE website: https://learningconnection.doe.in.gov/Library/FilingCabinet/ViewFileDetail.aspx?lfid=80366&et=USER_GROUP&eid=975&clid=&ret=~%2fUserGroup%2fGroupDetailFileBookmarks.aspx%3fgid%3d975%26ugfid%3d4504

3. We explore alternative model specifications with outcome variables defined to capture multiple suspensions/expulsion (likelihood of receiving multiple suspensions or expulsions in a year and a count variable indicating number of suspensions/expulsions in a year as alternative dependent variables in Table 3 and Appendix Table A5). Additionally, we explore outcome variables defined in a variety of ways: as any suspension, any out-of-school suspension, any expulsion, and duration of adverse disciplinary outcomes, in days, of suspensions, out-of-school suspensions, and expulsions. The results are qualitatively similar and are available upon request.

4. Another common measure of segregation used in past literature is the dissimilarity index of segregation, which measures spatial segregation and is usually interpreted as the proportion of a minority group that would need to hypothetically move to create a uniform distribution of the overall population. Therefore, the dissimilarity index is an aggregate measure calculated for a higher-order unit, such as city or state, based on racial composition in the lower-order units, such as census tracts or counties within the city or state. For example, in our sample, the Black-White dissimilarity index of segregation for students in Indiana is 0.64—which means that 64% of Black students would need to move across districts to achieve a uniform distribution of Black and White students in each district in Indiana. Given that this is an aggregate measure for the state, it cannot be used in the district- or student-level regressions as there is no variation across students within districts. Alternatively, we could calculate the dissimilarity index for each school district by analyzing the within-district dissimilarity index using variation across schools. However, we need to have enough schools within districts (at least five based on suggestions in prior literature) as well as a minimum number of students (overall as well as within each racial category) in those schools for meaningful interpretation. Hence, we have not included the district-level dissimilarity index in the regressions.

5. From a theoretical perspective, several reasons have been put forth in the larger literature regarding why we might expect age-level differences in discipline and, specifically, larger racial disciplinary gaps at higher grades. For example, Goff et al. (2014) found that Black boys are viewed as older and less innocent than their White peers. Developmentally speaking, we also expect that teenagers may be more likely to question authority and that harsher disciplinary sanctions may be more prevalent in high school. Our study finds higher rates of exclusionary discipline and wider racial disciplinary disparities at higher grade levels.

6. We believe that the most parsimonious model specification is the one with race indicators and year fixed effects only, carried out separately by grade level. The results are near identical when we include grade fixed effects in addition to year fixed effects in the elementary, middle, and high school baseline models (see Appendix Table A5).

7. Our lagged measures of disciplinary history can be considered a noisy proxy for behavioral differences. In the appendix, we show the decompositions without the inclusion of the past disciplinary history variable and for subsamples of male students as additional robustness checks. As mentioned, while the disciplinary history of a student positively predicts exclusionary disciplinary outcomes, it explains only between 0% to 3% of observed Black-White discipline gaps (we obtain this figure by comparing the endowment effects estimated with and without past disciplinary history variables; see Appendix Table A7).

8. Note that the race coefficients in our models cannot be interpreted as evidence of the causal effect of race on adverse disciplinary outcomes. Our models do not account for omitted variables that may be correlated with race and adverse disciplinary outcomes (e.g., the models do not account for time-invariant features of districts or schools, although students sort nonrandomly into district and school environments that differentially affect their disciplinary outcomes). Thus, those sources of unobserved variation reside in the error term and bias the estimated model coefficients.

9. Appendix Table A5 reports racial gaps estimated for 10th, 11th, and 12th graders separately. We find decreasing race-based gaps at higher grade levels within the high school model. Note that, on average, across the years in our sample, the share of Black students decreases from 14% in the 9th grade to 11% in the 12th grade. Similarly, the share of Hispanic students decreases from 7% to 5%. These patterns are suggestive of differential rates of dropout behavior among minority students.

10. Strict "no excuses" disciplinary policies in charter schools and high levels of exclusionary discipline in charter schools are a topic of ongoing scholarly and policy interest. An in-depth exploration of disciplinary outcomes in charter schools is beyond the scope of this article. However, we find that, at least in Indiana, high exclusionary discipline and high racial discipline gaps do not seem to be driven exclusively by charter schools.

References

Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a U.S. state. *Education Policy Analysis Archives*, *25*, 49. doi:10.14507/epaa.25.2787

Anderson, K. P., & Ritter, G. W. (2018). Do school discipline policies treat students fairly? Evidence from Arkansas. *Educational Policy*. Advance online publication. doi:10.1177/0895904818802085

Arcia, E. (2006). Achievement and enrollment status of suspended students. *Education and Urban Society*, *38*(3), 359–369. doi:10.1177/0013124506286947

Billings, S. B., Deming, D. J., & Rockoff, J. (2014). School segregation, educational attainment, and crime: Evidence from the end of busing in Charlotte-Mecklenburg. *Quarterly Journal of Economics*, *129*(1), 435–476. doi:10.1093/qje/qjt026

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human resources*, *8*(4), 436–455.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of black students in office disciplinary referrals. *Journal of Educational Psychology*, *102*(2), 508–520. doi:10.1037/a0018450

Brantlinger, E. (1991). Social class distinctions in adolescents' reports of problems and punishment in school. *Behavioral Disorders*, *17*(1), 36–46.

Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Washington, DC: Author.

Clotfelter, C. T. (2004). *After "Brown": The rise and retreat of school desegregation*. Princeton, NJ: Princeton University Press.

Coleman, J. S., Kelley, S. D., & Moore, J.A. (1975). *Trends in school segregation, 1968–73*. Washington, DC: The Urban Institute.

Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school students. *Journal of School Psychology*, *36*(1), 59–82. doi:10.1016/S0022-4405(97)00050-2

Elder, T. E., Goddeeris, J. H., & Haider, S. J. (2010). Unexplained gaps and Oaxaca-Blinder decompositions. *Labour Economics*, *17*(1), 284–290. doi:10.1016/j.labeco.2009.11.002

Figlio, D. N. (2006). Testing, crime and punishment. *Journal of Public Economics*, *90*(4–5), 837–851. doi:10.1016/j.jpubeco.2005.01.003

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59–68. doi:10.3102/0013189X09357621

Gregory, A., & Weinstein, R. S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology*, *46*(4), 455–475. doi:10.1016/j.jsp.2007.09.001

Hwang, N. (2018). Suspensions and achievement: Varying links by type, frequency, and subgroup. *Educational Researcher*, *47*(6), 363–374. doi:10.3102/0013189X18779579

Irwin, K., Davidson, J., & Hall-Sanchez, A. (2013). The race to punish in American schools: Class and race predictors of punitive school-crime control. *Critical Criminology*, *21*(1), 47–71. doi:10.1007/s10612-012-9171-2

Kinsler, J. (2011). Understanding the Black-White school discipline gap. *Economics of Education Review*, *30*(6), 1370–1383. doi:10.1016/j.econedurev.2011.07.004

Kinsler, J. (2013). School discipline: A source or salve for the racial achievement gap? *International Economic Review*, *54*(1), 355–383. doi:10.1111/j.1468-2354.2012.00736.x

Lacoe, J., & Steinberg, M. P. (2018). Do suspensions affect student outcomes? *Educational Evaluation and Policy Analysis*, *41*(1), 34–62. doi:10.3102/0162373718794897

Lindsay, C. A., & Hart, C. M. D. (2017). Exposure to same-race teachers and student disciplinary outcomes for Black students in North Carolina. *Educational Evaluation and Policy Analysis*, *39*(3), 485–510. doi:10.3102/0162373717693109

Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). *Descriptive analysis in education: A guide for researchers* (No. NCEE 2017-4023). Washington, DC: National Center for Education Evaluation and Regional Assistance.

Losen, D., Hodson, C., Keith, M. A., II, Morrison, K., & Belway, S. (2015). *Are we closing the discipline gap?* Retrieved from https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/are-we-closing-the-school-discipline-gap/AreWeClosingTheSchoolDisciplineGap_FINAL221.pdf

Losen, D. J., & Skiba, R. J. (2010). *Suspended education: Urban middle schools in crisis*. Los Angeles, CA: The Civil Rights Project at UCLA.

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, *67*(2), 281–315. doi:10.2307/2579183

Morris, E. W., & Perry, B. L. (2016, January). The punishment gap: School suspension and racial disparities in achievement. *Social Problems*, 68–86. doi:10.1093/socpro/spv026

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693–709.

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science*, *26*(5), 617–624. doi:10.1177/0956797615570365

Orfield, G., & Lee, C. (2007). *Historic reversals, accelerating resegregation, and the need for new integration strategies*. Retrieved from https://www.civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/historic-reversals-accelerating-resegregation-and-the-need-for-new-integration-strategies-1/orfield-historic-reversals-accelerating.pdf

Peguero, A. A., & Shekarkhar, Z. (2011). Latino/a student misbehavior and school punishment. *Hispanic Journal of Behavioral Sciences*, *33*(1), 54–70. doi:10.1177/0739986310388021

Ramey, D. M. (2018). The social construction of child social control via criminalization and medicalization: Why race matters. *Sociological Forum*, *33*(1), 139–164. doi:10.1111/socf.12403

Rausch, K. M., & Skiba, R. (2006). Discipline, disability, and race: Disproportionality in Indiana schools. *Education Policy Brief*, *4*(10). Retrieved from https://eric.ed.gov/?id=ED495751

Reardon, S., Ho, A., Shear, B., Fahle, E., Kalogrides, D., & DiSalvo, R. (2017). *Stanford Education Data Archive (Version 2.0)*. Retrieved from http://purl.stanford.edu/db586ns4974

Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 34–57. doi:10.7758/rsf.2016.2.5.03

Reardon, S. F., & Owens, A. (2014). 60 years after *Brown*: Trends and consequences of school segregation. *Annual Review of Sociology*, *40*(1), 199–218. doi:10.1146/annurev-soc-071913-043152

Rocha, R. R., & Hawes, D. P. (2009). Racial diversity, representative bureaucracy, and equity in multiracial school districts. *Social Science Quarterly*, *90*(2), 326–344.

Rocque, M. (2010). Office discipline and student behavior: Does race matter? *American Journal of Education*, *116*(4), 557–581. doi:10.1086/653629

Rueda, E. (2015). The benefits of being Latino: Differential interpretations of student behavior and the social construction of being well behaved. *Journal of Latinos and Education*, *14*(4), 275–290. doi:10.1080/15348431.2015.1025955

Skiba, R., & Edl, H. (2004). *The Disciplinary Practices Survey: How do Indiana's principals feel about discipline*. Bloomington, IN: Center for Evaluation and Education Policy.

Skiba, R. J., Chung, C., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion expulsion. *American Educational Research Journal*, *51*(4), 640–670. doi:10.3102/0002831214541670

Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, *40*(1), 85–107.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, *34*(4), 317–342. doi:10.1023/A:1021320817372

Steinberg, M. P., & Lacoe, J. (2018). Reforming school discipline: School-level policy implementation and the consequences for suspended students and their peers. *American Journal of Education*, *125*, 49.

Theriot, M. T. (2009). School resource officers and the criminalization of student behavior. *Journal of Criminal Justice*, *37*(3), 280–287. doi:10.1016/j.jcrimjus.2009.04.008

Townsend, B. L. (2000). The disproportionate discipline of African American learners: Reducing school suspensions and expulsions. *Exceptional Children*, *66*(3), 381–391. doi:10.1177/001440290006600308

U.S. Department of Education Office for Civil Rights. (2016). *2013–2014 Civil rights data collection: A first look*. Washington, DC: U.S. Government Printing Office.

Welch, K., & Payne, A. A. (2010). Racial threat and punitive school discipline. *Social Problems*, *57*(1), 25–48.

Welsh, R. O., & Little, S. (2018). The school discipline dilemma: A comprehensive review of disparities and alternative approaches. *Review of Educational Research*, *88*(5), 752–794. doi:10.3102/0034654318791582

Western, B. (2006). *Punishment and inequality in America*. New York, NY: Russell Sage Foundation.

Wright, J. P., Morgan, M. A., Coyne, M. A., Beaver, K. M., & Barnes, J. C. (2014). Prior problem behavior accounts for the racial gap in school suspensions. *Journal of Criminal Justice*, *42*(3), 257–266. doi:10.1016/j.jcrimjus.2014.01.001

Wu, S.-C., Pink, W., Crain, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review*, *14*(4), 245–303. doi:10.1007/BF02171974

Authors

MAITHREYI GOPALAN is an assistant professor at The Pennsylvania State University. Her research examines the causes and consequences of racial and socioeconomic disparities in student outcomes through interdisciplinary perspectives.

ASHLYN AIKO NELSON is an associate professor at the O'Neill School of Public and Environmental Affairs, Indiana University-Bloomington. Her research examines the causes and consequences of inequality in the overlapping housing and public education sectors.