ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-458

Desegregated but still separated? The impact of school integration on student suspensions and special education classification

Mark J. Chin
Harvard University

In this paper I study the impact of court-mandated school desegregation by race on student suspensions and special education classification. Simple descriptive statistics using student enrollment and outcome data collected from the largest school districts across the country in the 1970s and 1980s show that Black-White school integration was increasing for districts under court order, but not for a set of comparison districts. Similarly, Black student suspension rates were increasing at faster rates in integrating districts relative to comparison districts, and their classification rates as having an intellectual disability were decreasing at slower rates. Differences-in-differences and event study models confirm these patterns I observe in the raw data: after integration, school districts experienced statistically and practically significant reductions in racial isolation across schools and growth in racial disparities in discipline and special education classification. The impacts of integration are immediate, sustained, and robust for student suspensions in particular. My results thus provide causal evidence confirming prior descriptive and theoretical work suggesting that the racial composition of schools may influence measures of categorical inequality by race.

VERSION: September 2021

Suggested citation: Chin, Mark J. . (2021). Desegregated but still separated? The impact of school integration on student suspensions and special education classification. (EdWorkingPaper: 21-458). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/5r82-x803

Desegregated but still separated?

The impact of school integration on student suspensions and special education classification

Mark J. Chin[a]

[a]Harvard University, 1350 Massachusetts Avenue, Suite 350, Cambridge, MA 02138

Author Note

Correspondence concerning this article should be addressed to Mark J. Chin, 1709 N Alabama St., Indianapolis, IN 46202. E-mail: mark_chin@g.harvard.edu

Abstract

In this paper I study the impact of court-mandated school desegregation by race on student suspensions and special education classification. Simple descriptive statistics using student enrollment and outcome data collected from the largest school districts across the country in the 1970s and 1980s show that Black-White school integration was increasing for districts under court order, but not for a set of comparison districts. Similarly, Black student suspension rates were increasing at faster rates in integrating districts relative to comparison districts, and their classification rates as having an intellectual disability were decreasing at slower rates. Differences-in-differences and event study models confirm these patterns I observe in the raw data: after integration, school districts experienced statistically and practically significant reductions in racial isolation across schools and growth in racial disparities in discipline and special education classification. The impacts of integration are immediate, sustained, and robust for student suspensions in particular. My results thus provide causal evidence confirming prior descriptive and theoretical work suggesting that the racial composition of schools may influence measures of categorical inequality by race.

Keywords: school segregation, school discipline, special education, differences-in-differences, categorical inequality, *JEL No. I21, JEL No. I24*

Desegregated but still separated?

The impact of school integration on student suspensions and special education classification

1. Introduction

A rich literature shows that the desegregation of K-12 schools by race can improve the educational and adulthood outcomes of Black students without any negative consequence for White youth (e.g., Guryan, 2004; Johnson, 2011). But levels of Black-White isolation across schools have at best stagnated after decades of efforts to integrate (Reardon & Owens, 2014). Positive effects also appear to stem from the more equitable distribution of school resources, which can be directly achieved through other policies (e.g., school finance reforms) that arguably face fewer social, political, and judicial barriers. Furthermore, because Black communities often carry a disproportionate burden during efforts to reduce racial isolation in schools—Black teachers lose jobs, Black schools are closed, and Black students are the ones bused to desegregated schools (Schofield, 1991; Thompson, 2020)—school district leaders should fully understand the tradeoffs made when integrating before doing so.

In this study, I thus investigate how court-ordered school integration impacts two outcomes that have been understudied in this literature but contribute to observed educational inequalities in the present day: student suspensions and special education identification. I focus on a set of the largest districts in the country that were mandated by courts to racially desegregate schools following the U.S. Supreme Court ruling in *Brown v. Board of Education* (1954), and I analyze data on outcomes directly reported by these districts from the 1970s and 1980s. To recover causal effects, I use difference-in-differences (DiD) and explore the changes

in suspension and special education rates between Black and White students in desegregating districts to those in a set of comparison districts that are never under a court order.

My analyses contribute three main findings. First, I show that though suspension rates are rising for all students after *Brown v. Board*, rates increase faster in desegregating districts relative to comparison districts. Notably, changes in suspension rates for Black students are over double those of their White peers. Because I employ a quasi-experimental strategy and identify statistically significant differences, this study provides the first causal evidence that certain populations are negatively affected on a key educational measure—school suspensions—that predicts worse outcomes in adulthood (e.g., Bacher-Hicks, Billings, & Deming, 2020), though prior research shows overall positive effects of integration for Black youth.

Second, I find that relative to students in comparison districts, students in desegregating districts are classified as having an intellectual disability at higher rates, with impacts again being concentrated for Black students in particular. These results, however, are less robust than those for suspension rates, and effects are smaller in absolute terms. This study thus yields some evidence that integration leads to differences in the special education identification for Black and White students and provides support for conclusions from the one prior study that I am aware of that investigates this topic causally. Bergman (2019) specifically shows that a more recent cohort of minority students participating in an inter-district integration program are substantially more likely to be classified as needing special education services.

Third, I successfully replicate results from prior work investigating the impact of court-mandated desegregation on levels of school integration (e.g., Reber, 2005; Welch & Light, 1987) but use a more robust identification strategy. These previous studies similarly use DiD to recover causal effects, but specifically rely on comparing levels of racial isolation between only those

districts that are mandated by courts to integrate at different times. Recent advances in econometric theory on approaches that leverage staggered treatment timing for quasi-experimental variation have argued that DiD estimates may be biased if treatment effects change monotonically over time and/or if effects vary across the observations that experience treatment at different times (Baker, Larcker, & Wang, 2021; Callaway & Sant'Anna, 2020; Goodman-Bacon, 2018). By using a stacked DiD approach (see also Cengiz, Dube, Lindner, & Zipperer, 2019; Deshpande & Li, 2019), I compare integrating districts to only those never under court mandate and avoid these sources of bias. Reassuringly, I continue to find that court orders substantially equalize the racial composition of schools within a district and increase the exposure of White students to Black students.

My results speak specifically to two larger bodies of research. First, as noted earlier, prior work shows that school integration following *Brown v. Board* decreased the dropout rates of Black students (Guryan, 2004) and improved their labor market and health outcomes (Johnson, 2011). My study is the first to causally identify how student suspensions and special education classification changed during this period when schooling across the U.S. was radically altered. Furthermore, because one theoretical benefit of school integration is expanded opportunity for interaction between Black and White youth, knowing desegregation's impact on these outcomes is particular important. Social scientists in support of the Supreme Court's decision in *Brown v. Board* hoped that the ruling would ultimately improve outgroup racial attitudes, which the contact theory of psychology (Allport, 1954) predicts can happen when Black and White individuals more frequently interact with one another. Indeed, one recent study on school resegregation found evidence in support of this hypothesis (Billings, Chyn, & Haggag, 2021). If, however, Black and White students in "integrated" settings continue to be separated—

for example, if students are suspended from school and/or if they are sorted into different classrooms within schools based on programmatic needs resulting from students' special education classification—Black-White contact would be less pervasive than predicted, and the theoretical benefits of intergroup interaction would be less likely to accrue.

My study also builds on extant literature investigating how schools reproduce inequality by creating and sorting students into different categories, as seen by the disproportionate receipt of school discipline and special education by race (Domina, Penner, & Penner, 2017; Shores, Kim, & Still, 2020). Research shows, for example, that Black students are suspended 1.5 times more often than their White peers—even after accounting for key confounding covariates, including racial disparities in income (Shores et al., 2020). These differences by race, which have been observed for decades (Gordon, 2018), lead to other Black-White disparities in schools and contribute to the "school-to-prison pipeline"—or the link between school suspensions and adulthood encounters with the criminal justice system (Bacher-Hicks et al., 2020).

Prior work on special education rates is less conclusive on the implications of disproportionality by race. Like school discipline, Black students have consistently received special education services at higher rates than their White peers for decades (Fish, 2019). However, these overall differences may mask the contribution of key determinants of special education classification, with implications for students' outcomes in schools and in adulthood. On the one hand, classification may provide students with supports necessary to learn in schools. Hanushek, Kain, and Rivkin (2002) provide some evidence for this, finding that special education classification increases students' achievement. On the other hand, students needing special education may also suffer unintended consequences. The stigma of classification, in addition to the isolation from other students and higher quality teachers or curricular materials

often associated with special education receipt (Fish, 2019), may lead to negative outcomes in the long run. Furthermore, significant heterogeneity in the different categories of educational disabilities combined with racial differences in categorization may compound existing inequalities. Being identified as having an emotional disturbance or intellectual disability, for example, may lead to even more acute exclusion and stigma for students (Fish, 2019).

There are several reasons to believe that reducing racial isolation in schools through integration could lead to changes in student suspensions and special education classification. Given data limitations, testing the mechanisms behind any observed effects is beyond the scope of this paper. But it is worth highlighting that empirical research shows larger Black-White disciplinary inequalities in segregated contexts (Shores et al., 2020) and that Black students are overidentified for special education in schools with more White students (Elder, Figlio, Imberman, & Persico, 2021), especially for categories of disability that confer "lower status" (Fish, 2019). With respect to special education receipt, these authors argue that classification may increase in integrated schools because Black students are more distinct from their peers.

School desegregation could relate to school discipline and special education classification through several other channels. First, if predominantly Black schools are underfunded relative to their White counterparts, special education rates might be lower because fewer resources are available to spend on necessary programming. Suspension rates might be higher if Black students and their parents feel disenfranchised by institutional efforts to isolate them (Reardon & Owens, 2014). But theoretically, inequalities may also arise because of integration. Following *Brown v. Board*, many Black teachers lost their jobs when schools integrated (Thompson, 2020), and we know that Black students receive exclusionary discipline at lower rates when they are taught by Black teachers (Lindsay & Hart, 2017). This "race-match" effect may be potentially

due to Black teachers having weaker negative implicit and explicit racial attitudes and prejudices towards Blacks overall (Chin, Quinn, Dhaliwal, & Lovison, 2020). Racial biases might not just influence how educators perceive students' behaviors when choosing to discipline (Ferguson, 2000; Skiba, Michael, Nardo, & Peterson, 2002); they also might affect the evaluation of student work and change expectations such that Black students are overidentified for needing special education programming (Elder et al., 2021; Papageorge, Gershenson, & Kang, 2020; Quinn, 2020).

A more pernicious mechanism than those relying on unconscious processes like implicit racial bias is that educational stakeholders—whether they be teachers, system leaders, or White families—actively push for more exclusionary policies when schools integrate to ensure White students' access to limited resources (Ferri & Connor, 2005; Fish, 2019). This perspective aligns with the racial threat hypothesis (Key, 1949), which posits that as the relative size of the Black population grows, so does the use of different forms of social control over Black individuals. Some evidence highlights how states and districts in the U.S. South with stronger preferences for school segregation took longer to integrate and were more likely to do so only after court mandate (Cascio, Gordon, Lewis, & Reber, 2008). Forced integration in similarly resistant contexts might thus lead to increased suspension and special education classification rates for Black students.

In summary, much research has investigated the impacts of school desegregation. But very few studies provide direct causal evidence on whether reducing racial isolation between schools impacts the key outcomes of student suspension and special education classification, even though prior empirical work and theory put forth myriad reasons for why integration and exclusionary school policies may be linked. In the following text, I first describe the data,

sample, and empirical strategy I use to fill this gap in the literature. I then share results from my investigation of the impact of school desegregation following *Brown v. Board* on Black-White disparities in school suspension and special education classification rates before concluding.

2. Method

2.1 Data

All data used in this study come from the same source: the Office of Civil Rights (OCR) surveys administered to school districts across the U.S. in the 1960s through the 1980s. Like past research that used these data to demonstrate changes in the levels of school desegregation following *Brown v. Board* (e.g., Guryan, 2004; Reber, 2005; Welch & Light, 1987), I leverage the surveys' information on the racial composition of students at the school and district level. I specifically create two commonly used measures of school racial segregation for district-level analyses: the dissimilarity index, which captures the proportion of Black or White students in a district that would need to switch schools in order to achieve balanced racial representation across schools; and the White-Black exposure index, which captures the proportion of Black students in the average White student's school.

I also consider data from the surveys on student outcomes by race. Specifically, the OCR began asking schools and districts to report student suspensions and special education classification in the mid-1970s. The repeated data collection by the OCR allows me to investigate district-level patterns in these two outcomes over time, including the time periods immediately leading up to and immediately following when districts undergo court-mandated school integration. As I describe in my empirical strategy below, I use this data to specifically assess the similarity of integrating districts and non-integrating districts based on trends in suspension and special education classification rates. Prior studies have lacked such detailed data

when investigating the impacts of desegregation on student outcomes and have relied on trends in other measures to establish comparability (i.e., Guryan, 2004).

Though schools use several classifications for special education (Fish, 2019), I focus on the earliest, most consistently reported category for special education in the OCR: whether students receive services for intellectual disability. In survey years when schools report both total and subgroup numbers, I find that Black and White students with an intellectual disability make up 44 and 29 percent, respectively, of the total number of students receiving special education services for each group. Information on suspensions is available throughout the time panel.[1]

2.2 Sample

To identify the effects of school desegregation on student suspensions and special education classification, I focus on a set of school districts that were mandated by courts to integrate following *Brown v. Board*. The racial isolation of schools in these districts were first investigated by Welch and Light in 1987 but have since been considered in numerous studies (e.g., Guryan, 2004; Reber, 2005). These particular districts were sampled for exploration because they were large and because Black and White students each comprised at least 10 percent of each district's student body (Welch & Light, 1987); the 107 districts meeting these criteria represented one-fifth of total K-12 enrollment and almost half of all minority enrollment in the U.S. in 1968.

Because my empirical strategy (described in more detail below) relies on comparing student outcomes before and after school desegregation, I exclude from the 107 integrating

[1] The OCR survey data contain other measures of categorical inequality in schools, but these measures are less consistently collected and/or data collection begins later such that most districts ever under court mandate to integrate do not have baseline data to analyze. For example, the OCR data contain information on student identification as having a specific learning disability, speech-language impairment, and emotional disturbance. This data, however, begins in the 1976-1977 OCR survey. Only 25 of 107 desegregating districts analyzed in prior research that I also consider in my study have even one year of baseline data for these outcomes; I thus focus on the more consistently collected information on students' intellectual disability.

districts 67 that implemented court-mandated integration before the first year that schools began reporting student suspensions and special education classification to the OCR (i.e., the 1974-1975 school year).[2] I also exclude one additional district from analyses—New Castle County in Delaware—because the district integrated in 1978 through the consolidation of several districts, but not all comprising districts were consistently surveyed by the OCR before consolidation. In Table A1, I provide the names and desegregation dates for these 39 districts, which are largely located in urban contexts.

I provide more detail here about the OCR survey data collection from the 1960s through the 1980s because it affects the second group of districts I include in analyses. Data collection occurred in the fall semester for every school year from 1969 to 1975, and then again in 1977, 1979, 1981, 1983, 1987, and 1989. However, not all districts were surveyed every year though, once a district was surveyed, all schools in that district were included as well. Across survey administrations, sampling approaches did not remain the same, though efforts were made for national representativeness. In general, larger districts were consistently sampled with greater probability as were those with higher proportions of minority students and those of "high interest", which most notably included districts that were integrating schools by race.

As mentioned above, I focus on 39 districts under court order to integrate; 36 of these districts have data from all 13 OCR surveys. This coverage is unsurprising given the nature of how districts were sampled. I also identify a set of districts in the OCR data to compare student outcomes. These 84 districts: are never under court order to desegregate (as identified in their OCR reports in 1977 onward); have Black and White enrollment for all observed data points; and have data for all 13 surveys, which allows for consistent juxtaposition with the integrating

[2] For simplicity, I henceforth refer to school years by the calendar year of the associated spring semester.

districts. In this analysis, I favor results from this "full OCR" sample, but also show results when including all districts with Black and White enrollment never under court order to desegregate as a comparison group. Because results are robust across samples and because districts do not self-select into the OCR sample, I am thus not concerned with bias resulting from the missingness of data from districts across surveys.

[Insert Figure 1 about here.]

In Figure 1, I provide basic descriptives for the 39 districts in my sample that are under court mandate to integrate and for the 84 comparison districts with full OCR data. In the figure, I first plot the number of districts affected by mandates over time; as the vertical bars depict, the districts in my sample last to integrate do so by 1979. The figure then shows that levels of school segregation are decreasing for integrating districts over time but not for comparison ones. It also shows that suspension rates for Black students are increasing over time but more rapidly for integrating districts, and that rates for Black students being classified as having an intellectual disability are dropping but more rapidly for comparison districts. These district-level details (in addition to information on enrollment numbers and White students' outcomes) converge with statistics reported in Table A2 and suggest that integration may be leading to increases in exclusionary school outcomes for Black students. In the next section, I describe my empirical strategy, which allows me to test the significance of these changes and to investigate whether trends in outcomes are parallel leading up to integration between treated and comparison districts, despite Figure 1 and Table A2 showing absolute differences.

2.3 Empirical strategy

To identify the causal impact of school integration on student suspensions and special education classification, I use stacked difference-in-differences (DiD) (Cengiz et al., 2019;

Deshpande & Li, 2019) which helps to address concerns of bias in some standard DiD estimators raised by recent papers (e.g., Goodman-Bacon, 2018). For the first difference, I compare outcomes within integrating districts before and after courts mandate desegregation. For the second difference, I compare outcomes over time between integrating districts and those never mandated to desegregate.[3]

To operationalize this DiD, I first create a unique dataset for the 39 "treated" (i.e., mandated by courts to integrate) districts in my analytic sample. For each dataset, I include the treated district's OCR data from surveys between 1974 and 1989, but also include the same information from all 84 comparison districts, as well. I then stack all 39 unique datasets and estimate the following model (Equation 1):

$$Y_{ijt} = \sum_y \beta_y D_i 1\left(t - T_j^* = y\right) + \sum_y \psi_y (D_i 1\left(t - T_j^* = y\right) \times Treated_{ij}) + \delta_i + \gamma_j + \phi_t + \varepsilon_{ijt}$$

Where Y_{ijt} captures either the level of segregation in district i in year t, district-level suspension rates, or special education classification rates. In addition to fixed effects for district δ_i and year ϕ_t, I also include fixed effects for each unique dataset j, γ_j, where $\{j \in \mathbb{Z} | 1 \leq j \leq 39\}$. I cluster standard errors at the most conservative level—the dataset level.

The estimated coefficients of interests are ψ_y, which capture for treated districts the relationship between the outcome and a vector of indicator variables indicating how many OCR surveys have passed y since the latest OCR survey, T_j^*, before the court mandated desegregation

[3] As described in Baker et al. (2021), there are several appropriate methods to address bias in DiD estimators that result from variation in treatment timing. I prefer the stacked DiD for its intuitive, transparent, and easy-to-implement approach to removingpotential sources of bias, i.e., districts ever under court mandate to integrate are explicitly removed and never included as a comparison observation in the data. However, in Figures A1 and A2, I show results that leverage an alternative approach that makes two reasonable changes to the stacked DiD. Specifically, I employ the strategy proposed by Callaway & Sant'Anna (2020) and also use as comparison districts those that never face a court order to integrate *and* those that eventually integrate but have not yet faced their mandates. As seen in these figures, my results are robust to these changes—Black students are suspended at higher rates following integration.

for the treated district in the dataset.[4] Data from surveys after T_j^* (i.e., $y \geq 1$) in the treated district thus capture outcomes post integration.[5] Conversely, data from surveys before T_j^* (i.e., $y < 0$) can be used to conduct a partial test of the identifying assumption that, without court-mandated desegregation, the outcomes measured in the OCR surveys would have followed along parallel trends between integrating and comparison districts—even if the two groups substantially differ otherwise (see Figure 1 and Table A2). The model omits the dichotomous variable capturing data during the latest OCR survey before court mandate to integrate, $y = 0$.

In analyses, I also collapse these dichotomous time variables into a single indicator capturing whether outcome data come from surveys after T_j^* to provide standard DiD results as opposed to event study estimates. Finally, to assess whether or not integration impacts White and Black students differentially, I estimate the model represented in Equation (1) but additionally interact variables with an indicator for whether the outcome is measured for Black students—essentially estimating a triple DiD (or "triple" event study) model.

3. Results

I first present results of the impact of integration on district-level school segregation. In Figure 2, I plot coefficient estimates from estimation of the model represented by Equation (1) with one of the two measures of segregation—the dissimilarity index and White-Black exposure index—as the outcome. As noted above, segregation over time is compared to segregation

[4] No main effect for being a treated district in the dataset is included in the model because there is no variation in this main effect after accounting for district fixed effects.

[5] Because the OCR survey data of interest for my study spans eight school years (1974, 1975, 1977, 1979, 1981, 1983, 1987, and 1989), not all dichotomous time variables will be estimated for each district if integration occurs earlier or later in the panel. For example, if a district integrates between 1974 and 1975, this district will not have data for most pre-periods. The range of possible time periods for y are $\{y \in \mathbb{Z} | -4 \leq y \leq 7\}$. Of the 39 desegregating districts in my sample, 10%, 33%, 64%, 85%, 100%, 100%, 100%, 100%, 87%, 64%, 33%, and 15% have data for the time periods -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7, respectively. Finally, note that, because OCR surveys are not administered every year, differences between time periods across datasets will not necessarily reflect the same number of differences in school years.

measured at time period "zero" since desegregation (the omitted dichotomous variable in the model), the latest OCR survey before districts implement court-mandated integration.

[Insert Figure 2 about here.]

First, in the periods following court-mandated integration (i.e., positive periods), district level racial isolation across schools drops significantly, even after controlling for contemporaneous trends in integration indices in comparison districts. These drops are immediate and sustained in the post period. The average desegregating district experiences a 20 percentage-point drop in the dissimilarity index and between a 10 to 15 percentage-point increase in the White-Black exposure index. In layman's terms, the percentage of White or Black students needing to switch schools to achieve racial balance across schools dropped on average 20 percentage points after a district was mandated by a court to racially integrate; alternatively, the average White student saw a 10 to 15 percentage-point increase in the proportion of Black students in his or her school following court orders. Reassuringly, these estimates replicate those seen in prior studies with slightly different samples and identification strategies (Reber, 2005). To further put the magnitude of these effects in context consider that, in treated districts' earliest OCR survey, the means of the dissimilarity and White-Black exposure indices are .670 and .114 with standard deviations of .163 and .063, respectively (see Table A2).

The second pattern observable in Figure 2 is that, prior to court mandates, trends in racial isolation across schools in integrating school districts are slightly different than trends from non-integrating contexts. Specifically, districts in the former category are experiencing slight decreases in school segregation. However, these trends do not appear substantial enough to explain the sizable discontinuity in school segregation post mandate, which lends credence to internal validity of estimates from my DiD approach. Finally, results for the effect of court-

mandated desegregation on actual levels of school segregation are similar regardless of comparison group: point estimates are nearly identical regardless of whether I use data from just non-integrating districts with all 13 years of OCR surveys ("Sub Sample") or if I consider all non-integrating districts ("Full Sample").

[Insert Figure 3 about here.]

In Figure 3 Panel A, I plot coefficient estimates from separate estimations by race of the preferred version of my model represented by Equation (1), with student suspension rates as the outcome. The preferred model includes only Full OCR districts as a comparison group and weights observations by race-specific district enrollment counts. As seen in the figure, the proportion of Black and White students in integrating districts that are suspended is increasing for both groups after court-mandated desegregation and after accounting for contemporaneous trends by race identified in the comparison districts. Notably, these increases, which are immediate and sustained in the post period, are much higher for Black students, implying that school integration leads to increasing racial disparities in receiving exclusionary discipline. Like the results for district-level segregation, there is some evidence that trends in suspension rates differ prior to following court mandates between integrating and non-integrating districts (especially for White students), but they do not appear to fully account for the significant differences after "treatment".

In Figure 3 Panel B, I plot coefficient estimates from separate estimations by race of my preferred model represented by Equation (1) with student disability rates as the outcome. Unlike the results for suspension rates seen in Figure 3, here there is weaker evidence that integration leads to changes to student special education classification. No effects are observable for White students at all, but the figure shows some significant positive increases to classification rates for

Black students in later periods following court-mandated integration. The absolute magnitudes of these effects are much smaller when compared to those for suspensions. That Black students do not experience this significant change immediately further weakens the evidence of an effect.

In Table 1, I provide the point estimates for the preferred version of my event study and DiD model represented by Equation (1) as well as variations of this model to describe the magnitude of effects and test the sensitivity of my findings. Instead of presenting results from separate model estimations by race (i.e., the results shown in Figures 2 and 3), I present impact estimates of integration on suspension and special education classification rates by race concurrently (i.e., a triple DiD or triple event study model). This allows me to formally test, for example, whether after controlling for the observed increase in White student suspension rates I still observe a significant increase in Black student suspension rates.

[Insert Table 1 about here.]

The results in the Table 1 confirm the visual depictions of integration's impacts seen in the above figures and add additional detail. First, in all models, Black students have higher rates of representation in receipt of exclusionary discipline or special education. Specifically, Black students are three to five percentage points more likely to be suspended or identified as having an intellectual disability than White students in my sample of analyzed districts.

Second, Table 1 provides the specific magnitudes for the impact of court-mandated integration on suspension rates. In treated districts post court mandate, suspension rates are up to 2.7 percentage points higher. However, this increase is an additional .5 to 5.4 percentage-points greater for Black students. Panel B shows the simple DiD estimates and highlights that, in my preferred model (columns 3 and 6), Black students in treated districts are suspended 3.3 percentage points more than White students who themselves are 2.8 percentage points more

likely to be suspended following desegregation. This disparity is statistically significant and for the DiD models is fairly robust (i.e., positive and significant) even if I include all non-integrating districts in the comparison group or if I do not weight observations by enrollment counts.

Finally, there is weaker evidence that Black-White disparities in being classified as having an intellectual disability is growing after efforts to reduce racial isolation between schools following *Brown v. Board*. For the event study estimates in columns 4 through 6 of Table 1, we see across models suggestive evidence of increases in special education classification for Black students in later periods following integration (i.e., periods four onward). These effects are smaller relative to those seen for suspension rates in absolute terms (no effects greater than a one percentage-point increase), though effects for both student intellectual disability education classification and suspension are still sizable relative to baseline rates for Black students (three and 6.7 percent, respectively; see Table A2).

The DiD estimates in Panel B, however, are not only smaller in size, they are also less consistent across model specifications. Specifically, DiD estimates for special education classification for my preferred model (column 6) are positive but insignificant and precisely estimated. When removing enrollment count weights, they are no longer positive (but still insignificant and precisely estimated). Finally, when using weights but considering all non-integrating districts as viable comparison counties, estimates are positive and significant. All together, these results argue for more caution when concluding that school desegregation led to increased Black representation in special education programming relative to White students.

4. Discussion

In my study, I find that court-mandated school desegregation in the decades following the U.S. Supreme Court decision on *Brown v. Board* led to immediate, sustained significant

increases in student suspension rates in integrating districts. These increases were largely concentrated among Black students; in my preferred model, I find that White students were three percentage points more likely to be suspended in the post period, but Black students were 6.3 percentage points more likely. In some models, I also find suggestive evidence of increased racial inequity in student special education classification; Black students appear to be classified at higher rates for intellectual disability after efforts to reduce racial isolation across schools, but these differences are sensitive to modeling decisions and, in absolute terms, smaller.

My findings provide among the first pieces of causal evidence that school integration can expand categorical inequality in schools (Domina et al., 2017). Prior empirical and theoretical work has identified several links between school demographic composition and Black-White differences in student suspensions and special education classification (Elder et al., 2021; Fish, 2019). With other studies demonstrating the significance of these two key educational outcomes—especially for exclusionary discipline—for student success on other measures (Bacher-Hicks et al., 2020; Shores et al., 2020), my results thus fill an important gap in the literature.

Two additional points are worth reflecting on when considering my findings. First, though my empirical approach allows me to make causal claims, because I do not have data to track students in school districts longitudinally, estimates may be biased by mobility of students in and out of integrating school districts. For example, my results for Black students may be driven not by integration itself, but by Black students with higher propensities to be suspended migrating into districts under court decree. However, I am not significantly concerned that the internal validity of estimates is substantially affected by this issue for two main reasons. First, I observe large increases in suspension rates for Black students in the immediate period following

court mandated integration, limiting the time frame for which migration can change the student population within districts. Second, prior work using the exact same data as my study has found that, relative to White students (Reber, 2005), Black students were less likely to move following desegregation (Guryan, 2004; Shen, 2018).

Second, my findings appear on the surface to contradict the results from Guryan (2004) who use a similar sample of districts as my study and find that school integration following *Brown v. Board* decreased dropout rates for Black students. Some might expect that suspension and special education classification rates for Black students should be dropping, not increasing, if expanded school exclusion leads to worse future educational outcomes as many predict. To provide at least some suggestive evidence that my results do not directly counter Guryan's findings, I compare data from the 1980 Census on the dropout rates for teenagers that are 15, 16, or 17 years old with my OCR data. In Figure 4, I plot simple bivariate descriptives, weighted by persons, linking suspensions and special education classification from the 1979 OCR survey against dropout rates for Black youth, including only 1980 Census county groups that include districts in my analytic sample (including the Full OCR comparison districts).

[Insert Figure 4 about here.]

The figure shows that both educational outcomes predict higher dropout rates for Black teenagers, as expected. The relationship between dropouts and suspensions, however, is very weak, and the relationship between dropouts and intellectual disability rates, while stronger, would not necessarily lead to substantial higher dropout rates for Black students following integration given its marginal rise following integration.

My findings have a few implications for research and practice. For one, future studies should investigate the effects of more recent efforts that integrate—or resegregate (see Reardon,

Grewal, Kalogrides, & Greenberg, 2012)—schools on student suspensions and special education classification and whether these policies affect categorical inequality. Though it appears that substantial increases to suspension rates for Black students in particular following *Brown v. Board* did not translate to future educational inequities, that may not be true in the present day, when integration does not necessarily lead to the significant equalization of school resources as was seen from the 1960s through the 1980s (Johnson, 2011). Second, more research needs to unpack the mechanisms behind the link between school racial composition and categorical inequality in schools. How much of my observed findings is explained by the massive drop during school integration in the employment of Black teachers, who demonstrate less anti-Black bias? Alternatively, are educational stakeholders actively pushing for expanded use of exclusionary policies in schools to reduce competition for key resources as theorized? Finally, exclusion in present day schools may look very different than how it looked in the past. With so much attention being paid to racial disproportionality in disciplinary and special education classification rates, are schools currently adopting other policies to separate Black and White youth? Academic tracking (e.g., restricting access to gifted education or Advanced Placement courses) may be less pernicious that expulsion and suspension, but no less effective at separating students.

For practitioners and policymakers, my results urge careful consideration when adopting school integration policies. School integration improved racial equity in the past so there is no surprise in the renewed interest in attempts to reduce racial isolation across schools. However, without community support for these efforts (as is often the case), there may be unintended consequences from school integration that undermine potential positive effects. Without support, other educational policies with arguably more proponents and that have directly been shown to

improve the opportunities for historically underserved students (e.g., expanded early childhood education, school finance reform) may be more fruitful avenues to pursue.

Acknowledgements

I thank Nora Gordon and Sarah Reber for making the Office of Civil Rights survey data available, and to Christopher Cleveland and Matthew Lenard for valuable feedback.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Interest

Declarations of interest: none.

References

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Bacher-Hicks, A., Billings, S. B., & Deming, D. J. (2019). *The school to prison pipeline: Long-run impacts of school suspensions on adult crime* (No. w26257). National Bureau of Economic Research.

Baker, A., Larcker, D. F., & Wang, C. C. (2021). *How much should we trust staggered difference-in-differences estimates? Available at SSRN 3794018*.

Bergman, P. (2018). *The risks and benefits of school integration for participating students: Evidence from a randomized desegregation program* (IZA DP No. 11602). IZA Institute of Labor Economics.

Billings, S. B., Chyn, E., & Haggag, K. (2021). The Long-run effects of school racial diversity on political identity. *American Economic Review: Insights*.

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405-1454.

Cascio, E., Gordon, N., Lewis, E., & Reber, S. (2008). From Brown to busing. *Journal of Urban Economics*, *64*(2), 296-325.

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, *49*(8), 566-578.

Deshpande, M., & Li, Y. (2019). Who is screened out? Application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, *11*(4), 213-48.

Domina, T., Penner, A., & Penner, E. (2017). Categorical inequality: Schools as sorting machines. *Annual Rview of Sociology*, *43*, 311-330.

Elder, T. E., Figlio, D. N., Imberman, S. A., & Persico, C. L. (2021). School segregation and racial gaps in special education identification. *Journal of Labor Economics*, *39*(S1), S151-S197.

Ferguson, A. A. (2000). *Bad boys: Public school and the making of Black masculinity.* University of Michigan Press.

Fish, R. E. (2019). Standing out and sorting in: Exploring the role of racial composition in racial disparities in special education. *American Educational Research Journal*, *56*(6), 2573-2608.

Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing* (No. w25018). National Bureau of Economic Research.

Ferri, B. A., & Connor, D. J. (2005). Tools of exclusion: Race, disability, and (re) segregated education. *Teachers College Record*, *107*(3), 453-474.

Gordon, N. (2018). *Disproportionality in student discipline: Connecting policy to research.* Education Next.

Guryan, J. (2004). Desegregation and black dropout rates. *American Economic Review*, *94*(4), 919-943.

Johnson, R. C. (2011). *Long-run impacts of school desegregation & school quality on adult attainments* (No. w16664). National Bureau of Economic Research.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities?. *Review of Economics and Statistics*, *84*(4), 584-599.

Key, V. O. (1949). *Southern politics in state and nation.* Knopf.

Lindsay, C. A., & Hart, C. M. (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.

Papageorge, N. W., Gershenson, S., & Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, *102*(2), 234-251.

Quinn, D. M. (2020). Experimental evidence on teachers' racial bias in student evaluation: The role of grading scales. *Educational Evaluation and Policy Analysis*, *42*(3), 375-392.

Reardon, S. F., Grewal, E. T., Kalogrides, D., & Greenberg, E. (2012). Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools. *Journal of Policy Analysis and Management*, *31*(4), 876-904.

Reardon, S. F., & Owens, A. (2014). 60 years after Brown: Trends and consequences of school segregation. *Annual Review of Sociology*, *40*.

Reber, S. J. (2005). Court-ordered desegregation successes and failures integrating American schools since Brown versus Board of Education. *Journal of Human Resources*, *40*(3), 559-590.

Schofield, J. W. (1991). School desegregation and intergroup relations: A review of the literature. *Review of Research in Education*, *17*(1), 335-409.

Shen, M. (2018). How I met your mother: The effect of school desegregation on birth outcomes. *Economics of Education Review*, *63*, 31-50.

Shores, K., Kim, H. E., & Still, M. (2020). Categorical inequality in black and white: Linking disproportionality across multiple educational outcomes. *American Educational Research Journal*, *57*(5), 2089-2131.

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.

Thompson, O. (2020). School desegregation and Black teacher employment. *Review of Economics and Statistics*, 1-45.

Welch, F., & Light, A. (1987). *New evidence on school desegregation.* U.S. Department of Education.

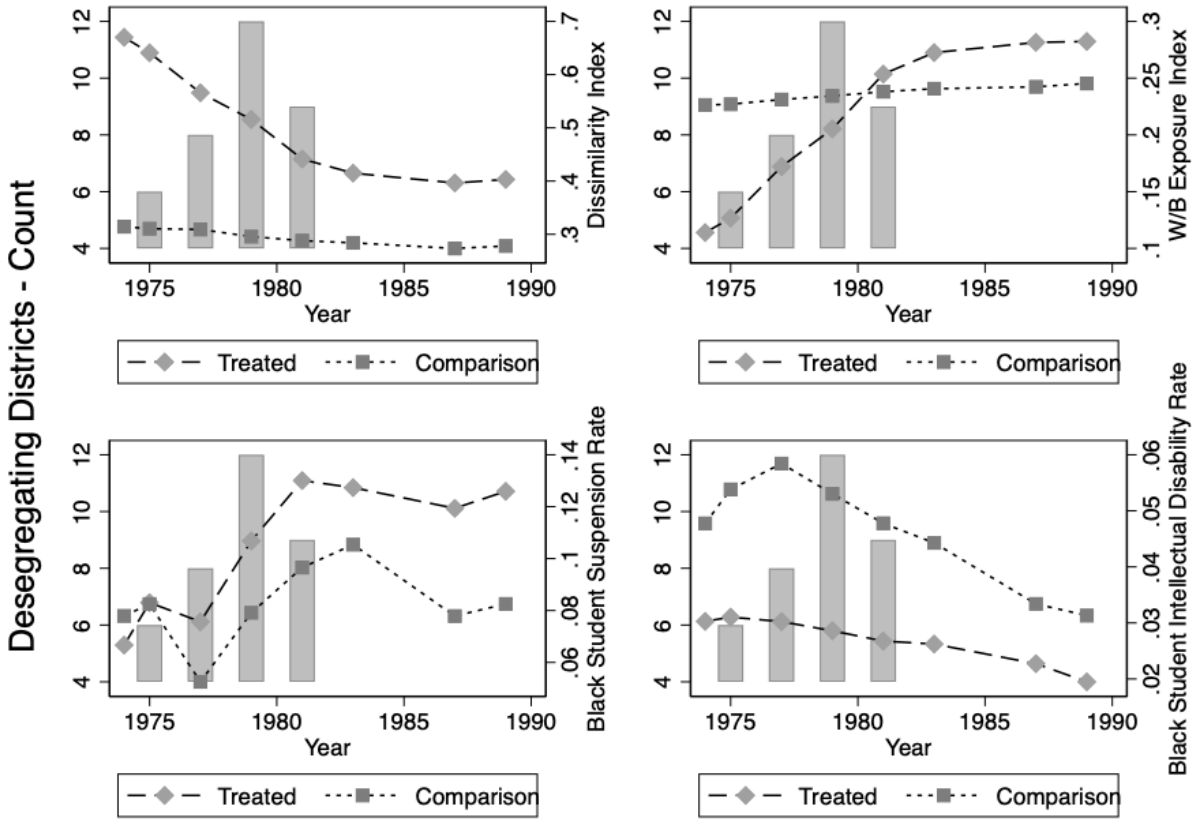

Figure 1. Trends in district-level averages for school segregation and Black students' outcomes, and the number of districts newly affected by court-ordered integration

Notes: The figure plots district-level averages over time for different outcomes from the Office of Civil Rights (OCR) survey data for districts that are ever under mandate by courts to integrate ("Treated") and districts never under mandate ("Comparison"). Included are any Treated districts in my sample (n=39) and Comparison districts with data from all OCR surveys (n=84). The outcome considered is identified by the label for the y-axis on the right side of each panel. The bars in each panel show the number of districts newly affected by court-ordered integration each year.

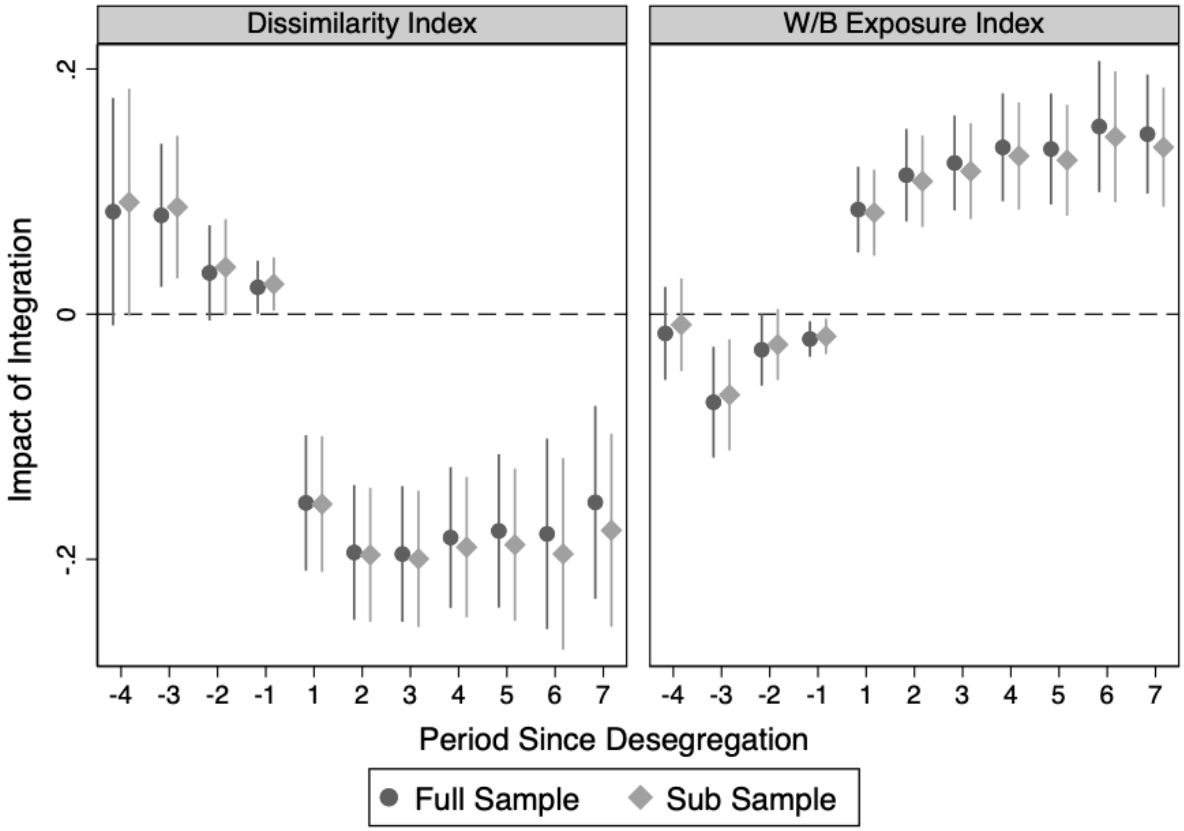

Figure 2. Average impacts of court-mandated integration on district-level segregation indices with 95% confidence intervals

Notes: The figure plots regression coefficient estimates for dummies that capture the period since court-mandated integration (omitted time period = 0) interacted with integration status (i.e., whether the district was ever under mandate to integrate). The dependent variable is a district-level measure of school racial segregation. The regression also controls for district, year, and dataset fixed effects. Standard errors are clustered at the dataset level. The "Full Sample" estimates come from models that include districts that ever integrate and all those that never do. The "Sub Sample" estimates come from models that include districts that ever integrate and those that never do who also have data from all Office of Civil Rights surveys.

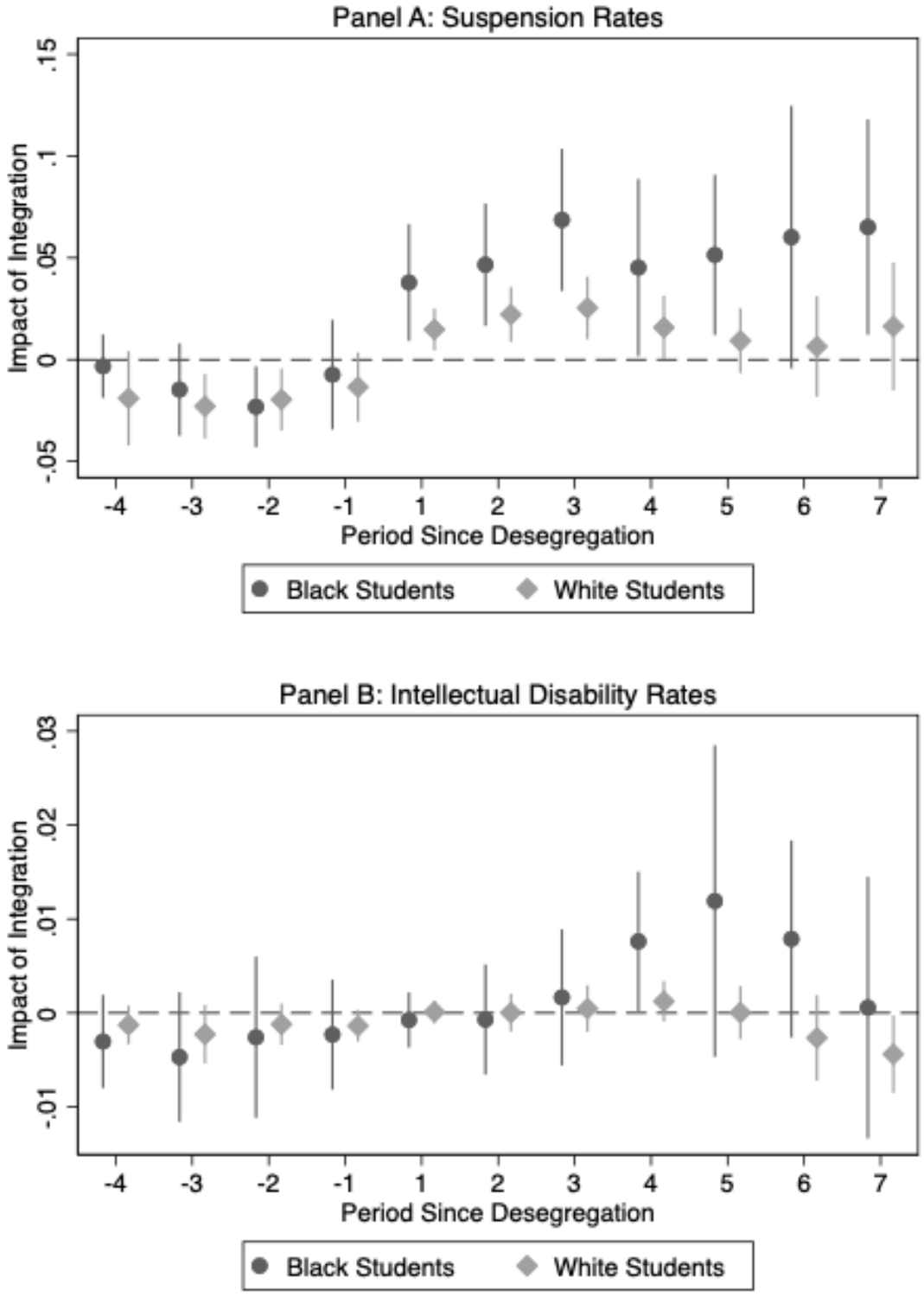

Figure 3. Average impacts of court-mandated integration on district-level suspension and intellectual disability rates with 95% confidence intervals

Notes: The figure plots regression coefficient estimates for dummies that capture the period since court-mandated integration (omitted time period = 0) interacted with integration status (i.e., whether the district was ever under mandate to integrate). Regression models are estimated separately by race. The dependent variable is a district-level measures of a student outcome. The regression also controls for district, year, and dataset fixed effects, and weights for district-race enrollment counts. Standard errors are clustered at the dataset level. Districts included are those that ever integrate and those that never do who also have data from all Office of Civil Rights surveys.

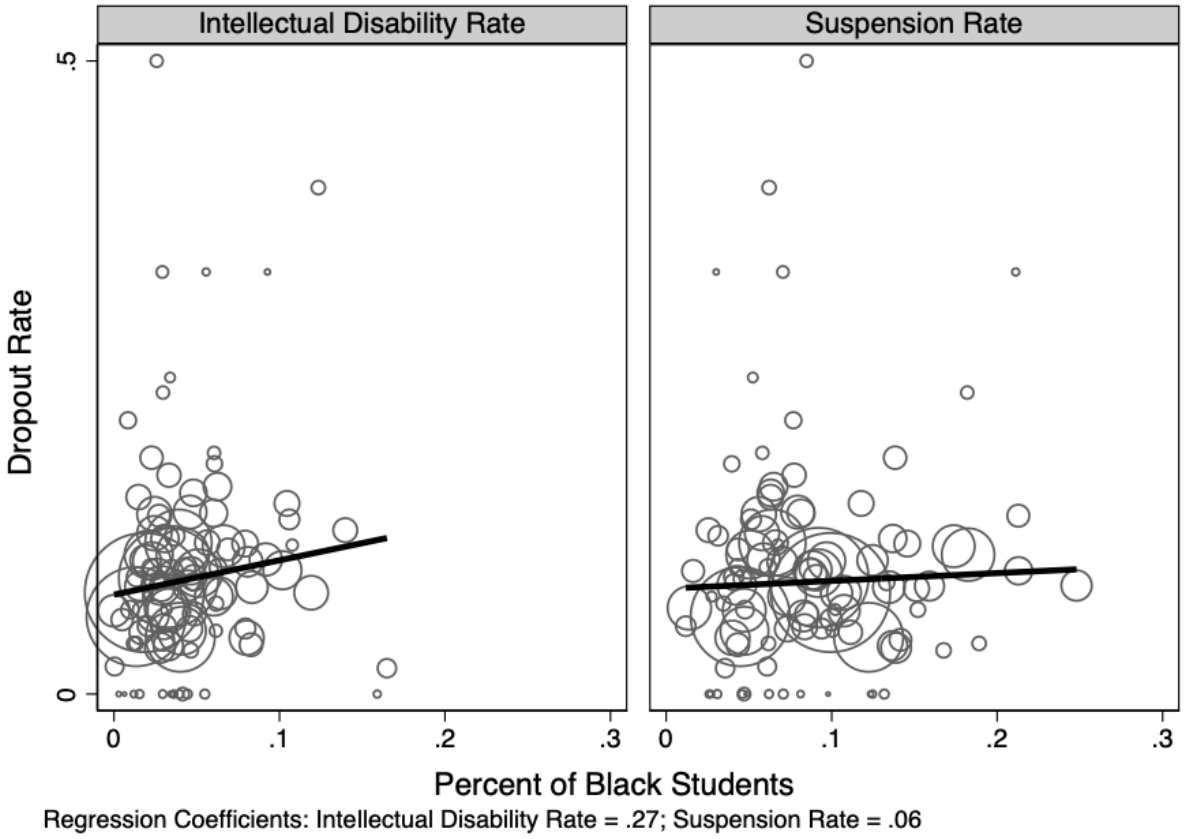

Figure 4. 1980 Census county group dropout rates against suspension and intellectual disability rates for Black youth, weighted by persons

Notes: The figure plots the county group level dropout rates for Black youth aged 15 to 17, measured in the 1980 Census, against suspension and intellectual disability rates for Black youth, measured in 1979 for districts ever under court order to integrate and those that never do who also have data from all Office of Civil Rights surveys. Points are weighted by the Black youth population count in the Census data. Regression coefficients come from county group level regressions with dropout rates as the dependent variable and either the suspension or the intellectual disability rate as the independent variable, weighted by population count.

Table 1. Impacts of court-mandated integration on district-level student outcomes by race

	Suspension Rates			Intellectual Disability Rates		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Event Study Estimates					
Black	0.0388***	0.0365***	0.0462***	0.0396***	0.0352***	0.0355***
	(0.00109)	(0.00165)	(0.000775)	(0.000623)	(0.000423)	(0.000372)
Treated#Period=1	0.00937**	0.0136**	0.0148***	-0.000945**	8.54e-06	9.50e-05
	(0.00412)	(0.00532)	(0.00508)	(0.000458)	(0.000525)	(0.000577)
Treated#Period=2	0.0183***	0.0196***	0.0211***	-0.00116	-0.000241	-8.84e-06
	(0.00572)	(0.00669)	(0.00667)	(0.000831)	(0.000895)	(0.000973)
Treated#Period=3	0.0247***	0.0226***	0.0247***	-0.00168*	8.17e-06	0.000406
	(0.00711)	(0.00763)	(0.00754)	(0.000982)	(0.00110)	(0.00118)
Treated#Period=4	0.0105	0.0118	0.0145*	0.000250	0.00113	0.00165
	(0.00756)	(0.00763)	(0.00801)	(0.00113)	(0.000990)	(0.00104)
Treated#Period=5	0.0124	0.00624	0.00823	0.00117	8.55e-05	0.000839
	(0.00853)	(0.00851)	(0.00839)	(0.00210)	(0.00152)	(0.00152)
Treated#Period=6	0.0189	0.00395	0.00698	-0.00137	-0.00353	-0.00254
	(0.0141)	(0.0132)	(0.0133)	(0.00215)	(0.00281)	(0.00278)
Treated#Period=7	0.0274*	0.0205	0.0216	-0.00226	-0.00435*	-0.00314
	(0.0160)	(0.0168)	(0.0168)	(0.00248)	(0.00252)	(0.00252)
Treated#Black#Period=1	0.00528	0.0236**	0.0227**	0.000748	-0.00197**	-0.000795
	(0.00712)	(0.00921)	(0.0104)	(0.000916)	(0.000938)	(0.000992)
Treated#Black#Period=2	0.0115	0.0254**	0.0262**	0.00301	-0.00290	-0.000575
	(0.00701)	(0.00945)	(0.0106)	(0.00180)	(0.00203)	(0.00205)
Treated#Black#Period=3	0.00816	0.0399***	0.0443***	0.00663***	-0.00129	0.00150
	(0.00702)	(0.0146)	(0.0147)	(0.00176)	(0.00289)	(0.00277)
Treated#Black#Period=4	0.00762	0.0290*	0.0342**	0.00912***	0.00469	0.00601**
	(0.00833)	(0.0147)	(0.0157)	(0.00212)	(0.00288)	(0.00281)
Treated#Black#Period=5	0.00920	0.0395***	0.0458***	0.00986***	0.00864	0.00979
	(0.00956)	(0.0126)	(0.0126)	(0.00229)	(0.00726)	(0.00699)
Treated#Black#Period=6	0.00498	0.0466**	0.0540**	0.00919***	0.00979**	0.00848**
	(0.0106)	(0.0210)	(0.0211)	(0.00233)	(0.00375)	(0.00393)
Treated#Black#Period=7	0.0208	0.0293*	0.0385**	0.00855***	0.000905	-2.98e-05
	(0.0238)	(0.0163)	(0.0165)	(0.00185)	(0.00356)	(0.00364)
	Panel B. Difference-in-Difference Estimates					
Black	0.0411***	0.0382***	0.0446***	0.0396***	0.0342***	0.0354***

	(0.000295)	(0.000374)	(0.000232)	(0.000167)	(0.000136)	(0.000105)
Treated#Post	0.0220***	0.0270***	0.0279***	9.19e-05	0.00138*	0.00158**
	(0.00521)	(0.00711)	(0.00730)	(0.000907)	(0.000722)	(0.000741)
Treated#Post#Black	0.0134**	0.0325***	0.0331***	0.00447***	-0.000938	0.00243
	(0.00545)	(0.00865)	(0.00885)	(0.00164)	(0.00143)	(0.00147)
Full OCR	X		X	X		X
Weighted		X	X		X	X

Notes: Panels A and B of the Table provides regression coefficient estimates showing the average impacts of integration on district-level student outcomes for all students in periods following integration (e.g., Period=1 reflects the first period following integration; Post reflects all periods following integration) and the additional impact for Black students specifically (e.g., coefficients capturing interactions with the Black indicator). All regression specifications include district, year, and dataset fixed effects. Standard errors clustered at the dataset level in parentheses. Full OCR (Office of Civil Rights) models includes districts ever under court mandate to integrate and those never mandated but who have data from all OCR surveys. Weighted models weight by district-race enrollment counts. *$p<.1$, **$p<.05$, ***$p<.01$.

Appendix

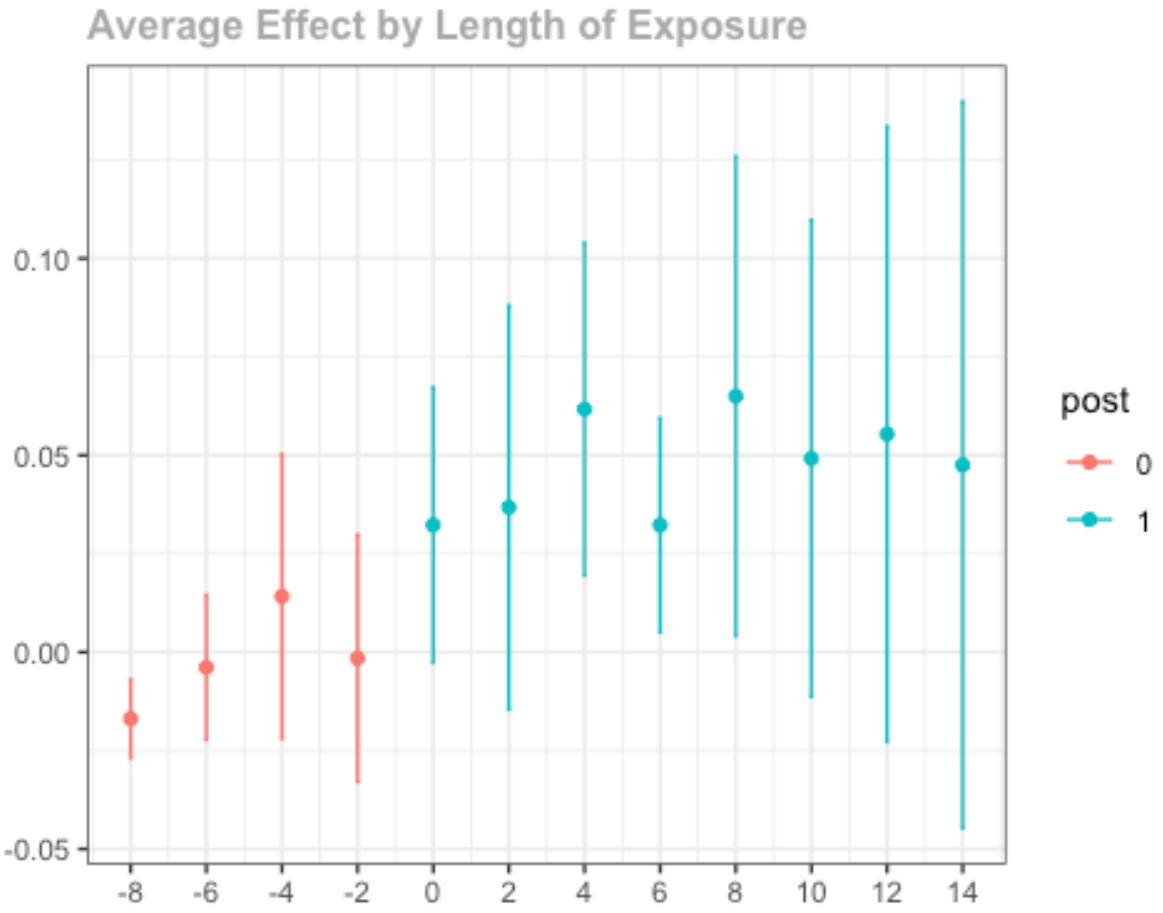

Figure A1. Average impacts of court-mandated integration on district-level suspension rates for Black students with 95% confidence bands

Notes: The figure plots coefficient estimates for years since court-mandated integration dummies interacted with treatment status. "Treated" districts are those in any given time period who have been mandated by courts to integrate. "Comparison" districts are those in any given time period who have not yet (if ever) been mandated by courts to integrate. The dependent variable is the district-level suspension rate for Black students. Standard errors are clustered at the district level. Districts included are both those that ever integrate and those that never do who also have data from all Office of Civil Rights surveys. Models are estimated using the *did* package in *R*.

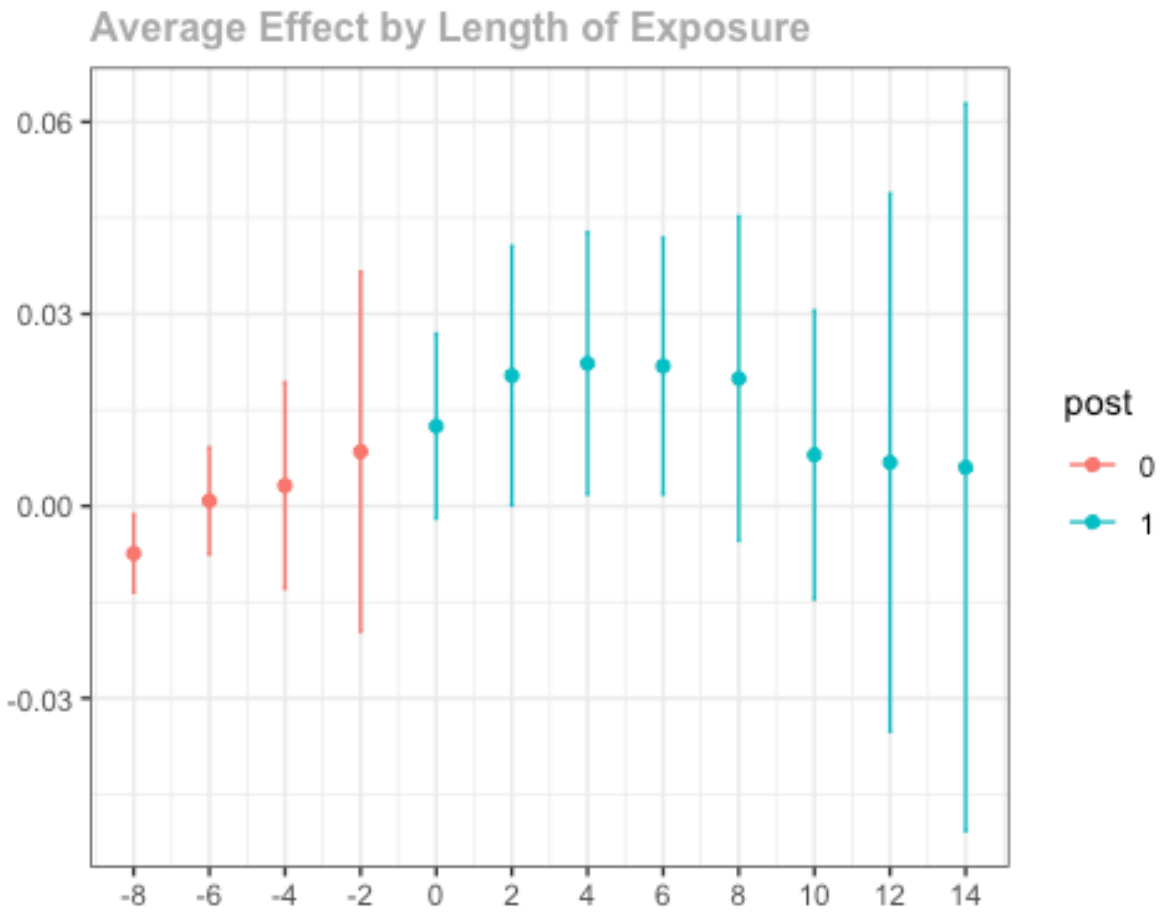

Figure A12 Average impacts of court-mandated integration on district-level suspension rates for White students with 95% confidence bands

Notes: The figure plots coefficient estimates for years since court-mandated integration dummies interacted with treatment status. "Treated" districts are those in any given time period who have been mandated by courts to integrate. "Comparison" districts are those in any given time period who have not yet (if ever) been mandated by courts to integrate. The dependent variable is the district-level suspension rate for White students. Standard errors are clustered at the district level. Districts included are both those that ever integrate and those that never do, but who also have data from all Office of Civil Rights surveys. Models are estimated using the *did* package in *R*.

Table A1. List of school districts in the sample under court mandate to desegregate

State	School district	Year of desegregation	First post-year of OCR data
ARIZONA	TUCSON	1979	1979
CALIFORNIA	FRESNO	1978	1979
CALIFORNIA	LONG BEACH	1980	1981
CALIFORNIA	LOS ANGELES	1978	1979
CALIFORNIA	SACRAMENTO	1976	1977
CALIFORNIA	SAN BERNARDINO	1978	1979
CALIFORNIA	SAN DIEGO	1977	1979
CALIFORNIA	SAN JOSE	1981	1983
COLORADO	DENVER	1974	1975
GEORGIA	DOUGHERTY COUNTY	1980	1981
ILLINOIS	CHICAGO	1982	1983
INDIANA	SOUTH BEND	1981	1983
KANSAS	KANSAS CITY	1977	1979
KENTUCKY	JEFFERSON COUNTY	1975	1977
MARYLAND	BALTIMORE	1974	1975
MASSACHUSETTS	BOSTON	1974	1975
MASSACHUSETTS	NEW BEDFORD	1976	1977
MASSACHUSETTS	SPRINGFIELD	1974	1975
MICHIGAN	DETROIT	1975	1977
MINNESOTA	MINNEAPOLIS	1974	1975
MISSOURI	KANSAS CITY	1977	1979
MISSOURI	ST. LOUIS	1980	1981
NEBRASKA	OMAHA	1976	1977
NEW JERSEY	JERSEY CITY	1976	1977
NEW YORK	BUFFALO	1980	1981
OHIO	AKRON	1977	1979
OHIO	CLEVELAND	1979	1981
OHIO	COLUMBUS	1979	1981
OHIO	DAYTON	1976	1977
OHIO	TOLEDO	1980	1981
OREGON	PORTLAND	1974	1975
PENNSYLVANIA	PHILADELPHIA	1978	1979
PENNSYLVANIA	PITTSBURGH	1980	1981
TEXAS	AUSTIN	1980	1981
TEXAS	ECTOR COUNTY	1982	1983

TEXAS	EL PASO	1978	1979
TEXAS	LUBBOCK	1978	1979
WASHINGTON	SEATTLE	1978	1979
WISCONSIN	MILWAUKEE	1976	1977

Notes: The table lists school districts mandated by courts to integrate schools used in the Welch and Light (1987) study that are also included in this study (*n*=39). The year of desegregation reports the year that the district implemented a major desegregation plan. The first post-year of OCR (Office of Civil Rights) data indicates the year of the first OCR survey where district-level data is considered collected "post" integration.

Table A2. District-level summary statistics

	Treated - 1974 (1)	Treated - 1989 (2)	Comparison - 1974 (3)	Comparison - 1989 (4)
N Black Students	39437.718	31515.333	2734.143	3158.357
	(64746.357)	(48525.071)	(4132.897)	(3845.765)
N White Students	51220.256	23646.949	11967.452	11656.750
	(46944.611)	(17330.328)	(17127.519)	(14525.017)
N Total Students	103313.821	79732.795	16043.464	17230.595
	(125308.737)	(109761.579)	(19679.824)	(18453.211)
Dissimilarity Index	0.670	0.403	0.314	0.278
	(0.163)	(0.175)	(0.196)	(0.178)
W/B Exposure Index	0.114	0.282	0.226	0.245
	(0.063)	(0.177)	(0.206)	(0.223)
Black Student Suspension Rate	0.067	0.126	0.078	0.083
	(0.052)	(0.084)	(0.073)	(0.055)
White Student Suspension Rate	0.033	0.066	0.032	0.044
	(0.027)	(0.042)	(0.029)	(0.031)
Black Student Intellectual Disability Rate	0.030	0.019	0.048	0.031
	(0.014)	(0.014)	(0.032)	(0.025)
White Student Intellectual Disability Rate	0.013	0.012	0.011	0.013
	(0.007)	(0.009)	(0.007)	(0.012)
Districts	39	39	84	84

Notes: The table provides average district-level summary statistics for observations in my study. “Treated” districts are those ever under court mandate to integrate. “Comparison” districts are those never under mandate that also have data from all Office of Civil Rights surveys. 1974 is the earliest year of data available, and 1989 is the latest year of data available.