



# Are students with disabilities suspended more frequently than otherwise similar students without disabilities?<sup>☆</sup>

Paul L. Morgan<sup>a, \*</sup>, George Farkas<sup>b</sup>, Marianne M. Hillemeier<sup>a</sup>, Yangyang Wang<sup>a</sup>, Zoe Mandel<sup>a</sup>, Christopher DeJarnett<sup>a</sup>, Steve Maczuga<sup>a</sup>

<sup>a</sup> The Pennsylvania State University, United States of America

<sup>b</sup> University of California, Irvine, United States of America

## ARTICLE INFO

Action Editor: Amy Briesch

### Keywords:

Suspension  
School-to-prison pipeline  
Special education  
Longitudinal  
Racial/ethnic minorities  
Disparities

## ABSTRACT

Students with disabilities (SWD) have been reported to be disproportionately suspended from U.S. schools and so more likely to experience the “school-to-prison pipeline” through suspension’s associations with lower academic achievement, dropout, juvenile delinquency, and adult criminality. Yet few studies have estimated SWD’s risk of more frequent suspension while simultaneously controlling for potential confounds. Negative binomial regression modeling of suspension count data from a nationally representative and longitudinal sample ( $N = 6740$ ) indicated that males, those from lower resourced families, and students attending more economically segregated schools were more frequently suspended. On average, students who are Black received about 1.6 times as many suspensions by the end of 8th grade as otherwise similar White students. In contrast, having a disability by 1st grade was not a risk factor for more frequent suspension by the end of 8th grade while simultaneously accounting for other risk factors (e.g., gender, race/ethnicity, family SES, prior history of externalizing problem behaviors, being from a English-speaking household, school-level economic composition). Students with specific disability conditions (e.g., emotional disturbance, speech or language impairment) were not at increased risk for more frequent suspension. Students with disabilities who are Black, Hispanic, or of other race/ethnicity were not more frequently suspended than SWD who are White.

## 1. Introduction

U.S. schools are systems for academically, behaviorally, and socially educating students. Schools are therefore expected to be safe, orderly, and civil environments where criminal, violent, or threatening behaviors rarely, if ever, occur. Schools may use suspensions to deter students from engaging in such behaviors. Specifically, suspension is designed to decrease the likelihood that violent or seriously disruptive students continue to engage in potentially dangerous behaviors as well as to protect the safety and ensure the education of other students (Lamont et al., 2013). Seriously disruptive students who remain in classrooms often adversely affect the learning and behavior of their peers (Carrell & Hoekstra, 2010; Figlio, 2007; Fletcher, 2010; Gottfried,

<sup>☆</sup> Funding support was provided by Spencer Foundation Midcareer Grant to the first author as well as an infrastructure grant (P2CHD041025), Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health. No official endorsement should be inferred.

\* Corresponding author at: Department of Education Policy Studies, 310E Rackley Building, The Pennsylvania State University, University Park, PA 16802, United States of America.

Email address: [paulmorgan@psu.edu](mailto:paulmorgan@psu.edu) (P.L. Morgan)

Egalite, & Kirksey, 2016; Horoi & Ost, 2015; Neidell & Waldfogel, 2010). About 7 million students, or about 14% of the school-aged population, were suspended from U.S. schools in 2011–2012 (U.S. Department of Education, 2016a, 2016b).

Despite its use by U.S. schools, suspension is increasingly questioned as a disciplinary practice because of its possible adverse effects on students who are suspended. For example, Morris and Perry's (2016) analyses of a longitudinal sample of students in Kentucky indicated that suspended students tended to later display lower reading and mathematics achievement. Noltemeyer, Ward, and McLoughlin's (2015) meta-analysis of 34 studies indicated that suspended students were more likely to drop out of school. Use of suspension may result in large fiscal costs for U.S. states through lost future tax revenue (Rumberger & Losen, 2016). Both cross-sectional and longitudinal studies indicate that students who are suspended are more likely to engage in substance abuse and violent behaviors as well as to be referred to courts during adolescence (Hemphill et al., 2009; Hemphill, Heerde, Herrenkohl, Toumbourou, & Catalano, 2012; Nicholson-Crotty, Birchmeier, & Valentine, 2009). Students who are suspended are also more likely to later engage in violent criminality and to be arrested in adulthood (Katsiyannis, Thompson, Barrett, & Kingree, 2013; Mowen & Brent, 2016). By increasing the risk of academic difficulties, school dropout, substance abuse, juvenile delinquency, and adult criminality and arrest, suspension is viewed as an entry point into the metaphorical "school-to-prison pipeline" in which school-based disciplinary practices and policies increase the risk that students interact with the juvenile and adult incarceration systems (Hemphill et al., 2009; Mowen & Brent, 2016). For example, being suspended from school is positively associated with committing crimes, including felonies, that result in arrest and referral to the juvenile justice system (Cuellar & Markowitz, 2015; Mittleman, 2018).

### 1.1. Disparities in the use of suspension by U.S. schools

In addition to suspension's associations with lower academic achievement, dropout, juvenile delinquency, and adult criminality, there is also suggestive evidence that U.S. schools may discriminate in their use of suspension. For example, students who are racial or ethnic minorities have been reported to be more likely to be suspended than students who are White (Skiba et al., 2014; U.S. Government Accountability Office [GAO], 2018). These disparities can occur even among otherwise similarly behaving students (Horner, Fireman, & Wang, 2010; Petras, Masyn, Buckley, Jalongo, & Kellam, 2011), as well as those attending similarly resourced schools (GAO, 2018). Racial disparities in suspension are thought to at least partially explain achievement gaps (Gregory, Skiba, & Noguera, 2010; Morris & Perry, 2016). Males, students being raised in low-income households, as well as those who are older for their grade level have also been found to be more likely to be suspended, including among otherwise similarly behaving students (Petras et al., 2011; Theriot, Craun, & Dupper, 2010), thereby suggesting that U.S. schools may be discriminating against some socio-demographic groups in their disciplinary practices (National Research Council [NRC], 2004). However, and alternatively, it may be that school contexts partially explain these disparities, particularly for racial or ethnic minorities and those from low-income families who more frequently attend schools where "zero tolerance" disciplinary practices are in use (Kinsler, 2011). Thus, what is attributed to discrimination due to differential treatment of similarly situated students may instead result from minority students being more likely to attend schools whose disciplinary practices involve greater use of suspension.

Students with disabilities (SWD) have also been reported to be disproportionately suspended (GAO, 2018; Sullivan, Klingbeil, & Van Norman, 2013), leading to suggestions that schools may be discriminating in regards to disability status (Kim, Losen, & Hewitt, 2010; Losen & Gillespie, 2012). For example, suspension rates for students with and without disabilities are estimated to be about 12–15% and 6–7%, respectively (Achilles, McLaughlin, & Croninger, 2007; Losen & Gillespie, 2012; U.S. Department of Education Office for Civil Rights [OCR], 2014). Analyses of cross-sectional data from multiple years suggest that the risk of suspension of SWD began increasing around 2000 (Krezmien, Leone, & Achilles, 2006; Zhang, Katsiyannis, & Herbst, 2004), possibly as SWD began spending more of their school days in general education classrooms due to mainstreaming (U.S. Department of Education, 2016a, 2016b). General education teachers have been shown to be more likely to self-report attitudes of concern, indifference, or rejection towards SWD than towards students without disabilities (Cook, Cameron, & Tankersley, 2007; Cook, Tankersley, Cook, & Landrum, 2000). Among SWD, those identified as having emotional disturbances (ED), specific learning disabilities (SLD), or speech or language impairments (S/LI) have been reported to be more likely to be suspended (Achilles et al., 2007; Krezmien et al., 2006; Sullivan et al., 2013). Racial and ethnic minorities with disabilities have been reported to be especially likely to be suspended (Achilles et al., 2007; GAO, 2018; Krezmien et al., 2006; Losen & Gillespie, 2012), and so to experience the school-to-prison pipeline (Behnken et al., 2014). This may occur because of a lack of training in effectively managing the classroom behaviors of SWD, especially those who are more likely to engage in behaviors that teachers view as disruptive or difficult to manage. This may also occur because of a lack of cultural or language sensitivity in working with SWD who are racial or ethnic minorities (Mendez & Knoff, 2003; Zhang et al., 2004).

To limit the possibility that SWD who are racial or ethnic minorities are being suspended due to discriminatory discipline practices, federal legislation requires U.S. states to monitor for significant disproportionality in the use of suspension by schools districts. Recently enacted federal regulations expand this monitoring by requiring school districts to use a standard methodology when monitoring for significant disproportionality in the use of suspension and other disciplinary practices (U.S. Department of Education Equity in IDEA Rule, 2016b). These new federal regulations expand the manifestation determinations and other legal protections already provided through the Individuals with Disabilities Education Act (IDEA) that ensure that schools continue to provide special education services to SWD whose disruptive behaviors are related to their underlying conditions (U.S. Department of Education Equity in IDEA Rule, 2016b). If significant disproportionality in the use of suspension is

found using the standard methodology for SWD who are among racial or ethnic minorities, school districts in the U.S. are required to review and revise policies, practices, and procedures that contribute to the disparities as well as reallocate their federal funding (U.S. Department of Education Equity in IDEA Rule, 2016b). Federal implementation of these regulations was recently delayed pending further review, including regarding whether the disparities are due to discriminatory practices or instead to other explanatory factors (U.S. Department of Education, 2018a, 2018b).

### 1.2. Limitations in the knowledge base regarding whether and which SWD are being disproportionately suspended

Current empirical evidence regarding the extent to which SWD, including those who are racial or ethnic minorities, are being disproportionately suspended as they attend U.S. schools is currently limited in several ways (Mendez & Knoff, 2003). First, most prior studies investigating whether SWD are disproportionately suspended have used limited statistical controls (e.g., year, race) (Krezmien et al., 2006; Vincent, Sprague, & Tobin, 2012; Zhang et al., 2004) or cross-sectional designs (Anderson & Ritter, 2017; GAO, 2018; Goran & Gage, 2011). Thus, the resulting risk estimates have been unadjusted for strong potential confounds (e.g., gender, age, family socioeconomic status [SES], prior behavioral or academic functioning). The risk attributed to having a disability may therefore be partly—or fully—explained by other factors (Duran, Zhou, Frew, Kwok, & Benz, 2011; Theriot et al., 2010).

Second, very few studies have attempted to identify risk factors for more frequent suspension. Most of the available studies instead have estimated the risk of being suspended dichotomously (i.e., “yes” or “no”) (e.g., Achilles et al., 2007; Bowman-Perrott et al., 2013; Krezmien et al., 2006; Duran et al., 2011; Hemphill et al., 2009, 2012; Martin, 2014). Yet dichotomizing a variable that could be measured continuously results in lost data, larger standard errors, reduced statistical power, and possibly biased risk estimates (Altman & Royston, 2006; MacCallum, Zhang, Preacher, & Rucker, 2002). Identifying risk factors for more frequent suspension is also substantively important. This is because students who are more frequently suspended are especially likely to experience (a) academic difficulties in school due to increased absenteeism and (b) involvement with the criminal justice system due to less adult supervision and more opportunities to interact with criminally active individuals (Anderson & Ritter, 2017). Mowen and Brent (2016) recently reported a strong relation between the frequency of suspension and children's risk of later arrest. Those who were suspended across two, three, or four of the study's survey waves were 136%, 252%, and 400% more likely, respectively, to later be arrested compared to those suspended only during one survey wave.

Third, and although a few available studies have used more extensive statistical controls, these studies have analyzed single district- (Sullivan et al., 2013) or state-level (Anderson & Ritter, 2017; Kinsler, 2011) samples not designed to generalize to the U.S. school-aged population. The few studies that have estimated the frequency of suspension among SWD find that those with the specific conditions of SLD, ED, or other health impairment (OHI), as well as students who are Black, male, or from lower SES families are suspended more frequently. However, generalizability of the samples and statistical control for potential confounds have been limited. For example, Sullivan et al.'s (2013) analyses indicating that SWD were more likely to experience multiple suspensions were based on data from a single, unrepresentative school district in Wisconsin where 50% of the parents had at least a college degree. Further, the cross-sectional data available to the investigators meant that prior behavioral functioning could not be controlled for as a potential confound. Limited generalizability and lack of control for potential confounds also characterize the study by Sullivan, Van Norman, and Klingbeil (2014) that analyzed data from the same district in Wisconsin. Both of these studies examined a constrained range of 0, 1, or > 1 suspensions. Neither Mowen and Brent's (2016), nor Petras et al.'s (2011), nor Mendez and Knoff's (2003) analytical samples included SWD. Anderson and Ritter (2017) examined school-level but not individual-level risk factors for more frequent suspension.

Because prior studies reporting on the suspension risk of SWDs have been unable to adjust for prior behavioral functioning, academic achievement, and other potential confounds (Sullivan et al., 2014), contrasts between otherwise similar students with and without disabilities typically have not been possible. Yet contrasts between similarly situated students who differ in their disability status would better evaluate whether schools are discriminating against SWD when using suspension (e.g., NRC, 2004; OCR, 2016). The very few studies reporting estimates designed to generalize to the U.S. school-aged population, which might be expected to best inform federal legislation and policy, have largely examined whether suspension risk varies by specific disability condition but only among samples of SWD (Achilles et al., 2007; Bowman-Perrott et al., 2013; Duran et al., 2011). Thus, it is currently unknown whether, among students who are otherwise similar in their behavior, academic achievement, and additional background characteristics, schools suspend those with disabilities more frequently. It is also unknown how suspension frequency varies by specific disability conditions as well as how disability status and race or ethnicity interact.

Establishing whether and to what extent SWD, including those who are racial or ethnic minorities, attending U.S. elementary and middle schools are being more frequently suspended than otherwise similar students without disabilities has important implications for federal legislation and policy as well as for educational research and practice. Doing so would better clarify whether, as hypothesized (Losen & Gillespie, 2012) and thought to be occurring widely (U.S. Department of Education Equity in IDEA Rule, 2016b), U.S. schools are discriminating against SWD in their use of suspension. This would be particularly important to examine for those SWD who are racial or ethnic minorities. If SWD who are racial or ethnic minorities are being more frequently suspended than SWD who are White, this would suggest the need for greater use of culturally and language sensitive behavioral interventions by school psychologists, special education teachers, and other practitioners to better meet the specific academic or behavioral needs of students with special needs. More broadly, such an investigation would help establish whether, as has been reported to occur for the specific condition of attention-deficit/hyperactivity disorder (ADHD), being identified as

disabled increases the likelihood of entering the metaphorical school-to-prison pipeline for students who are racial or ethnic minorities through an attendant greater risk for suspension (Behnken et al., 2014). Finding this to be the case would suggest that SWD who are racial or ethnic minorities may benefit from additional academic and behavioral supports to avoid suspension's associated life-course adversities. Investigating which students with specific disability conditions are more frequently suspended would also better inform federally mandated monitoring efforts.

### 1.3. Purpose of study

Given the aforementioned limitations of prior empirical work, the purpose of the current study was to evaluate the extent to which SWD are being more frequently suspended as they attend U.S. elementary and middle schools. To address substantive and methodological limitations in prior studies, we estimated the risk of being more frequently suspended using negative binomial regression modeling of the number of times suspended by eighth grade from a nationally representative and longitudinal sample. We statistically adjusted the risk factor estimates simultaneously for an extensive set of individual student-, family-, and school-level covariates. Doing so should better approximate contrasts between similarly situated students and so yield rigorously derived estimates of the risk for more frequent suspensions attributable to specific factors including disability status. We examined the following research questions:

1. Are students identified as having disabilities by the end of first grade at greater risk of being more frequently suspended by the end of eighth grade than students not so identified? Based on prior work (e.g., Sullivan et al., 2013), we hypothesized that SWD would initially (i.e., in unadjusted estimates) receive more suspensions than students without disabilities.
2. To what extent is any initially observed risk of more frequent suspension for SWD explained by other student- (e.g., race/ethnicity, gender), family- (SES), or school-level factors (racial, ethnic, and economic segregation), as well as plausibly exogenous individual student-level behavior and achievement measured at school entry? Based on prior work investigating for potential discriminatory disability identification practices by schools (Hibel, Farkas, & Morgan, 2010; Morgan, Farkas, Hillemeier, & Maczuga, 2017), we hypothesized that other explanatory factors would fully explain initial but unadjusted associations between having a disability and the frequency of suspension. We expected that these other explanatory factors would include individual student-level race or ethnicity, gender, family SES, prior history of externalizing problem behaviors, difficulties with attention and other learning-related behaviors, academic achievement, and school-level racial, ethnic, and economic segregation. We also hypothesized that these alternative explanatory factors would explain the greater suspension risk previously reported for students with specific disability conditions including SLD, ED, and S/LI (Bowman-Perrott et al., 2013; Sullivan et al., 2014).
3. Are SWD who are racial or ethnic minorities suspended more frequently than SWD who are White? We hypothesized that other explanatory factors (e.g., family SES, prior behavioral functioning) would also account for these interactions (Morgan et al., 2017).

## 2. Method

### 2.1. Data and analytical sample

We analyzed data from the Early Childhood Longitudinal Study-Kindergarten cohort 1998–1999 (ECLS-K), a data set collected and administered by the National Center for Education Statistics (NCES). The ECLS-K is a nationally representative, longitudinal cohort of children who were followed from kindergarten entry through the end of eighth grade. A multistage, probability sampling design was employed in which approximately 1300 public and private schools were sampled from 100 geographic regions. About 24 kindergarten students were recruited from each school. Data collection continued throughout elementary and middle school. Surveys were administered in kindergarten, first, third, fifth, and eighth grade. The NCES provided sampling weights to account for attrition across sample waves. The study's sample contained information on 6740 students who participated in the ECLS-K from kindergarten through eighth grade. As described below, we used multiple imputation to account for missing data for this sample.

### 2.2. Measures

#### 2.2.1. School suspension

The dependent variable of interest was the number of times a student had been suspended from school by the spring of eighth grade. During the parent interview portion of the ECLS-K eighth grade assessment, each parent was asked the number of times his or her child had received an in- or out-of-school suspension. This was reported by parents as a cumulative total regardless of the ages or grades at which each suspension occurred. Because this survey was administered to parents during the spring of the student's eighth grade year, we refer to this variable as the student's number of suspensions through eighth grade.

### 2.2.2. Disability status

School personnel reported whether an Individualized Education Program (IEP) was on file. We considered SWD as those students with IEPs on file who were receiving special education services due to formally identified disabilities. Along with the study's covariates, we used disability status (as indicated by an IEP being reported on file at the school) to predict the number of times students were suspended by the spring of eighth grade. To reduce the possibility of reverse causality (i.e., that suspended students missed instruction and thus performed less well academically, leading to being identified as disabled, such that any relation between these variables was due to suspension resulting in disability identification rather than vice versa), we conservatively measured disability status as having been identified by the spring of first grade. We did so because students identified as disabled received this designation prior to the preponderance of the time period when they were likely to be suspended (i.e., second to eighth grade). Doing so helped ensure that any significant relations between disability status and suspension frequency were most likely due to earlier disability identification affecting later suspension frequency. Prior work has found that most suspensions occur during middle school, followed by high school. Very few suspensions occur during elementary school, suggesting that even fewer occur by first grade (Mendez & Knoff, 2003).

Students who had an IEP by the end of first grade (i.e., 6.8% of the study's sample) represented only a subset of all students who had an IEP by the spring of eighth grade. However, and by restricting attention to students identified as having disabilities by the end of first grade, we limited the possibility that any relations observed between disability status and suspension frequency were the result of more frequently suspended students having been, as a consequence of suspension, identified as having disabilities. Further, students with an IEP by first grade should have had more severe impairments due to their earlier identification and so displayed greater academic or behavioral difficulties than students identified as having disabilities in the later grades. This made for a more conservative test of whether disability status increased the frequency of suspension. This is because, if having a disability increased the risk of suspension frequency, then student with more severe impairments should have been the most likely to be suspended more frequently. Put another way, if suspension frequency was not elevated for SWD with more severe impairments, it seems unlikely that suspension frequency would be elevated for SWD with less severe impairments. We also examined the risks of more frequent suspension associated with specific disability conditions. These specific conditions included SLD, ED, S/LI, intellectual disability (ID), OHI, developmental delay (DD), or other disability condition (ODC). We coded ODC as including the following rare disability conditions: deaf-blindness, deafness, hearing impairment, orthopedic impairment, traumatic brain injury, and visual impairment.

### 2.2.3. Head Start

Participating in Head Start has been reported to increase the risk for suspension by middle school, although this risk may itself result from children who participated in Head Start being more likely to later attend under-resourced schools (Aughinbaugh, 2001). We, therefore, used participation in Head Start as a covariate when examining whether SWD were more likely to be frequently suspended. Parents reported in the fall of kindergarten whether or not their child had been enrolled in federally sponsored Head Start. We used a dummy variable for having been previously enrolled in Head Start (1 = yes, 0 = no).

### 2.2.4. Non-English-speaking household

Students with limited English proficiency have been reported to be at decreased risk of suspension (Anderson & Ritter, 2017). Parents were asked about the primary language spoken in their home during the fall kindergarten survey. We coded those students whose parents reported a primary language other than English as "1" and those students reported as using English as their primary language as "0." We used the kindergarten measure of this variable to help ensure that it was measured prior to the time period when students were at risk for suspension.

### 2.2.5. Reading achievement

Greater academic achievement has been reported to decrease the risk for suspension (Arcia, 2006; Mizel et al., 2016). We accounted for reading achievement as a covariate using the ECLS-K Reading Test administered during the fall of kindergarten. All items were field-tested and psychometric properties of items were evaluated using item response theory (IRT) methods. Items assessed in kindergarten consisted of basic skills such as print familiarity, letter recognition, beginning and ending sounds, and rhyming sounds. The reliability of the IRT-scale Reading Test scores for the kindergarten assessment was 0.93 (Pollack, Atkins-Burnett, Rock, & Weiss, 2005). We used the fall of kindergarten measure of this variable to ensure that it was measured prior to when students were most likely to be suspended.

### 2.2.6. Mathematics achievement

Mathematics achievement in kindergarten was measured by the ECLS-K Mathematics Test. As with the Reading Test, all items were field-tested and their psychometric properties evaluated using IRT methods. The Mathematics Test's items during kindergarten emphasized basic skills such as identifying numbers and shapes and counting objects. The reliability of the IRT-scaled Mathematics Test scores in the kindergarten assessment was 0.92 (Pollack et al., 2005). As with the reading assessment, we used the fall of kindergarten measure of this variable to ensure that it was measured prior to when students were most likely to be suspended.

### 2.2.7. Behavior

Engaging in problem behaviors, particularly externalizing-type behaviors like fighting or being disruptive, has been reported to increase the risk for suspension (Mizel et al., 2016). Students are typically suspended due to recurrent acting-out behaviors (Skiba et al., 2014). General education teachers completed the Social Rating Scale, a modified version of the Social Skills Rating System (Gresham & Elliot, 1990) to rate an individual student's behavior in the fall of kindergarten. The teachers used a frequency scale to rate how often the student displayed a particular social skill or behavior (i.e., 1 = *never*; 4 = *very often*). Items used for the Approaches to Learning subscale measured how well a student self-regulated his or her behavior while completing learning-related tasks (e.g., attentive, task persistent, flexible and organized) in the classroom. Controlling for prior achievement, these learning-related behaviors best predict later achievement (Duncan et al., 2007; Tach & Farkas, 2006). The split-half reliability for the Approaches to Learning subscale was 0.89 (Pollack et al., 2005). The Externalizing Problem Behaviors subscale measures the frequency of acting out behaviors (e.g., argues with the teacher, fights, shows anger, disturbs the classroom). The split-half reliability for the Externalizing Behaviors scale was 0.90 (As with the achievement measures, we used the kindergarten measures of learning-related and externalizing problem behaviors to ensure that they were measured prior to when students were likely to be suspended).

### 2.2.8. Socio-demographic characteristics

Students who are Black or Hispanic have been reported to have a higher risk of suspension (GAO, 2018; Petras et al., 2011), as have boys (Sullivan et al., 2013). Students from lower SES families have also been reported to be more likely to be suspended than students from higher SES families (Hemphill et al., 2014). Children's age, race or ethnicity (coded as White, Black, Hispanic, or other), gender, and family structure (coded as whether the mother was married by the spring of first grade) have also been reported to increase the risk for suspension (Mizel et al., 2016). Data on these variables were collected from parents during the ECLS-K kindergarten interviews. Family SES was assessed using a multivariate parental self-report questionnaire measuring maternal, paternal, and/or guardian education levels and occupations as well as family income at each of the survey waves. Because our dependent variable was the cumulative number of suspensions received by the spring of eighth grade, we used the SES measure as experienced by students over this time period. Accordingly, we followed the common practice of averaging these variables across the survey waves.

### 2.2.9. School characteristics

Attending a racially or economically segregated school has been reported to increase the risk of suspension (Hughes, Warren, Stewart, Tomaskovic-Devey, & Mears, 2017). We therefore included covariates that captured information about the sociodemographic composition of the schools that students attended including the percentages of students who were Black or Hispanic and those eligible to receive free lunch. We averaged these variables across the survey waves so that each represented the combined school-level racial and economic conditions experienced by the study's students.

## 2.3. Analysis

We first descriptively examined the data including calculating the percent with different numbers of times suspended for students with and without IEPs by first grade. We then used negative binomial regression to model the number of times students were reported to have been suspended. Allison (2012) recommends use of negative binomial regression models when examining count data. This is because, unlike Poisson models that often fit count data relatively poorly, negative binomial regression models allow for over-dispersion (e.g., as here, where most students never experienced suspension and so there were a very large number of zero counts), while also often being easier to estimate than zero-inflated Poisson models. Negative binomial regression also allows for greater flexibility when modeling variance (Winkelmann, 2003). Negative binomial regression models estimate the log of the expected count as a function of the predictor variables. The resulting coefficients are interpreted as follows: a 1-unit change in the independent variable (e.g., having a disability) predicts a change in the difference in the logs of the expected counts of the dependent variable (i.e., being suspended) by the estimated coefficient, holding the other predictor variables in the model constant.

We corrected the standard errors in these models for clustering of sample students in the kindergarten schools using the Huber-White sandwich estimator, thereby adjusting for the non-independence of data from students in the same school when the sample was drawn in kindergarten (Cameron & Miller, 2010; Primo, Jacobsmeier, & Milyo, 2007). We used SAS PROC SURVEYREG to obtain these estimates. Multilevel models and Huber-White sandwich estimators provide similar values for standard errors as well as fully adjust for clustering. Difference in use is explained more by disciplinary training (e.g., economics vs. psychology) than by specific methodological advantages. For example, McNeish, Stapleton, and Silverman (2017) reported that "similar to HLM (hierarchical linear modeling), the CR-SE (cluster-robust standard error) estimates fully address the clustered nature of the data" (p. 118). Arceneaux and Nickerson (2009) reported that "clustered SEs, random effects, and hierarchical models all adequately account for the structure of the data" (p. 184). McNeish et al. (2017) further stated that cluster-adjusted standard errors made fewer assumptions than multilevel models. Primo et al. (2007), when contrasting clustered standard errors to multilevel modeling, stated that "calculating clustered standard errors is a more straightforward and practical approach, especially when working with large datasets or many cross-level interactions" (p. 446). Huang (2016) noted that use of cluster-ad-

justed standard errors based on Taylor linearization is “well accepted, easy to implement, and is often considered the gold standard for variance estimation using complex sample data” (p. 179). However, and as a robustness check, we re-estimated the negative binomial regression models using multi-level modeling instead of CR-SE. The study's main results (available from the study's first author) were consistent across both analytical methods.

We standardized the continuous predictor variables to facilitate comparison of effect sizes. We used weights supplied by the NCES to account for sample attrition by the eighth grade wave of the ECLS-K: 1998–99 and multiple imputation to account for missing data. We used Blimp software (Keller & Enders, 2017) when conducting the multiple imputation to account for the multilevel nature of the data (Enders, Mistler, & Keller, 2016; Mistler & Enders, 2017). About 7% of the data were missing. We conservatively imputed 20 data sets and then obtained parameter estimates for each. The resulting estimates were then averaged using the appropriate techniques in SAS programming software.

We addressed the first research question by describing the detailed distribution of number of times suspended, separately for students who did and did not have an IEP in first grade. We then estimated two regression models. The first simultaneous regression had IEP by first grade, the interaction between IEP and race or ethnicity, and a large number of control variables as additional predictors. The second simultaneous regression repeated this calculation, but in place of whether the student had an IEP by first grade we used dummy variables for the student's specific disability condition. Use of simultaneous regression is recommended when, as is the case here, the focus of the analyses is to assess the relative predicted effects of a set of possible explanatory variables on a criterion variable (Keith, 2015). We could not include interactions between specific disability conditions and race or ethnicity in this equation because of small sample sizes for these combinations. Instead, we continued the inclusion of interactions between IEP or not and race or ethnicity in the model. The results from these two models answer the study's second and third research questions.

### 3. Results

#### 3.1. Descriptive statistics

Table 1 shows the weighted descriptive statistics of the variables. In the sample of 6740 students, 16.6% had been suspended at least once by the spring of eighth grade. Of those suspended, 10.1% had been suspended once, 3.2% were suspended twice, 1.7% were suspended three times, 0.6% were suspended four times, and 0.9% were suspended five or more times. In the spring

**Table 1**  
Descriptive statistics of selected variables ( $N = 6740$ ), weighted.

	Percentage or $M$ ( $SD$ )
Ever suspended by 8th grade	16.6%
Suspended 1 time by 8th grade	10.1%
Suspended 2 times by 8th grade	3.2%
Suspended 3 times by 8th grade	1.7%
Suspended 4 times by 8th grade	0.6%
Suspended 5 times by 8th grade	0.9%
IEP, spring 1st grade	6.8%
White	57.3%
Black	17.3%
Hispanic	18.2%
Other race/ethnicity	7.3%
Male	51.7%
Age, fall kindergarten	68.5 (4.4)
IEP, spring 1st grade, learning disabled	0.7%
IEP, spring 1st grade, emotional disturbance	0.2%
IEP, spring 1st grade, speech or language impairment	2.1%
IEP, spring 1st grade, intellectual disability	0.3%
IEP, spring 1st grade, other health impairment	0.2%
IEP, spring 1st grade, developmental delay	0.5%
IEP, spring 1st grade, other/missing	0.5%
Enrolled in Head Start, fall kindergarten	17.4%
From non-English household, fall kindergarten	11.9%
Family SES, average	−0.0 (0.8)
Externalizing problem behaviors, fall kindergarten	1.6 (0.6)
Approaches to learning, fall kindergarten	3.0 (0.7)
Reading Test score, fall kindergarten, IRT-scale score	35.1 (10.1)
Mathematics Test score, fall kindergarten, IRT-scale score	26.4 (9.1)
School's percentage of Black students, average	16.0 (20.3)
School's percentage of Hispanic students, average	13.9 (18.7)
School's percentage of students receiving free lunch, average	33.1 (24.8)

Note. IEP = Individualized Education Program; IRT = item response theory; K = kindergarten; SES = socioeconomic status. Sample size rounded as per NCES requirements.

of first grade, 6.8% of the sample had an IEP. Of the entire sample, 0.7% percent had SLD, 0.2% had ED, 2.1% percent had S/LI, 0.3% had ID, 0.2% had OHI, 0.5% had DD, and 0.5% had ODC.

Table 2 shows the weighted distribution of the number of times students were suspended, separately by disability status in the spring of first grade. Overall, 20.1% and 16.3%, respectively, of students with and without disabilities in first grade had been suspended once or more by eighth grade. This answers the study's first research question and is consistent with prior studies. Specifically, when SWD are compared to students without disabilities and prior to controlling for other explanatory factors to make the two groups otherwise similar, SWD descriptively had higher rates of suspension.

### 3.2. Negative binomial regression model results

Table 3 shows two negative binomial regression models predicting the number of times students were suspended by the spring of eighth grade. Model 1 shows no significant coefficient of IEP on suspension. The coefficients of the interactions between IEP and race or ethnicity are also non-significant. Further, the coefficients of the interactions between IEP and being Black or Hispanic are negative. These results show that simultaneously controlling for an extensive set of alternative explanatory factors, we found no support for the hypotheses that SWD were more frequently suspended than otherwise similar students without disabilities. We also failed to find that SWD who are Black or Hispanic or of other race/ethnicity were more frequently suspended than otherwise similar SWD who are White.

**Table 2**

Times suspended by the end of 8th grade, students with and without an IEP in first grade and total sample, weighted.

Number of times suspended	Students with IEP by spring of 1st grade (N = 390)	Students without an IEP by spring of 1st grade (N = 6350)	Total (N = 6740)
0	79.9%	83.6%	83.4%
1	13.2%	9.9%	10.1%
2	3.1%	3.2%	3.2%
3	2.1%	1.6%	1.7%
4	0.9%	0.6%	0.6%
5	0.8%	1.0%	0.9%

Note. IEP = Individualized Education Program; sample sizes rounded as per NCES requirements.

**Table 3**

Weighted parameter estimates of negative binomial regression models of the number of times suspended by the end of 8th grade (N = 6740).

	Model 1	Model 2
Intercept	-2.44***	-2.45***
IEP in 1st grade	-0.01	–
Black	0.45**	0.45**
Hispanic	-0.06	-0.05
Other race/ethnicity	-0.33	-0.34*
Male	0.96***	0.97***
Age, fall kindergarten	0.11**	0.12**
Learning disabled		0.01
Emotional disturbance		-0.14
Speech or language impairment		0.42
Intellectual disability		0.36
Other health impairment		-1.46
Developmental delay		-1.30
Other disability		-2.31*
IEP in 1st grade × Black	-0.56	-0.53
IEP in 1st grade × Hispanic	-0.22	-0.33
IEP in 1st grade × Other race/ethnicity	0.16	0.10
Enrolled in Head Start, fall kindergarten	0.13	0.12
From non-English household, fall kindergarten	-0.65***	-0.65***
Family SES, average	-0.29***	-0.29***
Externalizing problem behaviors, fall kindergarten	0.32***	0.33***
Approaches to learning, fall kindergarten	-0.06	-0.06
Reading Test score, fall kindergarten	-0.01	-0.01
Mathematics Test score, fall kindergarten	-0.07	-0.07
School's percentage of Black students, average	0.06	0.05
School's percentage of Hispanic students, average	0.03	0.02
School's percentage of students receiving free lunch, average	0.21***	0.22***

Note. IEP = Individualized Education Program; SES = socioeconomic status. Multiple imputation and clustered standard errors used. All models are clustered on fall kindergarten school. Twenty imputed BLIMP (Keller & Enders, 2017) datasets used. Continuous variables standardized.

\* < 0.05.

\*\* < 0.01.

\*\*\* < 0.001.

In contrast, the coefficient for students who are Black is positive and significant. This suggests that students who are Black were suspended more frequently than otherwise similar students who are White. However, this increased risk was unrelated to having a disability as indicated by the non-significant (and, for Black and Hispanic students, directionally inconsistent) interaction terms. The negative binomial regression estimated the log of the number of events as an additive function of the predictors. Thus, to examine the multiplicative effect of a variable on the expected number of events, one should exponentiate its coefficient. Exponentiating the coefficient for Black students in Model 1 of Table 3 gives  $\exp(0.45) = 1.57$ . Thus, and on average, students who are Black received about 1.6 times as many suspensions by the end of eighth grade as otherwise similar students who are White. (We caution, however, that suspension is a relatively rare event. As shown in Table 1, 83.4% of the sample were never suspended and an additional 10.1% were suspended only once.)

As for the other variables in the equation, a number of them are statistically significant. Males had a much higher rate of suspension than females. Children who were older when they entered kindergarten had a higher rate than those who were younger. Students from non-English-speaking households and from higher SES families were suspended less frequently. Kindergarten students who engaged in externalizing problems more frequently were suspended more frequently during elementary and middle school. Schools with a higher percentage of low-income students had a higher suspension rate.

Model 2 replaced the general IEP variable with the specific disability condition for which students were receiving special education services. The IEP variable was deleted from this equation because it is redundant of having one of the specific disabilities. (Accordingly, students within each disability category are being compared to students without disabilities.) In general, the estimates in this model are very similar to those in Model 1. As for the specific disability conditions, only ODC is statistically significant. This had a large and negative coefficient. Thus, students with ODC were suspended less frequently during elementary and middle school than students without disabilities, possibly because this variable includes conditions such as blindness and other physical disabilities that may not result in disruptions of academic work in classrooms.

To summarize, the study's findings indicated that (a) even when disability status and many potential confounds were simultaneously controlled, students who are Black were more frequently suspended than otherwise similar students who are White; (b) the interactions between race or ethnicity and disability status were not statistically significant after controls including for family SES, prior behavior, and school characteristics; (c) these interactions were directionally consistent with SWD who are Black or Hispanic being less likely to be suspended than SWD who are White; and (d) the strongest relation between specific disability condition and frequency of suspension occurred for students with ODC, who were less frequently suspended than students without disabilities.

### 3.3. Robustness check

As a robustness check, we also estimated three additional regression models (i.e., Poisson regression, zero-inflated negative binomial regression, and zero-inflated Poisson regression, results available from the first author). We found that Akaike's Information Criteria (AIC) was lowest for the negative binomial regression of the four types of regressions that we estimated. Results from these other models were consistent with results reported in Table 3's negative binomial regression. This suggested that the negative binomial regression results were robust to alternative regression model specifications.

## 4. Discussion

Establishing whether and to what extent SWD are being suspended more frequently than otherwise similar students without disabilities is an important issue of policy, research, and practice because of suspension's adverse life-course associations with lower academic achievement, school dropout, substance abuse, juvenile delinquency, and adult criminality (e.g., Cuellar & Markowitz, 2015; Hemphill et al., 2012; Katsiyannis et al., 2013; Noltemeyer et al., 2015). For example, more frequent suspension greatly increases the risk for juvenile arrest (Mowen & Brent, 2016), thereby serving as a potential entry point into the metaphorical school-to-prison pipeline. Federal legislation and regulations require U.S. states to monitor for significant disproportionality in suspension, including for SWD who are racial or ethnic minorities (U.S. Department of Education Equity in IDEA Rule, 2016b). However, the field has lacked strong empirical evidence as to whether schools may be inappropriately suspending SWD more frequently than otherwise similar students without disabilities. This is because few studies have estimated the risk of being suspended more frequently, including as examined in analyses of a nationally representative sample that simultaneously accounted for a number of alternative explanatory factors when estimating the risk attributable specifically to having a disability. Establishing that schools are suspending SWD more frequently than otherwise similar students without disabilities would provide suggestive evidence that U.S. schools may be using suspension in ways that are discriminatory (e.g., NRC, 2004; OCR, 2016).

In order to avoid potential reverse causality between SWD and suspension, we coded IEP as 1 only if the student had an IEP by first grade. Because very few suspensions occur prior to middle school (Mendez & Knoff, 2003), these students were unlikely to have been classified as having a disability because they had previously been suspended. By measuring the relation between this definition of IEP and the number of times suspended through the spring of eighth grade, we were able to estimate the risk of having a disability on suspension while limiting any effect of suspension on the risk of having a disability.

Simple descriptive statistics showed that SWD have higher suspension rates than students without disabilities. However, multivariable regression models including controls for potential confounds fully explained this risk. Importantly, we found no signif-

inant interactions between IEP and race or ethnicity. Significant risk factors for more frequent suspension included being Black, male, older, raised in a English-speaking household, raised in a lower SES family, more frequently engaging in externalizing behavior problems, and attending a school in which a higher percent of students in the school are from low-income families. For example, on average, students who are Black receive about 1.6 times as many suspensions than otherwise similar students who are White across elementary and middle school. However, and despite currently being an explicit target of compliance monitoring by federal legislation and regulations (e.g., U.S. Department of Education Equity in IDEA Rule, 2016b), we found no evidence to suggest that SWD who are racial or ethnic minorities are suspended more frequently than SWD who are White. Thus, and although students who are Black are suspended more frequently, this risk is unrelated to whether they also are SWD.

#### 4.1. Limitations

Our study has several limitations. We report risk factor estimates. Our study does not allow for causal inferences. As in prior studies examining for disproportionality in suspension (Wright, Morgan, Coyne, Beaver, & Barnes, 2014), including for SWD (e.g., Achilles et al., 2007; Bowman-Perrott et al., 2013; Duran et al., 2011), our analyses relied on a parent's retrospective report of school suspension. Parents may not always have known how often their children had been suspended. However, and as seems likely, our risk factor estimates are conservative if the resulting measurement error was random. It is also possible that we would have observed other results if additional measures of the dependent variables (e.g., the number of days suspended for each disciplinary infraction, whether the suspension was in- or out-of-school) and data sources (e.g., school records, student reports, or direct observations) had been available.

Relatedly, data collection for the ECLS-K ended at the end of eighth grade. Although suspension occurs most frequently during middle school (Mendez & Knoff, 2003), we were unable to investigate whether the disparities that we observed continued to occur during high school. We adjusted for prior behavior and achievement using measures administered during the fall of kindergarten, which were plausibly exogenous to the occurrence of suspension throughout elementary and middle school. However, we may have observed other results had we accounted differently for prior behavior. Alternatively, and because the ECLS-K's measure of suspension surveyed parents about whether their children had been suspended by eighth grade, averaging the teacher behavioral ratings across the elementary grades would have resulted in a more tenuous assumption of exogeneity. We did average the family's SES and the racial and economic segregation of schools across kindergarten to eighth grade to account for the changing environmental contexts that students experienced as they moved from the fall of kindergarten to the spring of eighth grade. The ECLS-K data were collected until the spring of 2007. Analyses of more recently collected data are needed to replicate these findings.

#### 4.2. Contributions and implications

Our findings have implications for federal legislation and policy as well as educational research and practice. We find suggestive evidence that U.S. elementary and middle schools use suspension in ways that may be discriminatory, at least as indicated by the disparities in suspension frequency not being explained by measured confounds including student-level measures of behavior, family-level SES, and school-level racial and economic composition. For example, and consistent with some studies (e.g., Petras et al., 2011; Skiba et al., 2011), we find that on average students who are Black are more frequently suspended than similarly situated students who are White. Specifically, our analyses indicated that students who are Black receive about 1.6 times as many suspensions than otherwise similar students who are White by the end of middle school. This increased risk for students who are Black for more frequent suspension is not explained by potential confounds including student-level behavior, family SES, and school racial and economic segregation.

Variability in both sample characteristics and statistical analysis may explain differences between our findings and those of prior work. For example, we analyzed a nationally representative sample in which family SES was directly assessed using parental surveys of family income, occupation, and education level. In contrast, Kinsler (2011) analyzed state-level data and controlled for SES using school records of receipt of free or reduced-price lunch. The ECLS-K's survey of parental education, occupation, and income, which are averaged into a single composite variable, should have better controlled for family SES (Harwell & LeBeau, 2010), which is a strong confound of race and ethnicity (Patten & Krogstad, 2015). Because the ECLS-K data included these more proximal measures of family SES, our analyses should have better corrected for this strong confound and so allowed more rigorous estimates of the risk of suspension attributable to race or ethnicity.

Our findings conflict with those reported in Wright et al.'s (2014) study, in which student suspension was dichotomously measured and most control variables including disability status were contemporaneously assessed at the eighth-grade survey while prior behavioral functioning was averaged from kindergarten, first, and third grade. In contrast, our study's individual-level predictor variables were assessed in kindergarten and first grade while the dependent variable was instead the number of times suspended by the end of eighth grade. We identified students as having disabilities if they had an IEP by the spring of first grade in order to limit the possibility of reverse causality (i.e., suspension causing disability identification). We therefore did not analyze the number of times suspended for all students who ever had an IEP. Instead, we analyzed data from first grade students who already had an IEP to examine whether such identification increased their likelihood of being suspended. We believe that this allowed for a conservative analysis as this subgroup of students likely included an above average share of those with more severe impairments given their earlier disability identification. If SWD with more severe impairments were not at elevated risk

for suspension, then it seems unlikely that this would be the case for SWD with less severe impairments. We controlled for student-level behavior in kindergarten, which should have temporally preceded suspension to further limit reverse causality. In addition, our family- and school-level variables were averaged across all survey waves and so better accounted for changing environmental conditions students experienced from first through eighth grade. (Averaging across waves was also the best option to represent these environmental variables because the ECLS-K parent survey did not ask about specific dates for suspensions.) One resulting practical implication of our study is that school psychologists, teachers, and administrators should consider early interventions for kindergarten students already engaging in externalizing problem behaviors. These students are at increased risk of being more frequently suspended throughout the elementary and middle school grades.

Another practical implication of our study is that teachers, school psychologists, and administrators should consider whether students are being suspended in ways that may discriminate based on sex, race, age, or economic background. This is because, as others have found (Petrus et al., 2011; Wright et al., 2014), socio-demographic disparities in suspension frequency are not fully explained by variability in student-level externalizing problem behavior as well as other student-level indicators of school functioning that might reasonably be related to the frequency of suspension. Educational practices that might help address these disparities include increasing access to race- or gender-concordant teachers and/or those experienced in working with older students, as these teachers may be able to advise on how to appropriately manage problem behavior in culturally sensitive ways that do not result in suspension (Lindsay & Hart, 2017). Increasing the availability of evidence-based training of culturally sensitive practices may also be helpful. School-to-home outreach efforts to parents of children from traditionally marginalized populations may also help reduce socio-demographic disparities in suspension frequency by helping stressed or under-resourced parents better manage their children's externalizing problem behaviors as well as increasing the consistency of the behavior management approaches used across home and school (Mason et al., 2016).

We find little evidence that U.S. elementary and middle schools are suspending SWD more frequently than otherwise similar students without disabilities. Our findings do not support policies or reports (Kim et al., 2010; Losen & Gillespie, 2012; OCR, 2014) that having a disability is itself associated with an increased risk of entering the school-to-prison pipeline via school suspension (Behnken et al., 2014; Mowen & Brent, 2016) conditional on this study's other explanatory factors. This includes SWD who are racial or ethnic minorities, who were suspended no more frequently than SWD who are White. That SWD were not suspended more frequently than students without disabilities as they attend U.S. elementary and middle schools conflicts with some prior work (Krezmien et al., 2006), which did not adjust for potential confounds (Losen & Gillespie, 2012) including prior behavior (GAO, 2018; Sullivan et al., 2013). However, our findings are consistent with other studies (e.g., Theriot et al., 2010), including those few that have similarly accounted for the strong confound of prior behavior (Kinsler, 2011; Wright et al., 2014) and have also failed to find that SWD are more likely to be suspended than otherwise similar students without disabilities. We also fail to find empirical evidence to support federal legislation and policies (U.S. Department of Education Equity in IDEA Rule, 2016b) requiring monitoring for significant disproportionality in the extent to which SWD who are racial or ethnic minorities are being suspended as they attend U.S. schools.

## Uncited references

Berman and Wang, 2017  
Greene, 2003  
Kennedy, 2008  
U.S. Department of Education Office of Special Education Programs, 2016

## References

- Achilles, G.M., McLaughlin, M.J., Croninger, R.G., 2007. Socio-cultural correlates of disciplinary exclusion among students with emotional, behavioral, and learning disabilities in the SEELS national dataset. *Journal of Emotional and Behavioral Disorders* 15, 33–45. <https://doi.org/10.1177/10634266070150010401>.
- Allison, P., 2012 August 7. Do we really need zero-inflated models?. Retrieved from <https://statisticalhorizons.com/zero-inflated-models>.
- Altman, D.G., Royston, P., 2006. The cost of dichotomizing continuous variables. *BMJ* 332, 1080. <https://doi.org/10.1136/bmj.332.7549.1080>.
- 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, 1–32. <https://doi.org/10.14507/epaa.25.2787>.
- Arceneaux, K., Nickerson, D.W., 2009. Modeling uncertainty with clustered data: A comparison of methods. *Political Analysis* 17, 177–190. <https://doi.org/10.1093/pan/mpp00>.
- Arcia, E., 2006. Achievement and enrollment status of suspended students. *Education and Urban Society* 38, 359–369. <https://doi.org/10.1177/0013124506286947>.
- Aughinbaugh, A., 2001. Does head start yield long-term benefits?. *Journal of Human Resources* 36, 641–665. <https://doi.org/10.2307/3069637>.
- Behnken, M.P., Abraham, W.T., Cutrona, C.E., Russell, D.W., Simons, R.L., Gibbons, F.X., 2014. Linking early ADHD to adolescent and early adult outcomes among African Americans. *Journal of Criminal Justice* 42, 95–103. <https://doi.org/10.1016/j.jcrimjus.2013.12.005>.
- Berman, E., Wang, X., 2017. Essential statistics for public managers and policy analysts, 4th ed. Sage, Los Angeles, CA.
- Bowman-Perrott, L., Benz, M.R., Hsu, H.-Y., Kwok, O.-M., Eisterhold, L.A., Zhang, D., 2013. Patterns and predictors of disciplinary exclusion over time: An analysis of the SEELS national data set. *Journal of Emotional and Behavioral Disorders* 21, 83–96. <https://doi.org/10.1177/1063426611407501>.
- Cameron, C., Miller, D., 2010. Robust inference with clustered data. In: Ullah, A., Giles, D.E. (Eds.), *Handbook of empirical economics and finance*. CRC Press.
- Carrell, S.E., Hoekstra, M.L., 2010. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics* 2, 211–228. <https://doi.org/10.1257/app.2.1.211>.
- Cook, B.G., Cameron, D.L., Tankersley, M., 2007. Inclusive teachers' attitudinal ratings of their students with disabilities. *The Journal of Special Education* 40, 230–238. <https://doi.org/10.1177/00224669070400040401>.
- Cook, B.G., Tankersley, M., Cook, L., Landrum, T.J., 2000. Teachers' attitudes toward their included students with disabilities. *Exceptional Children* 67, 115–135. <https://doi.org/10.1177/001440290006700108>.

- Cuellar, A.E., Markowitz, S., 2015. School suspension and the school-to-prison pipeline. *International Review of Law and Economics* 43, 98–106. <https://doi.org/10.1016/j.irle.2015.06.001>.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., ... Sexton, H., 2007. School readiness and later achievement. *Developmental Psychology* 43, 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>.
- Duran, J.B., Zhou, Q., Frew, L.A., Kwok, O.-M., Benz, M.R., 2011. Disciplinary exclusion and students with disabilities: The mediating role of social skills. *Journal of Disability Policy Studies* 24, 15–26. <https://doi.org/10.1177/1044207311422908>.
- Enders, C.K., Mistler, S.A., Keller, B.T., 2016. Multilevel multiple imputation: A review and evaluation of joint modeling and chained equations imputation. *Psychological Methods* 21 (2), 222–240. <https://doi.org/10.1037/met0000063>.
- Figlio, D.N., 2007. Boys named Sue: Disruptive children and their peers. *Education Finance and Policy* 2, 376–394. <https://doi.org/10.1162/edfp.2007.2.4.376>.
- Fletcher, J., 2010. Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management* 29, 69–83. <https://doi.org/10.1002/pam.20479>.
- Goran, L.G., Gage, N.A., 2011. A comparative analysis of language, suspension, and academic performance of students with emotional disturbance and students with learning disabilities. *Education and Treatment of Children* 34, 469–488. <https://doi.org/10.1353/etc.2011.0035>.
- Gottfried, M.A., Egalite, A., Kirksey, J.J., 2016. Does the presence of a classmate with emotional/behavioral disabilities link to other students' absences in kindergarten? *Early Childhood Research Quarterly* 36, 506–520. <https://doi.org/10.1016/j.ecresq.2016.02.002>.
- Greene, W.H., 2003. *Econometric analysis*, 5th ed. Prentice-Hall.
- 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, 59–68. <https://doi.org/10.3102/0013189X09357621>.
- Gresham, P.M., Elliot, S.N., 1990. *Social Skills Rating System*. American Guidance Service, Circle Pines, MN.
- Harwell, M., LeBeau, B., 2010. Student eligibility for a free lunch as an SES measure in education research. *Educational Researcher* 39, 120–131. <https://doi.org/10.3102/0013189X10362578>.
- Hemphill, S.A., Heerde, J.A., Herrenkohl, T.I., Toumbourou, J.W., Catalano, R.F., 2012. The impact of school suspension on student tobacco use: A longitudinal study in Victoria, Australia, and Washington State, United States. *Health Education & Behavior* 39, 45–56. <https://doi.org/10.1177/1090198111406724>.
- Hemphill, S.A., Smith, R., Toumbourou, J.W., Herrenkohl, T.I., Catalano, R.F., McMorris, B.J., Romaniuk, H., 2009. Modifiable determinants of youth violence in Australia and the United States: A longitudinal study. *Australian and New Zealand Journal of Criminology* 42, 289–309. <https://doi.org/10.1375/acri.42.3.289>.
- Hibel, J., Farkas, G., Morgan, P.L., 2010. Who is placed into special education? *Sociology of Education* 83, 312–332. <https://doi.org/10.1177/0038040710383518>.
- Horner, S.B., Fireman, G.D., Wang, E.W., 2010. The relation of student behavior, peer status, race, and gender to decisions about school discipline using CHAID decision trees and regression modeling. *Journal of School Psychology* 48, 135–161. <https://doi.org/10.1016/j.jsp.2009.12.001>.
- Horoi, I., Ost, B., 2015. Disruptive peers and the estimation of teacher value added. *Economics of Education Review* 49, 180–192. <https://doi.org/10.1016/j.econedurev.2015.10.002>.
- Huang, F.L., 2016. Alternatives to multilevel modeling for the analysis of clustered data. *The Journal of Experimental Education* 84, 175–196. <https://doi.org/10.1080/00220973.2014.952397>.
- Hughes, C., Warren, P.Y., Stewart, E.A., Tomaskovic-Devey, D., Mears, D.P., 2017. Racial threat, intergroup contact, and school punishment. *Journal of Research in Crime and Delinquency* <https://doi.org/10.1177/0022427816689811>.
- Katsiyannis, A., Thompson, M.P., Barrett, D.E., Kingree, J.B., 2013. School predictors of violent criminality in adulthood: Findings from a nationally representative longitudinal study. *Remedial and Special Education* 34, 205–214. <https://doi.org/10.1177/0741932512448255>.
- Keith, T.Z., 2015. *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*, 2nd ed. Routledge, New York.
- Keller, B.T., Enders, C.K., 2017. *Blimp user's guide: Version 1.0*. Downloaded from <http://www.appliedmissingdata.com/blimpuserguide-4.pdf>, (on 2/15/2018).
- Kennedy, P., 2008. *A guide to econometrics*, 6th ed. Blackwell, Oxford, UK.
- Kim, C.Y., Losen, D.J., Hewitt, D.T., 2010. *The school-to-prison pipeline: Structuring legal reform*. NYU Press, New York.
- Kinsler, J., 2011. Understanding the black-white school discipline gap. *Economics of Education Review* 30, 1370–1383. <https://doi.org/10.1016/j.econedurev.2011.07.004>.
- Krezmien, M.P., Leone, P.E., Achilles, G.M., 2006. Suspension, race, and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders* 14, 217–226. <https://doi.org/10.1177/10634266060140040501>.
- Lamont, J.H., Devore, C.D., Allison, M., Ancona, R., Barnett, S.E., Gunther, R., ... Guinn-Jones, M., 2013. Out-of-school suspension and expulsion. *Pediatrics* 131, e1000–e1007. <https://doi.org/10.1542/peds.2012-3932>.
- Lindsay, C.A., Hart, C.M.D., 2017. Teacher race and school discipline. *Education Next* 17 (1), 72–78. Retrieved from [https://www.educationnext.org/files/ednext\\_xvii\\_1\\_lindsay\\_hart.pdf](https://www.educationnext.org/files/ednext_xvii_1_lindsay_hart.pdf).
- Losen, D.J., Gillespie, J., 2012. Opportunities suspended: The disparate impact of disciplinary exclusion from school. Retrieved from <https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/upcoming-ccrr-research/losen-gillespie-opportunity-suspended-2012.pdf>.
- MacCallum, R.C., Zhang, S., Preacher, K.J., Rucker, D.D., 2002. On the practice of dichotomization of quantitative variables. *Psychological Methods* 7, 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>.
- Martin, A.J., 2014. The role of ADHD in academic adversity: Disentangling ADHD effects from other personal and contextual factors. *School Psychology Quarterly* 29, 395–408. <https://doi.org/10.1037/spq0000069>.
- Mason, W.A., January, S.-A.A., Fleming, C.B., Thompson, R.W., Parra, G.R., Haggerty, K.P., Snyder, J.J., 2016. Parent training to reduce problem behaviors over the transition to high school: Tests of indirect effects through improved emotion regulation skills. *Children and Youth Services Review* 61, 176–183. <https://doi.org/10.1016/j.childyouth.2015.12.022>.
- McNeish, D., Stapleton, L.M., Silverman, R.D., 2017. On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods* 22, 114–140. <https://doi.org/10.1037/met0000078>.
- Mendez, L., Knoff, H., 2003. Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children* 26, 30–51.
- Mistler, S.A., Enders, C.K., 2017. A comparison of joint model and fully conditional specification imputation for multilevel missing data. *Journal of Educational and Behavioral Statistics* 42 (4), 432–466. <https://doi.org/10.3102/1076998617690869>.
- Mittleman, J., 2018. A downward spiral? Childhood suspension and the path to juvenile arrest. *Sociology of Education* 91, 183–204. <https://doi.org/10.1177/0038040718784603>.
- Mizel, M.L., Miles, J.N., Pedersen, E.R., Tucker, J.S., Ewing, B.A., D'Amico, E.J., 2016. To educate or to incarcerate: Factors in disproportionality in school discipline. *Children and Youth Services Review* 70, 102–111. <https://doi.org/10.1016/j.childyouth.2016.09.009>.
- Morgan, P.L., Farkas, G., Hillemeier, M.M., Maczuga, S., 2017. Replicated evidence of racial and ethnic disparities in disability identification in U.S. schools. *Educational Researcher* 46, 305–322. <https://doi.org/10.3102/0013189X17726282>.
- Morris, E.W., Perry, B.L., 2016. The punishment gap: School suspension and racial disparities in achievement. *Social Problems* 63, 68–86. <https://doi.org/10.1093/socpro/spv026>.
- Mowen, T., Brent, J., 2016. School discipline as a turning point: The cumulative effect of suspension on arrest. *Journal of Research in Crime and Delinquency* 53, 1–26. <https://doi.org/10.1177/0022427816643135>.
- National Research Council, 2004. *Measuring racial discrimination. Panel on methods for assessing discrimination*. National Academies Press, Washington, DC.
- Neidell, M., Waldfogel, J., 2010. Cognitive and noncognitive peer effects in early education. *The Review of Economics and Statistics* 92, 562–576.
- Nicholson-Crotty, S., Birchmeier, Z., Valentine, D., 2009. Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly* 90, 1003–1018. <https://doi.org/10.1111/j.1540-6237.2009.00674.x>.

- Noltemeyer, A.L., Ward, R.M., McIoughlin, C., 2015. Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review* 44, 224–240. <https://doi.org/10.17105/spr-14-0008.1>.
- Patten, E., Krogstad, J.M., 2015. Black child poverty rate holds steady, even as other groups see declines. Retrieved from <http://www.pewresearch.org/fact-tank/2015/07/14/black-child-poverty-rate-holds-steady-even-as-other-groups-see-declines/>.
- Petras, H., Masyn, K.E., Buckley, J.A., Ialongo, N.S., Kellam, S., 2011. Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual- and context-level influences. *Journal of Educational Psychology* 103, 223–237. <https://doi.org/10.1037/a0021545>.
- Primo, D.M., Jacobsmeier, M., Milyo, U., 2007. Estimating the impact of state policies and institutions with mixed-level data. *State Politics and Policy Quarterly* 7, 446–459.
- Rumberger, R.W., Losen, D.J., 2016. The high cost of harsh discipline and its disparate impact. Retrieved from <https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/the-high-cost-of-harsh-discipline-and-its-disparate-impact>.
- Skiba, R.J., Chung, C.-G., 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. *American Educational Research Journal* 51, 640–670. <https://doi.org/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, 85–107.
- Sullivan, A.L., Klingbeil, D.A., Van Norman, E.R., 2013. Beyond behavior: Multilevel analysis of the influence of sociodemographics and school characteristics on students' risk of suspension. *School Psychology Review* 42, 99–114. <https://doi.org/10.1177/1063426610377329>.
- Sullivan, A.L., Van Norman, E.R., Klingbeil, D.A., 2014. Exclusionary discipline of students with disabilities: Student and school characteristics predicting suspension. *Remedial and Special Education* 35, 199–210. <https://doi.org/10.1177/0741932513519825>.
- Tach, L.M., Farkas, G., 2006. Learning-related behaviors, cognitive skills, and ability grouping when schooling begins. *Social Science Research* 35, 1048–1079. <https://doi.org/10.1016/j.ssresearch.2005.08.001>.
- Theriot, M.T., Craun, S.W., Dupper, D.R., 2010. Multilevel evaluation of factors predicting school exclusion among middle and high school students. *Children and Youth Services Review* 32, 13–19. <https://doi.org/10.1016/j.childyouth.2009.06.009>.
- U.S. Department of Education, 2016a. School climate and discipline: Know the data. Retrieved from <https://www2.ed.gov/policy/gen/guid/school-discipline/data.html>.
- U.S. Department of Education, 2016b. Equity in IDEA Rule. Assistance to states for the education of children with disabilities; preschool grants for children with disabilities. Retrieved from <https://www2.ed.gov/policy/speced/reg/idea/part-b/idea-part-b-significant-disproportionality-final-regs-unofficial-copy.pdf>.
- U.S. Department of Education, 2018, July 3. Assistance to states for education of children with disabilities: Preschool grants for children with disabilities, final rule; delay of compliance. *Federal Register* 83 (128), 31306–31317. Retrieved from <https://www.regulations.gov/document?D=ED-2017-OSERS-0128-0393>.
- U.S. Department of Education, 2018b. Regulation postponed two years to ensure effective implementation. Retrieved from <https://sites.ed.gov/idea/regulation-postponed-two-years-to-ensure-effective-implementation>.
- U.S. Department of Education Office for Civil Rights, 2014. Civil rights data collection data snapshot: School discipline. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/data.html>.
- U.S. Department of Education Office of Special Education Programs, 2016. 2016 annual report to congress on the individuals with Disabilities Education Act. Retrieved from <https://ed.gov/about/reports/annual/osep/2016/index.html>.
- U.S. General Accountability Office, 2018. K-12 education: Discipline disparities for Black students, boys, and students with disabilities. Retrieved from <https://www.gao.gov/products/GAO-18-258>, (Washington, DC (GAO-18-258)).
- Vincent, C.G., Sprague, J.R., Tobin, T.J., 2012. Exclusionary discipline practices across students' racial/ethnic backgrounds and disability status: Findings from the Pacific Northwest. *Education and Treatment of Children* 35 (4), 585–601. <https://doi.org/10.1353/etc.2012.0025>.
- Winkelmann, R., 2003. *Econometric analysis of count data*, 4th ed. Springer, Berlin, Germany.
- 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, 257–266. <https://doi.org/10.1016/j.jcrimjus.2014.01.001>.
- Zhang, D., Katsiyannis, A., Herbst, M., 2004. Disciplinary exclusions in special education: A 4-year analysis. *Behavioral Disorders* 29 (August), 337–347. Retrieved from <http://biblos.lamar.edu:2254/ContentServer.asp?T=P&P=AN&K=14853600&S=R&D=a9h&EbscoContent=dGJyMmTo50SepA4y9fwOLCmrOqepq5Srqe4SLaWxWXS&ContentCustomer=dGJyMPGqsUi3prFKuePfyeyx44Dt6fIA%5Cnhttp://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=1485360>.