Who is Most at Risk for School Removal?

A Multilevel Discrete-Time Survival Analysis

of Individual and Contextual-Level Influences

Hanno Petras

JBS International, Inc.

Katherine E. Masyn

Harvard University

Jacquelyn A. Buckley and Nicholas S. Ialongo

Johns Hopkins University

Sheppard Kellam

American Institutes for Research

Author Note

Hanno Petras, JBS International, Inc.; Katherine E. Masyn, Harvard Graduate School of Education, Harvard

University; Jacquelyn A. Buckley and Nicholas S. Ialongo, each at Department of Mental Health, Bloomberg School

of Public Health, Johns Hopkins University; Sheppard Kellam, American Institutes for Research.

Correspondence concerning this article should be addressed to Hanno Petras, Ph.D.; JBS International, Inc.; 5515

Security Lane, Suite 800; North Bethesda, MD 20852; hpetras@jbsinternational.com

Abstract

The focus of this study was to prospectively investigate the effect of aggressive behavior and of classroom behavioral context, as measured in fall of first grade on the timing of first school removal across grades 1–7 in a sample of predominately urban minority youth from Baltimore, Maryland. Using a multilevel discrete-time survival framework, we found that demographic characteristics of the students as well as early individual and classroom level of aggression contribute to the onset of school removal. Although early individual aggression was positively associated with the risk of school removal, initially higher levels of classroom aggression corresponded to *lower* risk of school removal.

Keywords: school removal, suspension, aggression, event history, survival, multilevel

Who is Most at Risk for School Removal? An Application of Multilevel Discrete-Time

Survival Analysis to Understand Individual and Contextual-Level Influences

Removing students from the educational environment as a form of punishment has

become one of the most common forms of discipline in American public schools (Skiba &

Knesting, 2002). For example, the Elementary and Secondary School Civil Rights Compliance

Report noted that suspensions have increased steadily for all students, rising from approximately

1.7 million (3.7%) of all students in 1974, to more than 3 million (6.6%) during the 2002–2003

school year (Office for Civil Rights, U.S. Department of Education, 2002). This increase is

surprising since little evidence exists to suggest that the use of school removal practices

improves student behavior or contributes to overall school safety (Skiba & Knesting, 2002;

Stage, 1997). For example, up to 40% of students suspended are repeat offenders (Bowditch,

1993). Importantly, students who are removed from school are at a higher risk for several

negative outcomes, including academic failure, grade retention, negative school attitude,and,

consequently, high school dropout, juvenile delinquency, and incarceration (Dupper & Bosch,

1996; Ekstrom, Goertz, Pollack, & Rock, 1986; Nichols, 1999; Oppenheimer & Ziegler, 1988;

Osher, Woodruff, & Sims, 2002).

Although the trend denotes an overall increase in the use of school suspension, decades

of research indicate that the disciplinary procedure of school removal is not equitably applied

across race or ethnicity, sex, and socioeconomic status (SES) of students (Krezmien, Leone, &

Achilles, 2006; Leone, Mayer, Malmgreen, & Meisel, 2000; Mendez & Knoff, 2003; Skiba,

Michael, Nardo, & Peterson, 2002). Research has shown that those at highest risk for school

removal are young, minority males from low socioeconomic backgrounds (Imich, 1994;

McFadden, Marsh, Price, & Hwang, 1992; Morgan, 1991; Nichols, Ludwin, & Iadicola, 1999; Skiba, Peterson, & Williams, 1997.)

Various explanations have been offered for these well-documented differences in school removal rates across student race, sex, and SES. In concert with the fact that teachers show a tendency to center the problem "in" the student (Guttman, 1982; Johnson, Whitington, & Oswald, 1994), some studies report that African American males are perceived by teachers as more defiant and disruptive than other student groups (Newcomb et al., 2002; Wentzel, 2002). Psychological research, which has predominantly focused on students' characteristics, has linked the overrepresentation of African American students to increased stress (due to higher rates of victimization and exposure to neighborhood violence) as well as lack of parental support (Hinshaw, 1992; Reid & Eddy, 1997). Another line of research has emphasized the culture differences between African American students and the mainstream, White institutions (Giroux, 2001; Sheets, 1996; Townsend, 2000).

In addition to these accounts, it might also be tempting to explain the overrepresentation of poor minority males among the suspended population by assuming differences in frequency and severity of rule-breaking behavior across race or ethnicity, poverty, and sex. However, current research does not clearly indicate whether poor minority males' higher risk for school removal is due to higher levels of disruptive behavior in the classroom or whether their elevated risk can be explained by disparities in the application of school removal policies or other school-related practices (Civil Rights Project, 2000; Osher et al., 2004). In fact, most of the studies on school removal actually provide very little insight into or empirical analysis of the behavioral and contextual dynamics that may be involved in the school removal "process" (Morrison et al., 2001).

This study aimed to address this gap by examining the associations of student characteristics *and* student behavior on the occurrence of school removal while accounting for early classroom context. In particular, we sought to expand upon previous work, which focuses mainly on the demographic characteristics of removed students, by extending our analysis of school removal to also examine the influence of children's initial level of behavioral adaptation to the classroom, as indicated by their individual first grade aggression levels, as well as the overall level of aggression of their first grade classroom cohort.

In addition, we elaborated on the operationalization of school removal—typically modeled as a single dichotomous event spanning multiple school years (e.g., school removal in elementary school, yes or no)—by modeling it not as a single event but as an event *history*. That is, we modeled it as an event process unfolding over time, so that it was possible to incorporate information about not only *whether* a student experienced a school removal but also *when* that removal occurred.

To meet the aims described above, we posed the following specific research questions: (1) What is the association between behavioral adaptation to school (as measured by first grade aggression levels) and the risk of school removal across grades 1–7, controlling for the student demographic characteristics of race, sex, and SES? (2) What is the association between the student characteristics of race, sex, and SES and the risk of school removal, after accounting for first grade aggression levels? (3) What is the association between the first grade classroom behavioral context (as measured by classroom average levels of aggression) and the risk of school removal across grades 1–7? (4) What, if any, is the moderating effect of classroom aggression levels on the association between student aggression levels and the risk of school removal?

In order to investigate these questions, we utilized an advanced longitudinal modeling method—multilevel discrete-time survival analysis—uniquely suited for examining not only the occurrence of school removal but also the *timing* (i.e., grade) of school removal, while accounting for clustering of students within the classroom and explicitly incorporating the estimation of covariate effects at both the student level and classroom level on the event history process.

Background

In this section, we provide an overview of the existing research on school removal related to the student characteristics of race, sex, and SES, followed by a discussion of behavioral adaption in a classroom context and its association with school-related outcomes.

Student Demographic Characteristics

Racial or ethnic disparities in the use of school-based disciplinary procedures such as school removal have been documented consistently for more than 30 years (Children's Defense Fund, 1975; Costenbader, & Markson, 1998; Gregory, 1997; Kids First Coalition, 1999; McFadden et al., 1992; Skiba et al., 2002; Thornton, & Trent, 1988; Townsend, 2000; Verdugo, 2002). Specifically, data from the Department of Education (Office for Civil Rights, U.S. Department of Education, 2002) indicate that African American children, who represent 17% of the public school enrollment nationally, constitute 36% of out-of-school suspensions. In comparison, White students, who comprise 60% of student population, represent only 44% of suspensions and expulsions.

Students' level of poverty is also associated with the likelihood of being suspended. A report by the American Academy of Pediatrics, based on the 2000 U.S. Census data, indicated that children living at or near the poverty line were more likely to be expelled than their peers

from higher socioeconomic backgrounds (American Academy of Pediatrics, 2003). Poverty,

typically measured in educational research by participation in a free or reduced-price school

lunch program, may also interact with race or ethnicity. For example, poverty has been shown to

increase the risk of removal for African American students, but not necessarily for non-minority

students (Nichols, Ludwin, & Iadicola, 1999).

Lastly, research also indicates a sex discrepancy in school removal practices. Only a few

published reports include sex when investigating school removal, but they have found a disparity

in the rates at which of males and females were removed (Imich, 1994; Mendez, & Knoff, 2003;

Morgan, 1991; Nichols, Ludwin, & Iadicola, 1999). Data from the Department of Education

(Office for Civil Rights, U.S. Department of Education, 2002) indicate that males, who represent

51% of the public school enrollment nationally, constitute 70% of out-of-school suspensions. In

comparison, females who comprise 49% of the student population represent only 30% of

suspensions and expulsions.

Student Behavioral Adaptation to School

The start of elementary school marks an important time of transition in children's

development, characterized by a redefinition of social roles by some non-familial authority (e.g.,

teachers, peers) as well as by requirements to assume new social expectations and obligations.

Research has clearly documented that the ability of children to respond adequately to these new

behavioral and academic demands will have a long-lasting impact on their adolescent and adult

developmental course (e.g., Ensminger, Kellam, & Rubin, 1983; Ensminger & Slusarcick, 1992;

Kellam, Brown, Rubin, & Ensminger, 1983; Robins, 1978).

There is ample evidence that the degree to which students can successfully respond to the

behavioral and academic demands of the first grade classroom is highly predictive of various

later school-related outcomes (Alexander, Entwisle, & Horsey, 1997; Alexander, Entwisle, &

Kabbani, 2001; Entwisle & Hayduk, 1988; Entwistle et al., 2005; Hamre & Pianta, 2001). For

example, cognitive skills and math/reading achievement during first through third grade tend to

be maintained into early and late adolescence (Reynolds, 1994; Stevenson & Newman, 1986);

however, the stability tends to be moderate in the early grades. In addition, Alexander and

colleagues showed that early academic problems place children at risk for grade retention and

school dropout and argue that few opportunities to alter students' academic trajectories occur

after third grade (Alexander, Entwisle et al., 2001). In a similar fashion, Olweus (1979), in his

review of 16 separate studies, documented the high stability of aggressive behavior over short

time periods. Huesman, Eron, Lefkowitz, & Walder (1984) report, in their study of 600 subjects

over 22 years, that peer-rated aggression at age 8 correlates with peer-rated aggression at age 19,

with a coefficient of 0.44. These data are comparable to those reported by Patterson (1982)

regarding the stability of antisocial behavior from school age to late adolescence.

Taken together, these results indicate that aggression, cognitive skills, and reading/math

achievement show moderate to strong continuity, beginning in early childhood. Recognition of

that stability and of the developmental relevance of behavioral adaption at entry to a formal

classroom setting motivated the use of a measure of first grade aggression as the student-level

behavioral predictor in our model of school removal.

Given that school removal may be "justified" by teachers and administrators (based on

zero tolerance policies) as a response to behavioral infractions, it would seem that students who

exhibit more disruptive behavior as they start their formal schooling in first grade will have a

greater likelihood of being removed from school during the later elementary and middle school

years, given their early behavioral and academic deficits (Civil Rights Project, 1999; Seymour,

1999). For example, Sprague & Walker (2000) found that, without any intervention, students

who do exhibit chronic patterns of problem behavior are more likely to experience future school

behavior problems. However, given the preponderance of research indicating that school

removal practices are not equitably distributed across different students, further investigations

into the interplay between students' demographic characteristics and their behavioral

characteristics are warranted.

Early Classroom Behavioral Context

When considering student behavioral adaption to school and an outcome such as school

removal, which is directly determined by student actions relative to behavioral norms and rules

that are both implicit and explicitly defined within the school, it is essential to account in some

way for variability in the social context in which these behaviors and consequences occur. It is

possible that behavioral dynamics in the classroom could influence not only individual behavior

but also school removal practices to the same extent that individual behavior influences the risk

of school removal

Research on aggressive behavior suggests that exposure to aggressive classrooms

increases the risk for persistent aggressive behavior problems (Barth, Dunlap, Dane, Lochman, &

Wells, 2004; Kellam, Ling, Merisca, Brown, & Ialongo, 1998) and conceivably also for

disciplinary removals. Three mechanisms have been cited to explain the negative impact of

elevated levels of classroom aggression: social norm, deviancy training, and coercive teacher

control strategies (Thomas, 2006).

From a social norm or person-group similarity perspective (Tversky, 1977), it can be

argued that groups with a high concentration of aggressive members create a social milieu which

normalizes aggressive behavior (Henry et al., 2000; Wright, Giammarino, & Parad, 1986). From

a deviancy training perspective, it can be argued that aggressive children, when paired, tend to

model, provoke, and reinforce antisocial behavior (Dishion, McCord, & Poulin, 1999). Finally,

research also suggests that highly aggressive classrooms decrease the chances for teachers to

forge positive relationships with students and to use effective classroom management strategies

(Brophy, 1996; Hawkins, VonCleve, & Catalano, 1991; Hughes, Cavell, & Jackson, 1999).

These three mechanisms may occur in combination, contributing to the escalation of early

aggressive behavior and consequently to increased likelihood for disciplinary removals.

Although we did not directly examine the functioning of these mechanisms in the current study,

they provided a conceptual framework which informed our hypotheses and interpretations of the

results.

Present Study

The focus of this study was to prospectively investigate the relationship between

individual and contextual information collected when children transitioned to the elementary

classroom (i.e., fall of first grade) and the timing of first school removal across grades 1–7 in a

large sample of predominately urban minority youth from Baltimore, Maryland.

We hypothesized that students with more maladaptive behavior (i.e., higher aggression

levels) in first grade would be at consistently higher risk of school removal across grades 1–7.

We expected that the student characteristics of race, sex, and SES would still be associated with

the risk of school removal across time, even after accounting for first grade aggression levels.

We hypothesized that, for students who begin school in classroom cohorts where the

majority of students are following the rules and not being disruptive, a student who is disruptive

would most likely "stand out" to the teacher and perhaps be more likely than his/her classmates

to be removed for disciplinary purposes. If that same child is placed in a classroom where there

are many other students being disruptive, it is possible that the child's behavior is viewed according to the norm of behavior in that particular context. In other words, a child's behavior may be evaluated by teachers and administrators in an absolute sense, but also relative to the behavior of other fellow students.

Finally, in accordance with the above cited three mechanisms (social norm, deviancy training, coercive teacher control strategies), we anticipated that first grade classroom levels of aggression would be positively related to the likelihood of removal due to environmental influences on individual levels of aggressive behavior (Kellam et al., 1998).

Method

Sample Selection and Study Participants

Data for this study are from a larger study of two school-based, universal preventive interventions targeting early learning and aggression in first and second grade (Dolan et al., 1993; Kellam & Rebok, 1992) in Baltimore City public schools. The use of data from the Baltimore prevention trial provides the unique opportunity to study variation in the time to first school removal for several reasons.

First, unlike previous studies that often examined samples with a primarily Caucasian student population (e.g., Mendez & Knoff, 2003; Mendez, Knoff, & Ferron, 2002), the sample used in this study is comprised of primarily African American urban individuals who participated in a randomized prevention trial in five urban areas of Baltimore, thus not a sample of convenience. Based on this sample, we were able to maximize the validity of conclusions regarding the association between the chosen covariates and the timing to first school removal by minimizing selection bias.

Importantly, although the sample was not selected specifically to be at high risk for negative school outcomes, it is representative of all students entering the first grade classroom in the 1986–87 school year in urban areas comprised of neighborhoods at high risk (due to high rates of poverty and crime) for many negative outcomes. Overall, this study responds to the U.S. Surgeon General's call for more research among ethnic minority populations (U.S. Public Health Service, 2001).

Participants included 1339 male and female students across 19 schools who were first assessed at age 6 as part of the evaluation of the two preventive interventions. Students receiving one of the two intervention conditions were excluded from these analyses in order to estimate the relationship between covariates and grade of first removal in the absence of any intervention. The selected 1339 students were members of the control group within the evaluation design. The 19 schools were drawn from five geographic areas consisting of three to four schools within the eastern half of Baltimore City, defined by census tract data and vital statistics obtained from the Baltimore City Urban Planning Department and the Baltimore City Public School System. Baltimore City consists of three school district areas (A, B, C) and area B was selected since it coincided with the Epidemiological Catchment Area Baltimore sample area. It represents about a third of the City covering the northern City limits boundary down to the south boundary. The five areas varied by ethnicity, type of housing, family structure, income, unemployment, violent crime, suicide, and school dropout rates. However, each area was defined so that the population within its borders was relatively homogenous with respect to each of the above characteristics. Homogeneity was accomplished by comparing rates in multivariate t-tests. Additionally, equivalence was tested regarding school characteristics including achievement test scores, percent minority and rates of free or reduced school lunch.

Students in the control group came from two sequential cohorts of students; therefore, the second cohort covered 1 year less of school information, such as school removal. The analysis for this paper is therefore restricted to the occurrence of school removal between first and seventh grade in order to make the two cohorts comparable. Of the total available control group, 87% (n=1169) had complete observations on the variables of interest and are consequently used for the analyses in this paper. No significant differences between the selected and the non-selected groups were found regarding sex, race, SES and age in fall of first grade, indicating that the selected sample is representative of the complete sample of 1339 students.

The data did not allow for a comparison of selected and non-selected students regarding behavioral measures in fall of first grade. Of the 170 excluded students, 94% had missing values in fall of first grade on the measure used to assess aggressive behavior. Alternatively, it was tested whether non-selected students differed in terms of their likelihood of school removal, a consequence of aggressive/disruptive behavior. Non-selected compared to selected students had lower log odds of removal, but this difference was not statistically significant (β=-0.330, *Std. Error*=0.256, *p*=0.198).

Of the 1169 students included in the study, males comprised a slight majority of participants, with 594 males (50.8%) and 575 females (49.2%). The ethnicity of the sample was as follows: 761 (65.1%) African American, 393 (33.6%) Caucasian, 9 (0.8%) Native American, 4 (0.3%) Hispanic/Latino, and 2 (0.2%) Asian students. Slightly more than half (51.6%) of the participating students were eligible for free or reduced lunch. Overall, 22.8% of the selected youth were removed from school at least once, with a mean age of 11 years at first removal. The students in the analysis sample were distributed across 48 first grade classrooms with an average classroom size of 25.

When taking a closer look at the distribution of school removal between first and seventh grade separately by race and sex (see Table 1), it can be seen African American males contribute more than half of the removal cases (52.1%), followed by African American females (27.7%). In addition, the sex disparities in school removal are apparent, with males contributing more than two-thirds of the removal cases (67.8%). Finally, the majority (60%) of the school removal cases occur when the sample participants are in middle school (i.e., sixth and seventh grade).

Measures

Teacher Observation of Classroom Adaptation-Revised (TOCA-R). Teacher ratings of aggressive/disruptive behavior, attention/concentration problems, and peer rejection were obtained from fall of first grade until spring of seventh grade using the TOCA-R (Werthamer-Larsson, Kellam, & Wheeler, 1991). In this study, we use information from the earliest assessment (i.e., fall of first grade). The TOCA-R is a structured interview with the teacher, which is administered by a trained assessor. Teachers respond to 36 items pertaining to the child's adaptation to classroom task demands over the last three weeks. Adaptation is rated by teachers on a six-point frequency scale (1 = almost never to 6 = almost always). The *aggressive/disruptive behaviors* subscale includes (1) breaks rules, (2) harms others and property, (3) breaks things, (4) takes others property, (5) fights, (6) lies, (7) has trouble accepting authority, (8) yells at others, (9) is stubborn, and (10) teases classmates. The coefficient alphas for the aggressive/disruptive behaviors subscale ranged from .92 to .94 over grades 1 through 7 (ages 6–13). The 1-year test-retest intraclass reliability coefficients ranged from .65 to .79 over grades 2–3, 3–4, and 4–5. Scores on the aggressive/disruptive behavior subscale were significantly related to the incidence of school suspensions within each year from grades 1 through 7, i.e., the higher the score on aggressive/disruptive behavior, the greater the likelihood

of being suspended from school that year. The TOCA-R measure has been used in numerous studies using this sample (e.g., Petras, Chilcoat, Leaf, Ialongo, & Kellam, 2004a; Petras et al., 2004b; Petras et al., 2005; Petras et al., 2008).

Classroom level of aggression. The classroom aggression was constructed by aggregating the sum of individual aggression scores in fall of first grade within classrooms. Importantly, to avoid inflating the dependence between the individual and classroom aggression variable, the overall classroom aggression mean was subtracted from the individual classroom means. This method is called centering around the Level 2 means, or group-mean centering (Raudenbush & Bryk, 2002).

Students' demographic information. Baltimore City Public School System records provided information on each students' sex, race, socioeconomic background, and age at the start of first grade. Along with classroom and individuals levels of aggression, students' demographic variables, particularly sex and race, are viewed as risk factors in this analysis, and not merely as *control* variables. Given the fact that the act of school removal is subjectively applied to students, a given student's perceived identity could have a direct influence on his/her risk of being removed. However, these demographic variables may also function as risk indicators given their correlations with potential unmeasured predictors of school removal, for example, parental monitoring or neighborhood influences.

Considering the racial makeup of the sample, we treated race as a binary variable, comparing African American students to non-African American students. As noted earlier, the majority of non-African American students were Caucasian. School records of each student's free lunch status were used to indicate the socioeconomic background of the student. Subsidized and partially subsidized lunch categories were collapsed into one category of "free/reduced

lunch" and self-paid lunch was the remaining category. In a study of the reliability and validity of measures of socioeconomic status of adolescents, Ensminger et al. (2000) showed that participation in school lunch programs was highly negatively correlated with family income and other measures of SES.

In our analyses we also included the youth's age in first grade because of the relationship between being over-aged and later educational and behavioral problems (Graue & DiPerna, 2000; May & Kundert, 1995). The majority of children were 6 years old or younger (14.5% age 5, 48% age 6). Of the remaining 37.5% of students, 33.3% were 7 years old in first grade, with the remaining 4.2% students being 8 or 9 years old. Given the bimodal as well as upper and lower truncated distribution, age was treated as a binary rather than a continuous variable. Categorization of the students' age was additionally motivated by the fact that the increase in risk associated with age is not hypothesized to progress in a smooth linear fashion. Therefore, age was coded "0" for being 6 or 5 years old (i.e., age-appropriate) in first grade and "1" for being older than 6 years old (i.e., age 7–9).

School removal. The number of school removals and the grade in which the first removal occurred were obtained from school records. For this study, school removal includes short-term periods out of school, typically referred to as suspension, as well as more long-term removals, such as expulsions, that are usually for the remainder of the term or school year.

"Short-term suspensions" are suspensions of less than 10 consecutive school days for regular and special education students. First, a principal or vice principal will meet with the student to explain why he or she is being suspended. Within 3–5 days, another meeting will be held between the principal, the student, and the parents regarding the suspension.

"Long-term suspensions" means the removal of a student from school for disciplinary reasons for a period of more than 10 consecutive days. A meeting will be held between the principal, student, and parents regarding the suspension and the principal will make a recommendation to the person chosen by the superintendent. Within 10 days of the student's first removal from school, a meeting will be held between the person chosen by the superintendent, the student, and the parents. If this meeting is not held within the first 10 days, the suspension must be rescinded; that is, the child may go back to school. Notice of the meeting with the person chosen by the superintendent must be in writing. It must inform the parents and student of the charges and the policy allegedly violated. Parents and student have the right to have witnesses present and to bring an advocate/attorney to the meeting. In these data, the majority of removals fell into the category of suspensions; that is, they were short term in nature. Ninety-four percent were suspended and only 5.2% were expelled.

As noted earlier, school suspension is not a perfectly measured variable, and its observed distribution may vary due to variation in teacher and administrative practices. We therefore compared the information about school removal extracted from school records to teacher reports about suspension. In grades 6 and 7, teachers were asked if a particular student had been suspended. In cross-tabulating these responses with the school records, we found that 91% of the 267 students were also identified as a case by the teacher interview, indicating sufficient levels of reliability.

Data Analysis

This study integrates the strength of survival analysis and hierarchical linear modeling by conducting a multilevel discrete-time survival analysis to model school removal in grades 1–7

(Reardon, Brennan, & Buka, 2002). The complete model is shown in Figure 1, which uses

squares for observed variables and circles for latent variables.

Survival analysis, also known as event history modeling, is an approach that allows the

investigation of not only *if* but also *when* a given event occurred (Singer & Willett, 1993). For

this study, the event of interest was defined as *first* school removal experienced by a given

student. Although the event of school removal may occur to an individual more than once, we

focus here on the first removal because of its implication for later problems and its relevance for

school-based preventive interventions. The time scale of the first school removal event was

recorded in discrete-time intervals—grade of first removal—so, although the time-to-event

process may be more continuous in nature, with removal able to occur on any school day during

the year, the data limitations required that the process be modeled using discrete-time.

Seven binary event indicators (represented by SR_{G1}–SR_{G7} on Level 1 in Figure 1),

corresponding to the seven time periods, were coded such that an event indicator had a value of 1

if first school removal occurred during that grade and 0 if the first school removal had *not*

occurred by the end of that grade. Once first school removal occurred or a participant was lost to

follow-up, the remaining event indicators were coded as missing. Thus, the sequence of seven

event indicators represents the entirety of each participant's event history with respect to first

school removal across grades 1–7.

Further, the probability of each event indicator being equal to unity is equal to the hazard

probability for that time period (Muthén & Masyn, 2005). In this analysis, the hazard probability

for a given grade was defined as the probability of a student experiencing his or her first school

removal in that grade, provided that a student had not experienced a school removal in an earlier

grade. The hazard probability's relationship to time may be left unstructured such that the hazard

probability for each time interval is freely estimated rather than being constrained by, for example, a linear function of time. The hazard probability is then related to covariates through a logit link function—that is, logistic regression—so that the effect of a covariate on the timing of the first school removal is parameterized by its effect on the log hazard odds of an event during a given time interval.

Thus, we may describe a covariate's effect on the likelihood of event occurrence in terms of the hazard odds ratio (hOR). For example, if sex, coded 1 for males and 0 for females, was estimated to have a hazard odds ratio of 2.0 for the third grade, we would say that the odds of a first school removal in third grade for males was twice the odds of first school removal in third grade for females.

Covariate effects may be time-varying or time-invariant. When covariate effects are constrained to be time-invariant, the effects on the hazard probability are the same for each time interval; in other words, the hazard odds ratio is constant over time. This constraint is sometimes referred to as the proportional hazard odds assumption. When this constraint is relaxed, the covariate effects are permitted to by time-varying. For example, a time-varying effect of sex would indicate that the risk of school removal associated with being male changes as students grow older and is potentially overwritten by other, more proximal processes.

In Figure 1, the survival aspect is shown in the top part of the figure, where a latent variable "F_W" is measured by seven binary event indicators. This top portion is the model for the *within* cluster, or student-level, variability in time-to-school-removal process. The regression path representing effects of the Level 1 or student-level predictors (age, sex, race, lunch status) on "F_W" probe for proportional or time-invariant effects on the hazard probabilities. In addition, the regression path from individual aggression to "F_W" is modeled as a random effect; that is, the

influence of aggression on the time-specific hazard of removal is allowed to vary by Level 2

units (i.e., classrooms). This random effect is labeled "S".

To investigate whether the classroom level of aggressive behavior moderated the

influence of individual-level aggressive behavior, the regular discrete-time survival analysis was

extended to accommodate the multilevel structure of the data (see Figure 1, *between*—or Level

2—part). This extension was required due to the clustering of students in different classrooms. In

a multilevel discrete-time survival analysis, as was used in this study, the hazard probability for

an event may also depend on cluster-level (Level 2; in this case, classroom-level) variables.

In addition, as with other multilevel models, the effects of covariates on the individual

level (Level 1) may also vary according to cluster-level variables. This is often referred to in the

multilevel literature as a cross-level interaction. In the *between,* or Level 2, part of the model, a

latent variable "F_B" is used to estimate the hazard probability of removal among the classrooms.

The Level 2 predictor "fall of first grade classroom aggression" is allowed to predict classroom

differences in the likelihood of removal. In addition, the effect of individual aggression on school

removal is allowed to vary by classroom context, which is captured by the regression path of "S"

on "classroom aggression."

In this study, the classroom was chosen as the Level 2 unit for two reasons. From a

policy standpoint, while suspensions are a schoolwide problem, suspension is most commonly

related to behavior in the classroom and is initiated by the classroom teacher. Secondly, while

there is evidence that the leadership culture in a particular school may impact the number of

suspensions, the number of schools used in this study is too small (i.e., 19) to reliably estimate

between-school-level effects. To do that accurately would require at least 25–30 Level 2 clusters.

As noted earlier, the focus in this study is on predictor variables collected at the time of school entry and their relationship to the later school removal process. Substantively, we have argued that entering the first grade elementary classroom is a significant developmental marker where higher levels of aggression—indicating, in part, a lack of school readiness—are strongly predictive of later problems. A similar rationale applies to the classroom aggression variable. In our statistical argument we focus on one of the requirements for causality. Temporal order of predictor and outcome variables is one of the essential criteria for establishing causality (Rubin, 1974).

In this study, information regarding school removal was collected longitudinally instead of relying on retrospective teacher or student reports, while covariate information was collected prior to the occurrence of any school removal (i.e., fall of first grade), thus providing the opportunity to draw inferences about the process of school removal over time as influenced by student characteristics in fall of first grade. In addition, data in this study come from a randomized control trial, and students were randomly assigned to classrooms, teachers and conditions, thus minimizing any prior differences among students.

It can be argued that the relationship between prior aggression and consequent removal, as well as between removal and consequent aggression, can be best characterized as a reciprocal process. Using aggression as a time-varying covariate, however, disallows establishing a temporal ordering and makes it impossible to tease apart the likely within-grade reciprocal relationship between aggression and school removal. Given the high correlation between teacher ratings of aggressive behavior (as previously discussed), we argue that first grade information is a reasonable proxy for the early individual process. In addition, we focus on first grade levels of classroom aggression as a proxy for the general cohort/peer context that student's experience,

especially in the earlier years of schooling, regardless of whether they have the same teacher or exactly the same classmates from year to year.

Importantly, the population in this sample is highly mobile, and students not only change teachers from year to year but also change classrooms and schools. This type of movement requires multilevel modeling with cross-classification. In these data, later cluster sizes (i.e., classrooms) are too small to estimate such a model, since only the original study participants were followed.

All discrete-time survival analysis models were estimated in Mplus, Version 5.2 (Muthén & Muthén, 1998–2007), as described by Muthén and Masyn (2005).

Results

In building the model for time to first school removal, we began with a multilevel discrete-time survival analysis with no predictors at the classroom level, that is, an unconditional Level 2 model with only proportional main effect terms at the student level (Level 1). For all models, the baseline hazard probabilities (i.e., the hazard probabilities for each time interval at the points where all covariates are zero, analogous to intercepts in a regular regression model) were left unstructured so that we estimated seven separate baseline hazard probabilities, one for each grade. Initially, the student-level model included proportional main effect terms for sex, race, lunch status, and first grade aggression (classroom-mean centered); that is, the effects for each covariate were constrained to be time-invariant.

As the next step in the model building, we relaxed the proportionality assumption for each covariate in turn, allowing the effects of the selected covariate to differ across all grade levels, and, using the likelihood ratio chi-square test, compared the fit of the model to the proportional hazard odds model. The only covariate that showed evidence of non-proportionality

(i.e., time-varying effects) was sex, χ^2*(6, N=1169) =14.60, p<0.05*, and thus, in all subsequent

models, sex was included with non-proportional effects. This finding indicates that the risk

associated with being male changes from first to seventh grade.

In the final step of exploring the fixed effects at the student level, moderating effects of

sex and of aggression were explored. No significant interaction effects in the model were found

between sex and any of the other covariates nor between aggression and any of the other

covariates. After finalizing the student-level model (Level 1), we included first grade classroom

aggression as a predictor of variability in time to first school removal at the classroom level and

found a significant effect. We also explored allowing the effect of individual aggression to vary

across classroom, but there was no significant variability in the effect to be modeled.

Results for the final multilevel model are given in Table 2. The baseline hazard

probabilities—that is, the hazard probabilities for White females, with a grade-appropriate age,

not on free/reduced lunch, with an average level of aggression in first grade, in a classroom of

average aggression levels—were relatively stable in grades 1, 2, and 3, increasing twofold in

grades 4 and 5 and then increasing almost tenfold in grades 6 and 7. African American students

had 2.02 times the hazard odds of first school removal at any given grade compared to White

students, controlling for the effects of SES, sex, and aggression. Those students on free or

reduced lunch had 1.68 times the hazard odds of first school removal at any grade compared to

students of higher SES levels, holding the effects of race, sex, age, and aggression constant.

Students with higher levels of aggression in first grade had higher hazard odds of school removal

(1.84 times the hazard odds for every one-point increase on the TOCA-R scale), holding the

effects of race, SES, age, and sex constant. Sex was found to have a time-varying effect, with

males having higher hazard odds of first school removal compared to females across grades 1–7,

with the greatest differences in grades 3 and 4. Finally, students who were older than 6 years in

fall of first grade had 54% higher hazard odds of first school removal, holding all other

predictors constant. At the classroom level, average first grade classroom aggression was found

to have a negative effect on the hazard of first school removal, with a 50% decrease in the

average hazard odds of first school removal for a one-unit increase in the classroom average

aggression on the TOCA-R scale.[1]

In order to better understand the overall impact of the student-level and the classroom-

level covariates, we produced estimated hazard and survival probability plots of *high-* and *low-*

risk males and females from *high-* and *low-aggression* first grade classrooms. "High-risk"

students were defined as those who were African American, receiving free/reduced lunch, older

than 6 years of age at the start of first grade, and one standard deviation above the sample mean

in first grade aggression level. "Low-risk" students were defined as those who were White, not

receiving free/reduced lunch, 5 or 6 years of age at the start of first grade, and one standard

deviation below the sample mean in first grade aggression level. A "high-aggression" classroom

was defined as a classroom with an average aggression level one standard deviation above the

mean average aggression level across all first grade classrooms in the sample, while a "low

aggression" classroom was defined as a classroom with an average aggression level one standard

deviation below that mean.

Examining Figure 2 depicting the hazard probabilities for males, we can see that the

subpopulation estimated to be at greatest risk of first school removal across grades 1–7 was high-

risk males in low-aggression classrooms, followed by high-risk males in high-aggression

classrooms. In Figure 3, the discrete-time hazard probabilities are translated into cumulative

---

[1] We also explored whether variability in classroom aggression was related to the likelihood of school removal. The results revealed that classroom variation with and without the mean in the model were unrelated to school removal. Thus, we report models which include the mean of classroom aggression only.

survival probabilities. By the end of seventh grade, nearly 80% and 60%, respectively, of high-risk males in high- and low-aggression classrooms are estimated by the fitted model to experience a first school removal, compared to less than 10% of all the low-risk males (see Figure 3).

In comparison, females showed a different pattern in the hazard profile of school removal. Unlike males, high-risk females in low-aggression classrooms only had noticeably higher risk during the middle school years, followed by high-risk females in high-aggression classrooms (see Figure 4). The difference between high- and low-risk females, although large, is not as dramatic as it is for males. By the end of seventh grade, a little over 40% of high-risk females in low-aggression classrooms and a little over 20% of high-risk females in high-aggression classrooms are estimated by the fitted model to experience a first school removal, compared to less than 10% of all the low-risk females (see Figure 5).

Discussion

The purpose of this paper was to investigate the variation in the likelihood to be removed from school by race or ethnicity, sex, and poverty level of students, taking into account children's individual levels of first grade aggression as well as the level of first grade classroom aggression. Using a multilevel discrete-time survival framework, we extended previous analyses of school removal to explore the effects of early individual- and classroom-level aggression by modeling the clustering of students within classrooms and schools.

Overall, we found that demographic characteristics of the students as well as initial individual and classroom level of aggression contribute to the first onset of school removal. We replicated the findings of previous studies indicating that African American youth who live in poverty, as compared to Caucasian youth, represent the group at highest risk for school removal,

but also found those risk differences to remain when controlling for early individual levels of aggressive/disruptive behavior. Furthermore, important sex differences in the timing of risk emerged. Finally, an important finding, enabled by the multilevel model, was that the first grade classroom level of aggression contributed to the risk for school removal in a direction opposite to the effect of individual-level first grade aggression.

Demographics and Individual Aggression

Results support previous research suggesting that the disciplinary procedure of school removal may not be equitably applied across race, sex, or poverty level of students. These findings are consistent with previous research indicating that the students most likely to be removed from school are low-income, African American males (e.g., Mendez & Knoff, 2003; Skiba et al., 2002). A student's individual level of aggression is important as well, as these results also indicate that students who are initially more aggressive when they begin their formal schooling are at higher risk for school removal.

These findings are in line with existing research, but when we examined demographic characteristics combined with initial individual levels of aggression, we expanded previous research knowledge about school removal practices. Specifically, including early individual levels of aggression in the analysis did not result in a reduction in the impact of the demographic variables on the risk to be removed from school. In other words, controlling for individual levels of first grade aggression, African Americans who live in poverty were *still* more likely to be removed from school.

Results show that males as compared to females, African American students as compared to Caucasian students, and students living in poverty as compared to those not living in poverty are at much greater risk for school removal, and this phenomenon is not fully accounted for by

differences in students' initial levels of aggression. Also, given the stability in aggression levels that has been previously established, the phenomenon is not likely to be fully accounted for by students' time-concurrent levels of aggression (although not explicitly tested in these analyses).

Upon further examination of our high-risk group of students (i.e., minority status, free/reduced lunch recipient, elevated levels of aggression, more than 6 years old in first grade), important sex differences emerged in regard to the timing of school removal. For males and females overall, high-risk youth beginning in low-aggression classrooms were the most likely to be removed, followed by high-risk youth beginning in highly aggressive classrooms. While for males this ordering was true from grades 1 through 7, important differences were detected for females. During the elementary school years, low- and high-risk females, independent of the initial classroom environment, showed an almost identical risk of being removed. However, once middle school started, females showed a pattern similar to that of males: high-risk females beginning in low-aggression classroom displayed the highest likelihood for removal.

One reason for this pattern may be the lower overall removal rate for females compared to males, especially in elementary school. Although the variables associated with risk become more apparent for females in middle school, for both sexes there was an increase in the incidence of school removal in middle school, which is consistent with previous literature indicating that the greatest numbers of suspensions occur during the middle school years (e.g., Skiba et al., 1997).

This study does not shed light on the reasons for the persistent disparities by race and ethnicity, sex, and poverty in school removal practices. Although we could speculate on possible reasons, we cannot ascertain from our data precisely what accounts for this disparity because it

would require access to observational data. What the results do argue for, however, is more research that specifically addresses reasons for the disproportionality.

For example, perhaps demographic characteristics represent variables we did not include in our model, such as parent involvement. Families living in low-income environments live with multiple stressors and, as a result, may not be as able to be integrally involved in their child's schooling. The decision to remove a student from the classroom is often at the discretion of the teacher; if a teacher works with a highly involved parent, perhaps a unique behavioral reinforcement plan will be developed for a student with behavioral difficulties, whereas a teacher working with a parent viewed as "uninvolved" may see suspension as the only option to intervene with the child.

Alternatively, some schools are more successful than others in engaging parents or in enhancing the capacity of teachers and staff to employ positive approaches to discipline. Similarly some teachers have a better capacity to provide authoritative behavior management and/or to work with families (Quinn, Osher, Hoffman, & Hanley, 1998). This is speculation, but it suggests that there may be decision making processes involved in school removal practices that go beyond demographic characteristics that need to be examined to truly start understanding this long-standing documentation of disparities in school removal practices.

Influence of Classroom Levels of Aggression

When the beginning classroom context was examined (i.e., first grade classroom levels of aggression), we found that the risk of school removal varied by these initial classroom level of aggression; however, the direction of the influence was unanticipated. Specifically, students in highly aggressive first grade classrooms were *less likely* to be removed than their peers in low-aggression classrooms. Furthermore, examination of our student risk groups revealed that

students considered at high risk (i.e., African American males who receive free/reduced lunch and show above-average individual aggression) in highly aggressive first grade classrooms were *less likely* to be removed from school when compared to similar high-risk youth in low-aggression classrooms. We also did not find the anticipated cross-level interaction effect between classroom aggression levels and individual aggressive behavior on school removal.

These results stand in contrast to the findings of Kellam et al. (1998), which showed that aggressive males who were in badly managed classroom had a significantly greater likelihood to also be highly aggressive in early middle school (i.e., sixth grade) than their highly aggressive peers in low-aggression classrooms. However, all these findings together document the importance of the early classroom context for later development.

Although we do not have data in this study related to classroom practices or teacher disciplinary approaches, we can speculate about reasons for the negative association between early classroom aggression levels and risk of school removal. Since teachers rather than administrators assign the majority of discipline consequences (Skiba et al., 2002), it is possible that the explanation for the association lies in the response of teachers to the behavioral context of the classroom as well as to the behavior of individual students. There is a good empirical literature emphasizing that teacher behavior can set the stage for disciplinary problems as well as influence student–teacher interactions (Osher, Cartledge, Oswald, Artiles, & Coutinho, 2004; Osher et al., 2004). There is also evidence for great variability in teacher use of disciplinary actions such as school removal. For example, Skiba et al. (1997) found that, in one middle school, 66% of all disciplinary referrals came from 25% of the teachers.

Perhaps when looking at the fall of first grade classroom context as a whole, teachers in classrooms composed of highly aggressive students develop higher tolerances or have a higher

threshold for what is considered "removable" behavior. Alternatively, a teacher in a high-aggression classroom may be more likely to attribute individual rule-breaking behavior to the classroom context and, thus, be less likely to apply removal to individual students as a behavioral consequence.

Further, the association found here may also be related to the link between classroom behavior and classroom management strategies. For example, recent work has shown that those males who were on a stable aggression trajectory were also at risk for an array of different negative outcomes (e.g., school dropout, juvenile and adult arrest, etc.) and that they benefited the most from the intervention aiming to reduce aggressive behavior. This intervention tested the hypothesis that negative outcomes can be prevented by improving classroom behavior management and thereby improving proximal (aggressive, disruptive behavior) and distal (school dropout, arrest) outcomes (Petras et al., 2008).

Clearly, there is an ample opportunity for future studies to probe the association found here between early classroom aggression levels and risk of school removal in addition to exploring other contextual sources of influence at the teacher, classroom, and school levels.

Implications

From an educational policy perspective, it is important to know who is at greatest risk for school removal, when the greatest risk for students to be removed exists, and what variables influence that risk. This knowledge will assist researchers and practitioners in identifying early behavioral risk factors and time periods that may be potential targets for classroom- or school-based preventive interventions with the goal of preventing school removal and ultimately preventing the resulting negative educational outcomes.

Schools are considered "risk-prone contexts" (Steinberg & Avenevoli, 1998) where children with behavioral problems elicit punitive reactions from teachers and peers (Reid & Eddy, 1997; Walker & Buckley, 1973), but school environments are also important contexts that may foster resilience (Nettles, Mucherah, & Jones, 2000; Rutter, Maughan, Mortimore, & Ouston, 1979) and where students can succeed and receive necessary services. Importantly, from a prevention perspective, school settings are unique settings where interventions can be introduced to a large number of children at the lowest cost to society (Black & Krishnakumar, 1998; Woodruff et al., 1999). Alternatively, if factors other than students' behavior influence the risk for school removal, interventions aiming only at the student level cannot be expected to be effective. This argues for additional interventions targeting teacher, school, and/or district practices and policies.

Preventive interventions. The pattern of school removal by grade levels is informative for researchers and practitioners interested in preventive intervention to reduce reliance on school removal procedures as a disciplinary tactic. For example, it was clear that the use of school removal practices increased at times of transition, such as moving from elementary to middle school, indicating that these may be prime times to specifically target interventions to improve students behavior, including classroom interventions that may help teachers better manage student behavior without an overreliance on school removal.

Also, given the impact of aggressive behavior on the risk for school removal, school-based interventions aimed at early onset of aggressive behavior would be expected to show an effect on decreasing removal rates. However, as discussed earlier, given the impact of demographic characteristics, interventions only aiming at the students' behavior are not expected to eliminate the disparities in school removal practices completely.

School policies. Although teachers and administrators commonly view suspensions as a result of specific student behavior, our results support previous reports that other factors—such as a child's race, sex and socioeconomic status, or other yet unaccounted for variables—, play an important role. Given that the Elementary and Secondary Education Act/No Child Left Behind Act mandates that schools maximize the opportunity to learn for all children, regardless of their backgrounds, this differential application of school discipline policies is troublesome. It indicates that schools are facing the difficult challenge of maintaining an orderly environment and promoting student learning by ultimately tending to exclude low-income minority males and females from the educational endeavor.

The finding that high-risk males in above-average aggression first grade classrooms are less likely to be removed than their high-risk peers in low-aggression classrooms suggests that there may be contextual variables at play in the decision to remove students from school, and further investigations should explore the potential influence of contextual variables on school removal practices and the extent to which they are malleable by policy-focused interventions. For example, the results are consistent with the possibility that teacher-level factors, such as methods of classroom management, may be influencing the school removal process. If teachers feel ill-equipped to prevent and handle behavior problems in the classroom, they may be more likely to fall back on the use of suspension to address misbehavior or engage students in power struggles that may serve to only escalate behavior which could lead to increased use of suspension (Skiba et al., 2002).

In addition to the range of student misbehavior and possible teacher response by classroom, schooling policies and practices may add great variation to the numbers of recorded school removal cases. Although we were not able to directly examine school-level effects in the

analysis, the data for the schools in the study sample illustrate the high degree of variability in

school removal rates, even in the same school district. In Figure 6, the distribution of single and

repeat removal among the removed students across the different schools in the study is shown.

The prevalence of single school removals across the schools averages about 49.8%, while

ranging from 16.7 to 75%, and that for repeat offenses averages around 50%, ranging from 25 to

83.3%.

Since there is good evidence that schools with high suspension rates typically have high

student-teacher ratios, low academic quality ratings, administrative indifference to school

climate, reactive disciplinary programs, and ineffective governance (Christle, Nelson, &

Jolivette, 2003), it would be worthwhile to explore in future research how sources of variability

in school removal rates on the school level affect individual student risk of school removal

across time. Such work would build upon that of Wu and colleagues (1982), who reported that

suspension rates varied as a function of school and district characteristics, such as teacher

attitudes, administrative centralization, quality of school governance, teacher perception of

student achievement, and racial makeup of the school. They concluded that students could more

successfully reduce their future risk of removal by transferring to a school with lower removal

rates than by improving their attitudes or reducing their misbehavior (Wu, Pink, Crain, & Moles,

1982).

Limitations

Although we have tried to more comprehensively examine the inequity in school removal

practices than previous research has, the problem is more complex than our study could address.

There could be many possible reasons for the disparity in removal practices that could not be

examined with these data. For example, a teacher may have prior information about a student

through interactions with the student or reports from other teachers that may increase or decrease the teacher's likelihood to remove that student. Furthermore, there may be aspects related to a student's dress, language, or nonverbal behaviors that influence a teacher's attitude/willingness to remove the student from school in a nonrandom fashion.

In addition, although the sample was representative of all students entering first grade in a large urban district, the sample was relatively homogeneous with respect to race in that the African American and White students comprised almost 99% of the sample. Therefore, a comparison of other race or ethnic groups, such as Latino populations, was not possible. A more diverse sample might reveal other differences in school removal practices for other races or ethnicities.

Last, but not least, the primary explanatory predictors (individual- and classroom-level aggression) are one point-in-time assessment and focus only on the level of aggressive/disruptive behavior displayed at school entry, thus emphasizing the importance of this early transitional period. Further research is required to study the potentially reciprocal interplay of aggressive behavior with the occurrence of school removal by treating aggression as a co-occurring longitudinal process. Importantly, this would help elucidate whether the effect of aggression on school removal is constant over time as well as clarify the effect of school removal on later aggression.

In addition, other measures of students' challenges to adjust to the elementary classroom, such as shy/withdrawn behavior as well as concentration and attention problems, should be examined. For example, it is possible that students' misbehavior is caused by their inability to follow the academic expectations and a consequent tendency to act out to relieve their frustration. Moreover, measures of teacher responses to students' misbehavior could provide

valuable information about the ways in which the classroom behavioral context ultimately impacts the school removal process.

Summary and Conclusion

The current study showed that race and ethnicity, sex, poverty level, and early individual levels of aggression all have strong relationships to school removal. Furthermore, it showed that the early classroom context matters: the overall level of classroom aggression also was significantly related to the likelihood of school removal.

A prime advantage of this study was the use of an epidemiological sample that reduced selection bias. A second advantage was the application of an advanced modeling technique—multilevel discrete-time survival analysis—that allowed us to make full use of the longitudinal data regarding the timing and occurrence of school removal while accounting for the nested structure of the data and covariates measured on different levels. The ongoing documentation of disparities in school removal practices, and the recent discussions regarding the renewal of the Elementary and Secondary Education Act, taken together, represent an excellent opportunity to address these issues on a federal level and to put in place policies to promote equal learning opportunity for children of different ethnic and socioeconomic backgrounds.

References

Alexander, K. L., Entwisle, D. R., & Horsey, C. (1997). From first grade forward: Early

      foundations of high school dropout. *Sociology of Education, 70*, 87-107.

Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life

      course perspective: Early risk factors at home and school. *Teachers College Record,

      103*(5), 760-822 . Retrieved from http://www.tcrecord.org/

American Academy of Pediatrics (2003). Out-of-school suspension and expulsion. *Pediatrics,

      112*, 1206-1209.

Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom

      environment influences on aggression, peer relations, and academic focus. *Journal of

      School Psychology, 42*, 115-133.

Black, M. M., & Krishnakumar, A. (1998). Children in low income, urban settings: Interventions

      to promote mental health and well-being. *American Psychologist, 53*, 635-646.

Bowditch, C. (1993). Getting rid of troublemakers: High school disciplinary procedures and the

      production of dropouts. *Social Problems, 40*, 493-509.

Brophy, J. E. (1996). *Teaching problem students*. New York, NY: Guilford Press.

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

      DC: The Washington Research Project, Inc.

Civil Rights Project (2000). *Opportunities suspended: The devastating consequences of zero

      tolerance and school discipline policies.* Cambridge, MA: Harvard University.

Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school

      students. *Journal of School Psychology*, *36*, 59-82

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and

problem behavior. *American Psychologists, 54*, 755-764.

Dolan, L., Kellam, S. G., Brown, C. H., & Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S.,
et al. (1993). The short-term impact of two classroom-based preventive interventions on
aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental
Psychology, 14*, 317-345.

Dupper, D. R., & Bosch, L. A. (1996). Reasons for school suspensions. *Journal for a Just and
Caring Education, 2*, 140-150.

Ekstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who drops out of high
school and why? Findings from a national study. *Teachers College Record, 87*, 357-373.

Ensminger, M. E., Forrest, C. B., Riley, A. W., Kang, M., Green, B. F., Starfield, B., et al.
(2000). The validity of measures of socioeconomic status of adolescents. *Journal of
Adolescent Research, 15*, 392-419.

Ensminger, M. E., Kellam, S. G., & Rubin, B. R. (1983). School and family origins of
delinquency: Comparisons by sex. In K. T. Van Dusen & S. A. Mednick (Eds.),
Prospective studies of crime and delinquency (pp. 73-97). Boston, MA: Kluwer-Nijhoff
Publishing.

Ensminger, M. E., & Slusarcick, A. L. (1992). Paths to high school graduation or dropout: A
longitudinal study of first grade cohort. *Sociology of Education, 65*, 95-113.

Entwisle, D. R., & Hayduk, L. A. (1988). Lasting effects of elementary school. *Sociology of
Education, 61*, 147-159.

Entwistle, N., Nisbet, J., & Bromage, A. (2005). Teaching-learning environments and student
learning in electronic engineering. In L. Verschaffel, E. De Corte, G. Kanselaar & M.
Valcke (Eds.), *Powerful environments for promoting deep conceptual and strategic*

learning (pp. 175-198). Leuven, Belgium: Leuven University Press.

Giroux, H. A. (2001). *Theory and resistance in education: Towards a pedagogy for the*

       opposition. Westport, CT: Bergin & Garvey.

Graue, M. E., & DiPerna, J. (2000). Redshirting and early retention: Who gets the "Gift of Time"

       and what are its outcomes? *American Educational Research Journal, 37*, 509-534.

Gregory, J. F. (1997). Three strikes and they're out: African American boys and American

       schools' responses to misbehavior. *International Journal of Adolescence and Youth, 7*,

       25-34.

Guttman, J. (1982). Pupils', teachers' and parents' causal attributions for problem behavior in

       schools. *Journal of Educational Research, 76*(1), 14-21.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of

       children's school outcomes through eighth grade. *Child Development, 72*, 625- 638.

Hawkins, J. D., VonCleve, E., & Catalano, R. F. (1991). Reducing early childhood aggression:

       Results of a primary prevention program. *Journal of the American Academy of Child and*

       Adolescent Psychiatry, 30, 208-217.

Henry, D., Guerra, N., Huesmann, R., Tolan, P., VanAcker, R., & Eron, L. (2000). Normative

       influences on aggression in urban elementary school classrooms. *American Journal of*

       Community Psychology, 28, 59-81.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in

       childhood and adolescence: Causal relationships and underlying mechanisms.

       Psychological Bulletin, 111(1), 127-155.

Huesman, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, J. (1984). Stability over aggression

       over time and generations, *Developmental Psychology, 20*, 1120-1134.

Hughes, J. N., Cavell, T. A., & Jackson, T. (1999). Influence of the teacher–student relationship on childhood conduct problems: A prospective study. *Journal of Clinical Child Psychology, 28*, 173-184.

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

Johnson, B., Whitington, V., & Oswald, M. (1994). Teachers' view on school discipline: A theoretical framework. *Cambridge Journal of Education, 24*(2), 261-276.

Kellam, S. G., Brown, C. H., Rubin, B. R., & Ensminger, M. E. (1983). Paths leading to teenage psychiatric symptoms and substance abuse: developmental epidemiological studies in Woodlawn. In S. B. Guze, F. J. Earls, & J. E. Barrett (Eds.), *Childhood Psychopathology and Development* (pp. 17-51). New York: Raven Press.

Kellam, S. G., & Ensminger, M. E. (1980). Theory and method in child psychiatric epidemiology. In F. Earls (Ed.), *Studying children epidemiologically* (pp. 145-180). *International Monograph Series in Psychosocial Epidemiology*, Vol. 1: New York: Neale Watson Academic Publishers.

Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology, 10*, 165-186.

Kellam S. G., & Rebok G. W. (1992). Building developmental and etiological theory through epidemiological based preventive intervention trials. In: J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behavior: Interventions from birth through adolescence* (pp. 162-195). New York: Neale Watson Academic Publishers.

Kids First Coalition (1999). *Locked out: Exposing the suspension epidemic in the Oakland*

Public Schools. Oakland, CA: Kids First Coalition.

Krezmien, M. P., Leone, P. E., & Achilles, G. M. (2006). Suspension, race, and disability: Analysis of statewide practices and supporting. *Journal of Emotional and Behavioral Disorders, 4*, 217-226.

Leone, P. E., Mayer, M. J., Malmgren, K., & Meisel, S. M. (2000). School violence and disruption: Rhetoric, reality, and reasonable balance. *Focus on Exceptional Children, 33*, 1-20.

May, D. C., & Kundert, D. K. (1995). Does delayed school entry reduce later grade retentions and use of special education services? *Remedial & Special Education, 16*, 288-295.

McFadden, A. C., Marsh, G. E., Price, B. J., & Hwang, Y. (1992). A study of race and gender bias in the punishment of school children. *Education and Treatment of Children, 15*, 140-146.

Mendez, L. M. R., & Knoff, H. M. (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.

Mendez, L.M.R., Knoff, H. M., & Ferron, J. M. (2002). School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district. *Psychology in the Schools, 39*, 259-277.

Morgan, H. (1991). *Race and gender issues in school suspension.* Paper presented at the annual meeting of the American Education Research Association, Chicago, Il.

Morrison, G. M., Anthony, S., Storino, M., Cheng, J., Furlong, M. F., & Morrison, R. L. (2001). School expulsion as a process and an event: Before and after effects on children at-risk for school discipline. *New Directions for Youth Development: Theory, Practice,*

Research, 92, 45-72.

Muthén, B., and Masyn, K. (2005). Discrete-time survival mixture analysis. *Journal of Educational and Behavioral Statistics, 30*, 27-58.

Muthén, L., & Muthén, B. (1998-2007). *Mplus User's Guide, 5th Edition.* Los Angeles, CA: Muthén & Muthén.

Nettles, S., Mucherah, W., & Jones, D. (2000). Understanding resilience: The role of social resources. *Journal of Education for Students Placed at Risk, 5*, 47-60.

Newcomb, M. D., Abbott, R. D., Catalano, R. F., Hawkins, J. D., Battin-Pearson, S., & Hill, K. (2002). Mediational and deviance theories of late high school failure: Process roles of structural strains, academic competence, and general versus specific problem behavior. *Journal of Counseling Psychology, 49*(2), 172-186.

Nichols, J. D. (1999). *An exploration of discipline and suspension data.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Nichols, J. D., Ludwin, W. G., & Iadicola, P. (1999). A darker shade of gray: A year-end analysis of discipline and suspension data. *Equity & Excellence in Education, 32*, 43-55.

Office for Civil Rights, U.S. Department of Education (2002). *Elementary and Secondary School Civil Rights Compliance Reports*. Retrieved from http://www.ed.gov/offices/OCR/data.html/

Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86*, 852- 875.

Oppenheimer, J., & Ziegler, S. (1988) *Suspension, alternatives to suspension and other approaches to discipline*. Toronto, Ontario: Toronto Board of Education.

Osher, D., Cartledge, G., Oswald, D., Artiles, A. J., & Coutinho, M. (2004). Issues of cultural

and linguistic competency and disproportionate representation. In R. Rutherford, M.

Quinn, & S. Mather (Eds.), *Handbook of research in behavioral disorders* (pp. 54-77).

New York, NY: Guilford Publications.

Osher, D., Woodruff, D., & Sims, A. (2002). Schools make a difference: The relationship

between education services for African American children and youth and their

overrepresentation in the juvenile justice system. In D. Losen (Ed.), *Minority issues in*

special education (pp. 93-116). Cambridge, MA: The Civil Rights Project, Harvard

University and the Harvard Education Publishing Group.

Osher, D., Van Aker, R., Morrison, G., Gable, R., Dwyer, K., & Quinn, M. (2004). Warning

signs of problems in schools: Ecological perspectives and effective practices for

combating school aggression and violence. *Journal of School Violence, 2*(3), 13-37.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Petras, H., Chilcoat, H., Leaf, P., Ialongo, N., & Kellam, S. (2004a). The utility of teacher ratings

of aggression during the elementary school years in identifying later violence in

adolescent males. *Journal of the American Academy of Child and Adolescent Psychiatry,*

1, 88-96.

Petras, H., Ialongo, N., Lambert, S. F., Barrueco, S., Schaeffer, C. M., Chilcoat, H., & Kellam, S.

(2005). The utility of elementary school TOCA-R scores in identifying later criminal

court violence amongst adolescent females. *Journal of the American Academy of Child*

and Adolescent Psychiatry, 8, 790-797.

Petras, H., Kellam, S., Brown, H., Ialongo, N., Muthén, B., & Schaeffer, C. (2008).

Developmental epidemiological courses leading to antisocial personality disorder and

violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and Alcohol Dependence, 95*, S45-S59.

Petras, H., Schaeffer, C., Ialongo, N., Hubbard, S., Muthén, B., Lambert, S., Poduska, J., & Kellam, S. (2004b). When the course of aggressive behavior in childhood does not predict antisocial outcomes in adolescence and young adulthood: An examination of potential explanatory variables. *Development and Psychopathology, 16*, 919-941.

Quinn, M. M., Osher, D., Hoffman, C. C., & Hanley, T. V. (1998). *Safe, drug-free, and effective schools for ALL students: What Works!* Washington, DC: Center for Effective Collaboration and Practice, American Institutes for Research.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Reardon, S. F., Brennan, R. T., & Buka, S. L. (2002). Estimating multi-level discrete-time hazard models using cross-sectional data: Neighborhood effects on the onset of adolescent cigarette use. *Multivariate Behavioral Research, 37*, *297-330.*

Reid, J. B., & Eddy, J. M. (1997). The prevention of antisocial behavior: Some considerations in the search for effective interventions. In D. Stoff, J. Breiling, & J. D. Maser (Eds.), *The handbook of antisocial behavior* (pp. 343-356). New York: Wiley.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology, 30*, 787-804

Robins, L. N. (1978). Sturdy childhood predictors of adult antisocial behavior: Replications from longitudinal studies. *Psychological Medicine, 8*, 611-622.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology, 66*, 688-701.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children.* Cambridge, MA: Harvard University Press.

Seymour, P. H. K. (1999). Cognitive architecture of early reading. In I. Lundberg, F. E. Tùnnessen, & I. Austad (Eds.), *Dyslexia: Advances in theory and practice* (pp. 59-73). Dordrecht, Netherlands: Kluwer.

Sheets, R. H. (1996). Urban classroom conflict: Student-teacher perception, ethnic integrity, solidarity and resistance. *Urban Review, 28*(2), 165-183.

Singer, J. D. and Willett, J. B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics, 18*, 155-195.

Skiba, R. J., & Knesting, K. (2002). Zero tolerance, zero evidence: An analysis of school disciplinary practice. In R. J. Skiba & G. G. Noam (Eds.), *Zero tolerance: Can suspension and expulsion keep schools safe?* (pp. 17-43). San Francisco, CA: Jossey-Bass.

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*, 317-342.

Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children, 20*, 295-315.

Sprague, J., & Walker, H. (2000). Early identification and intervention for youth with antisocial and violent behavior. *Exceptional Children, 66*, 367-379.

Stage, S. A. (1997). A preliminary investigation of the relationship between in-school suspension and the disruptive classroom behavior of students with behavior disorders. *Behavioral Disorders, 23*, 57-76.

Steinberg, L., & Avenevoli, S. (1998). Disengagement from school and problem behavior in adolescence: A developmental-contextual analysis of the influences of family and part-time work. In R. Jessor (Ed.), *New perspectives on adolescent risk behavior* (pp. 392-424). Cambridge, UK: Cambridge University Press.

Stevenson, H. W., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading, *Child Development, 57*(3), 646-659.

Thomas, D. (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology, 18*(2), 471-487.

Thornton, C. H., & Trent, W. T. (1988). School desegregation and suspension in East Baton Rouge Parish: A preliminary report. *Journal of Negro Education, 57*(4), 482-501.

Townsend, B. L. (2000). The disproportionate discipline of African American learners: Reducing school suspension and expulsion. *Exceptional Children, 66*(3), 381-391.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-354.

U.S. Department of Education, National Center for Education Statistics. (2003). *Status and trends in the education of Hispanics* (NCES 2003–008). Washington, DC: U.S. Department of Education.

U.S. Department of Education, Office for Civil Rights. (2002). *Elementary and secondary school survey: State and national projections for enrollment and selected items by race/ethnicity, and sex*. Washington, DC: U.S. Department of Education.

U.S. Public Health Service. (2001). *Mental health: Culture, race, and ethnicity. A supplement to mental health: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

Verdugo, R. R. (2002). Race-ethnicity, social class, and zero-tolerance policies: The cultural and structural wars. *Education and Urban Society, 35*, 50-75.

Walker, H. M., & Buckley, N. K. (1973). Teacher attention to appropriate and inappropriate classroom behavior: An individual case study. *Focus on Exceptional Children, 5*, 5-11.

Wentzel, K. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73*, 287-301.

Werthamer-Larsson, L., Kellam, S. G., & Wheeler, L. (1991). Effect of first-grade classroom environment on child shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19*, 585-602.

Woodruff, D. W., Osher, D., Hoffman, C. C., Gruner, A., King, M. A., Snow, S. T., & McIntire, J. C. (1999). The role of education in a system of care: Effectively serving children with emotional or behavioral disorders. In S*ystems of Care: Promising Practices in Children's Mental Health,* 1998 Series, Volume III (pp. 13-16). Washington, DC: Center for Effective Collaboration and Practice, American Institutes for Research.

Wright, J. C., Giammarino, M., & Parad, H. W. (1986). Social status in small groups: Individual-group similarity and the social misfit. *Journal of Personality and Social Psychology, 50*, 523-536.

Wu, S. C., Pink, W. T., Crain, R. L., & Moles, O. (1982). Student suspension: A critical

reappraisal. *The Urban Review, 14*, 245-303.

Table 1

Distribution of School Removal between First and Seventh Grade by Sex and Race (n=267)[*]

	Male African American	Male Non-African American	Female African American	Female Non-African American
	f (%)	f (%)	f (%)	f (%)
Grade 1	8 (3.0%)	4 (1.5%)	1 (0.4%)	0 (0.0%)
Grade 2	15 (5.6%)	3 (1.1%)	1 (0.4%)	0 (0.0%)
Grade 3	13 (4.9%)	8 (3.0%)	1 (0.4%)	1 (0.4%)
Grade 4	13 (4.9%)	6 (2.2%)	5 (1.9%)	1 (0.4%)
Grade 5	16 (6.0%)	7 (2.6%)	2 (0.7%)	2 (0.7%)
Grade 6	39 (14.6%)	11 (4.1%)	29 (10.9%)	4 (1.5%)
Grade 7	35 (13.1%)	3 (1.1%)	35 (13.1%)	4 (1.5%)
Total	139 (52.1%)	42 (15.7%)	74 (27.7%)	12 (4.5%)

Note. [*] Sex differences were significant at the 0.05 level in every grade for both ethnic groups. In addition, for non African-American gender differences were only significant in third grade, while for African American gender differences were significant in first to sixth grade and only marginally significant in seventh grade.

Table 2

Results for Final Multilevel Survival Model (n=1169)

Covariates	Estimate[^]	SE	p-value	hOR[^^]
Student Level (Level 1)				
African American	0.70*	0.16	<0.01	2.02
Free/Reduced Lunch	0.52*	0.12	<0.01	1.68
Aggression in Fall 1st Grade (classroom-mean centered)	0.61*	0.07	<0.01	1.84
Age in 1st Grade (1=age7-9)	0.43	0.16	<0.01	1.54
Grade 1: Male	2.28*	1.03	0.03	9.82
Grade 2: Male	2.80*	0.98	<0.01	16.36
Grade 3: Male	2.33*	0.73	<0.01	10.31
Grade 4: Male	1.18*	0.47	0.01	3.24
Grade 5: Male	1.84*	0.55	<0.01	6.27
Grade 6: Male	0.58*	0.30	0.05	1.78
Grade 7: Male	0.14	0.27	0.61	1.15
Residual Variance	0.01	0.14	0.96	---
Classroom Level (Level 2)				
Classroom Aggression	-0.68*	0.16	<0.01	0.51
Residual Variance	0.03	0.06	0.61	---

Note. *p<.05.
[^] Estimates are adjusted for cohort status.
[^^] hOR is the abbreviation for hazard odds ratio.

Figure Captions

Figure 1. Multilevel discrete-time survival model diagram.

Figure 2. Model-estimated hazard probabilities for males by risk status and classroom aggression.

Figure 3. Model-estimated survival probabilities for males by risk status and classroom aggression.

Figure 4. Model-estimated hazard probabilities for females by risk status and classroom aggression.

Figure 5. Model-estimated survival probabilities for females by risk status and classroom aggression.

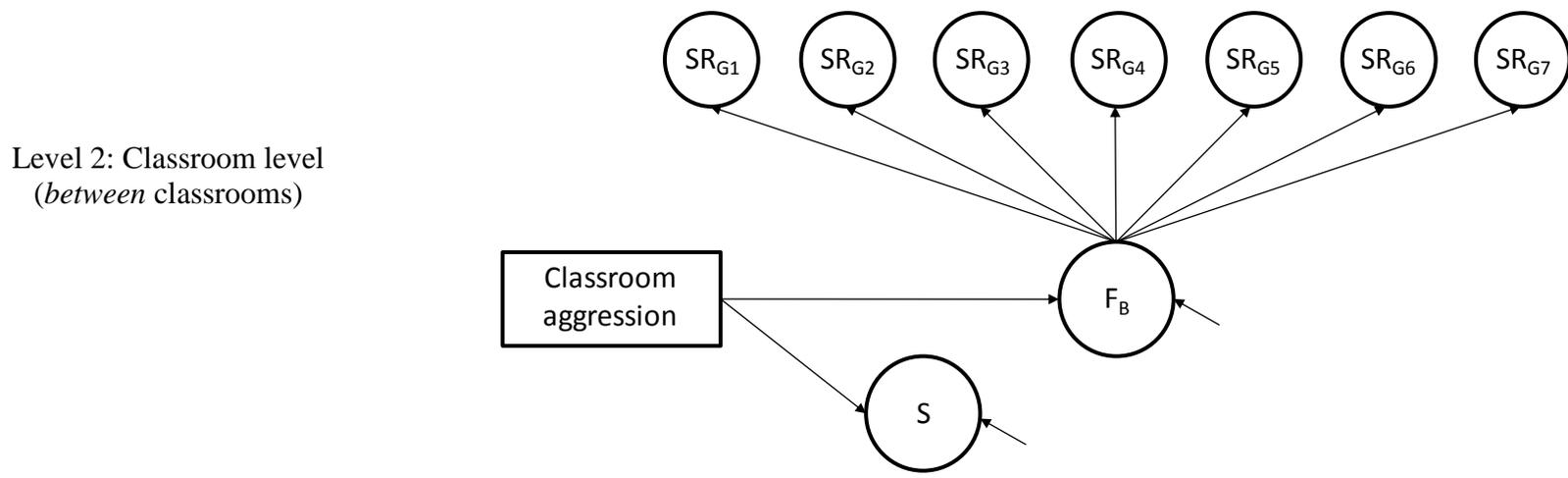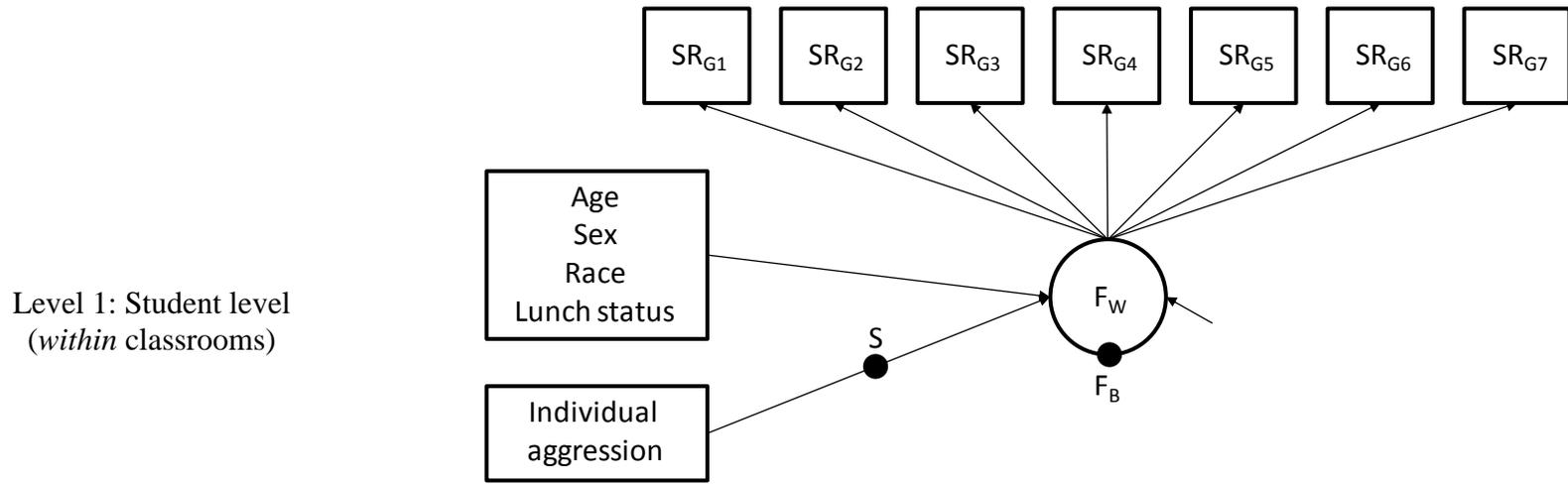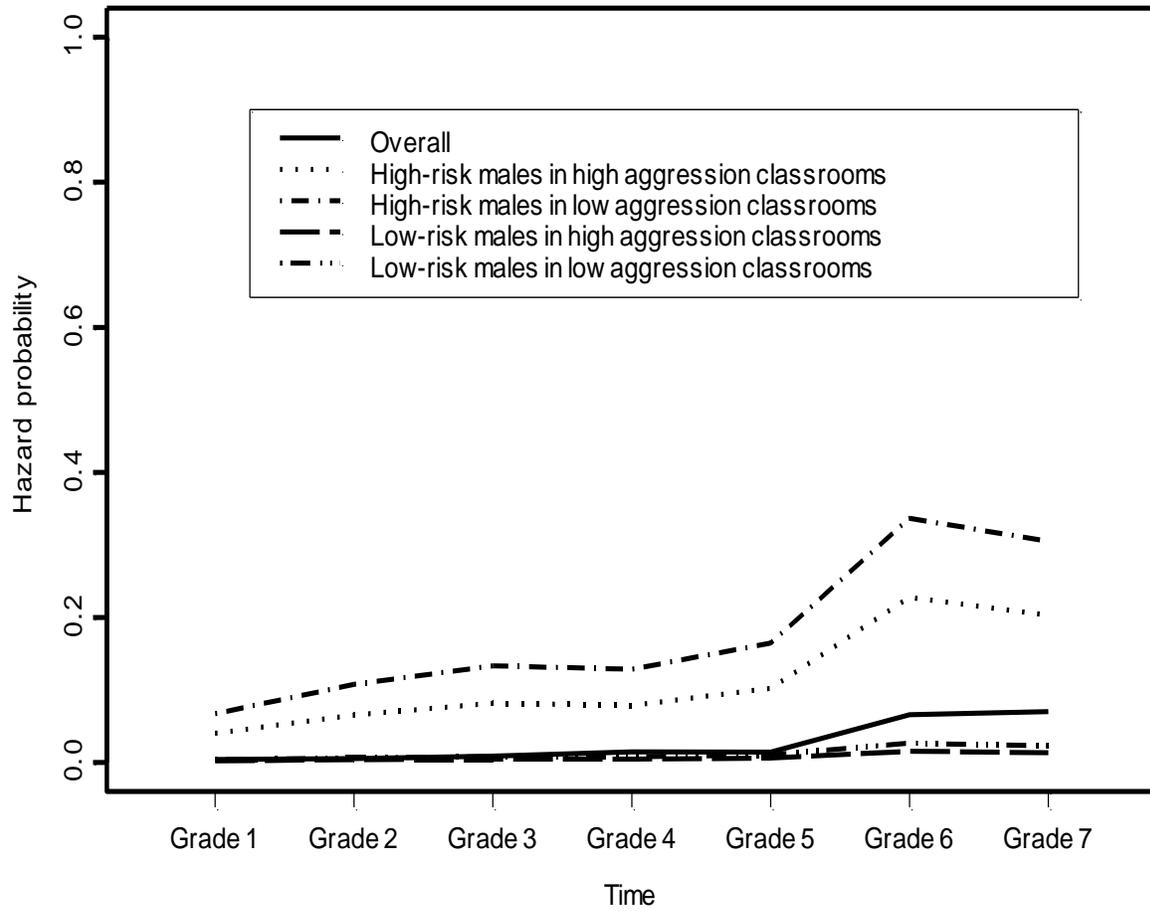*Figure 6*. Distribution of school removal across schools.

Figure 1.

Level 1: Student level
(*within* classrooms)

Level 2: Classroom level
(*between* classrooms)

Figure 2.

Figure 3.

Figure 3 shows a survival analysis plot. The y-axis is labeled "Survival probability" ranging from 0.0 to 1.0, and the x-axis is labeled "Time" with categories Grade 1 through Grade 7. The legend identifies five lines:

- Overall
- High-risk males in high aggression classrooms
- High-risk males in low aggression classrooms
- Low-risk males in high aggression classrooms
- Low-risk males in low aggression classrooms

Figure 4.

Figure 5.

Figure 6.