Life course turning points: The effect of grade retention on physical aggression

DANIEL S. NAGIN, LINDA PAGANI, RICHARD E. TREMBLAY, AND FRANK VITARO

Carnegie Mellon University; and University of Montreal

Abstract

Our objective is to advance the life course analytical framework by demonstrating a model for testing two of its tenets. The first is whether the individual’s developmental history conditions the response to a turning point event. The second is whether the influence of a major life event upon an individual’s developmental course depends upon the timing of the event. We test both propositions in an analysis of the effect of grade retention on a child’s trajectory of physical aggression. Our analysis is based on data from a longitudinal study of 1,037 boys from schools in the lowest socioeconomic areas in Montreal, Canada. We find clear evidence that a developmental history of physical aggression conditions the child’s response to grade retention. The evidence on whether the timing of retention affects this response is less clear.

Life course research aims to broaden inquiry into the effects of life’s turning points by analyzing transitions in the context of an individual’s developmental course. In this paper we demonstrate an analytical framework for testing two tenets of the life course paradigm—the effect of a turning point on a specified behavior depends upon (a) the individual’s developmental history of that behavior and (b) the timing of the turning point.

The point of departure for this analysis is the four developmental trajectories of physical aggression from childhood to adolescence depicted in Figure 1: a chronic physical aggression trajectory comprising boys (4%) displaying persistently high levels of physical aggression; a high level declining trajectory comprising boys (28%) who in kindergarten displayed a high level of physical aggression, but who thereafter declined to virtual cessation; and a “low” trajectory comprising boys (17%) who throughout rarely displayed physical aggression. These trajectories, which were first reported in Nagin and Tremblay (1999), derive from annual teacher assessments of physical aggression at ages 6 and 10–15 for over 1,000 French-speaking Canadian males from Montreal, Canada. Details of the data set and the method used in estimating the trajectories are described below.

This study was made possible by grants from Québec’s CQRS and FCAR funding agencies, Canada’s NHRDP and SSHRC funding agencies, the Molson foundation, the National Consortium on Violence Research, and the National Science Foundation (SBR-9511412 and SES-9911370). We thank Lisa Broidy, David Farrington, Holly Foster, John Laub, Terrie Moffitt, anonymous reviewers, and seminar participants at the University of Maryland and the University of Cambridge for their input; Hélène Beauchesne and Lucille David for supervising the data collection; and Lyse Desmarais–Gervais, Pierre McDuff, and Muriel Rorive for managing the data bank.

Address correspondence and reprint requests to: Daniel Nagin, 2105 Hamburg Hall, Carnegie Mellon University, Pittsburgh, 15213 PA; E-mail: dn03@andrew.cmu.edu.
We test the two propositions outlined above in an analysis of the effect of grade retention on each of these physical aggression trajectories. Our aim is to examine whether the trajectories depicted in Figure 1 are deflected by grade retention. Specifically, the first proposition is tested by whether the direction and magnitude of the change in physical aggression resulting from grade retention depends upon physical aggression trajectory group. The second dependency is tested by whether the change in physical aggression for each trajectory group is conditioned by the child’s age when retained.

In addition we have a methodological objective of generalizing the group-based statistical method that was used to identify the trajectories depicted in Figure 1 (Jones, Nagin, & Roeder, 2001; Nagin, 1999; Nagin & Land, 1993; Nagin & Tremblay, 2001; Roeder, Lynch, & Nagin, 1999). Sampson and Laub (1992, p. 66) observe that “trajectories refer to the long-term patterns and consequences of behavior.” They go on to observe that “transitions are marked by specific life events . . . that are embedded in trajectories and evolve over shorter time spans.” The generalization of the method allows for estimation of the impact of a transition like grade retention on each trajectory while controlling for enduring individual characteristics that may affect both the individual’s trajectory and the likelihood of experiencing the transition. This methodological objective aims to provide statistical expression to the concepts of trajectories and turning points.

Elder (1985, pp. 31–32) defines a turning point as “a change in state that is more or less abrupt.” Marriage, parenthood, migration, entering the work force and military service are important examples of major life turning points that have been the central focus of life course studies (Farrington & West, 1985; Hagan, MacMillan, & Wheaton, 1996; Rutter, 1989; Sampson & Laub, 1990, 1993; Wheaton, 1990). Being held back in school also represents a life event that fits this definition of a turning point. Such an event alters a child’s social milieu and networks, expectations, and self-perceptions. Children experience diminished self-esteem, peer ridicule, labelling, and long-lasting perceptions of retention as punishment (Gottfredson, Fink, & Graham, 1994; Kaplan, Peck, & Kaplan, 1994). The response of parents and school authorities may exacerbate the child’s negative reactions. Byrnes (1989) finds that some parents punish a child for being retained. Also, labeling theory (Lemert, 1972; Rosenthal & Jacobson, 1968, 1992) suggests that such a visible demonstration of failure may alter perceptions and expectations about the child on the part of teachers and other school authorities.

Studies of the linkage of school failure and conduct problems most commonly measure school failure by grade point average or achievement tests, not by the event of being placed off grade level (Maguin & Loeber, 1996). Only a small number of studies have specifically focused on the effect of grade retention on antisocial behavior. Findings from

![Figure 1. The trajectories of physical aggression.](image-url)
Maguin, Loeber, and LeMahieu (1993); LeBlanc, Vallières, and McDuff (1992); and Pagani and her colleagues (Pagani, Boulerice, Vitaro, & Tremblay, 1999; Pagani, Tremblay, Vitaro, Boulerice, & McDuff, 2001) on the effects of being held back mirror the conclusion of the larger literature: students who have academic difficulties are more prone to antisocial behavior.

The study by Pagani et al. (2001) is particularly germane to this analysis. It used autoregressive modeling for the purpose of examining the effect of grade retention on academic and behavioral adjustment from ages 6 to 12 years old. The authors found that children with a history of behavior problems were more likely to be retained and that grade retention did not improve behavioral adjustment. To the contrary, maladjustment not only persisted but, on average, worsened after grade retention. Particularly relevant for our analysis was the finding that the maladjustment was more pronounced when grade retention occurred early in primary school and that boys were particularly vulnerable to large, long-lasting negative effects on academic performance and classroom disruptiveness. In this study we build upon Pagani et al. (2001) by investigating whether grade retention affects adjustment beyond primary school and also whether effects of retention depend on the child’s prior developmental course.

For several reasons we focus on physical aggression and on boys in particular. First, several leading explanations of the nexus of school failure and delinquency emphasize the frustration and disappointment engendered by school failure (Agnew, 2000; Cohen, 1955; Hirschi & Hindelang, 1977; Lemert, 1972; Moffitt, Gabrielli, Mednick, & Schulsinger, 1981; among others). Second, boys are more at risk of being held back than girls (Shepard, 1997), are retained earlier (Mâsse & Tremblay, 1999), and have more undercontrolled responses to being held back (Pagani et al., 2001). Third, compared to other forms of misbehavior, physical aggression against peers is far more likely to be visible and problematic to teachers. In this study, our measurements of school misconduct are based on teacher ratings from ages 6 to 15. Fourth, studies of toddlers show that most humans, especially males, begin acts of aggression as soon as they are physically able (Tremblay, 2000) and that violent behavior is present throughout the life span whereas other forms of misconduct such as stealing or drug use are more age dependent (Loeber & Stouthamer–Loeber, 1998). Fifth, and most importantly, violence is among the most socially damaging and feared forms of antisocial behavior. Not only does it cause great harm to victims but the perpetrators themselves also commonly suffer. Their violence heightens their own risk of being victims of violence, as well as militates against the acquisition of important life skills in both cognitive and social domains (Caspi, Wright, Moffitt, & Silva, 1998; Moffitt, 1993).

A Life Course Perspective on Grade Retention and Physical Aggression

The linkage of school failure and delinquency is long established (Gottfredson, 1981; Maguin & Loeber, 1996; Silberberg & Silberberg, 1971). Potential explanations for this inter-relationship can be found in social control (Hirschi, 1969; Hirschi & Hindelang, 1977), social learning (Akers, 1985), strain (Agnew, 2000; Aseltine, Gore, & Gordon, 2000), labeling (Lemert, 1972), and anomie (Merton, 1968) theories of delinquency. In this section, we examine implications of this literature for whether developmental history and timing of a turning point condition the influence of grade retention on physical aggression.

Among the most robust empirical regularities about aggression is its stability over time (Caspi & Moffitt, 1991; Loeber, 1982; Moffitt, 1993; Nagin & Paternoster, 1991; Olweus, 1979). Sampson and Laub (1990, p. 27) observe that this regularity allows “an impressive generalization that is rare in the social sciences.” In light of this persistence, does the developmental course of conduct problems moderate the impact of an event like grade retention? Both empirical evidence (Caspi & Moffitt, 1991; Loeber, 1982) and theory (Moffitt, 1993; Patterson, 1982; Patterson, DeBaryshe, & Ramsey, 1989; Patterson & Yoerger, 1993) strongly point to an affirmative response.
In her developmental taxonomy of antisocial behavior, Moffitt (1993) posits that the behavioral constancy of those showing life course persistent conduct problems is the outcome of an ongoing process in which environmental reaction to the individual’s poor impulse control reinforces the individual’s propensity toward antisocial behavior. Her theory predicts decreasing malleability with age. This suggests that youngsters engaged in a longstanding antisocial behavioral trajectory will be less affected by school failure and its concomitant, grade retention. An inversion of Moffit’s reasoning also suggests that individuals with a longstanding prosocial trajectory would be less susceptible to deflection toward conduct problems than individuals in a behavioral trajectory already including some degree of antisocial activity. The former group’s deeply ingrained prosocial disposition and strong adherence to conformity may serve to protect them from deflection toward an antisocial pathway.

Other theoretical perspectives, however, suggest different predictions. The theories of Agnew (1991, 2000) and George (1999) suggest that a trajectory of chronic conduct problems, particularly physical aggression, reflects a vulnerability to lashing out at stressful events. As such, grade retention may have a particularly aggravating effect on violent conduct problems in individuals with a long history of such problem behavior. This propensity for violence may be further aggravated by increased numbers of vulnerable targets due to the retained boy being physically larger, on average, that his younger classmates. Those with no history of physical aggression may be least affected by retention. In between would be individuals with more moderate trajectories of physical aggression.

These differing perspectives on how developmental history of physical aggression may condition the impact of grade retention share the assumption that grade retention will aggravate violence. However, the work of Wheaton and colleagues (Wheaton, 1990, 1996; Wheaton, Roszell, & Hall, 1997) suggests that if being held back in school results in the child being removed from stressful social interactions and circumstances, physical aggression among children with a history of violence in school could be reduced.

Thus, theory provides no clear guidance on how developmental history may affect the response to grade retention. Empirical resolution of this issue would not only further develop the life course analytical framework, but also provide valuable information for decision-makers on the effectiveness of retention as a corrective measure for academic failure.

Another central theme of life course studies involves the timing of life’s turning points. Elder (1998, p. 6) observes that “the personal impact of any change depends on where people are in their lives at the time of the change.” For key life events such as marriage or parenthood, age dependencies are obvious: marriage and parenthood at ages 16 versus 30 have manifestly different implications for the life course (Hagan & Palloni, 1990).

Empirical evidence on whether timing conditions the influence of grade retention is limited. Pagani et al. (2001) found that over the period from kindergarten to sixth grade, retention’s most persistent and enduring negative influence on classroom performance and disruptiveness was greatest in kindergarten and first grade. Whether this pattern of diminished impact with age extends beyond primary school is not clear.

Strain and social control theories of delinquency suggest that it might not. In strain theory (Agnew, 1991; Cloward & Ohlin, 1960; Cohen, 1955), delinquency is a response to the realization that chances of success through legitimate means are remote. Being held back in school is likely a more salient event in establishing this realization in older children. For them the prospect of entering the adult world is more proximate and tangible than for a grade school child. Similarly, social control theory suggests retention in adolescence may have a larger effect on problem behaviors than in earlier childhood. Adolescents typically have far larger investments in social bonds than younger children. They are also less easily and effectively supervised and monitored (Statin & Kerr, 2000). Thus, the fracturing of social bonds that attends grade retention may have a larger immediate impact on older children because they have lost more
social capital and they have more opportunities to engage in delinquency. Although empirical evidence points to diminishing impact through primary school, theory suggests the impact of retention on antisocial behavior may increase beyond primary school. Hence, our second objective is to test whether timing conditions the influence of grade retention beyond grade school.

In addition, we control for individual and early life circumstances that are associated with both physical aggression and poor school performance and which thereby bias our estimate of the impact of retention on physical aggression. Specifically, in this analysis we account for six such characteristics: maternal age and education, intelligence, inattentiveness, prosociality, and nonintact family status at age 6. Nagin and Tremblay (2001) found that each of these characteristics was predictive of membership in the trajectory groups shown in Figure 1. As will be shown, each is also correlated with grade retention, a finding that accords with much research documenting that school failure is associated with low parental education, family breakup, low IQ, and attention-deficit/hyperactivity disorder (Faraone, Biederman, & Kiely, 1996; Faraone, Biederman, Lehman, & Spencer, 1993; Fergusson, Lynskey, & Horwood, 1997; Pagani et al., 1999).

Method

Subjects

The subjects in this study were part of a longitudinal study that started in the spring of 1984, when all teachers of kindergarten classes in 53 French language elementary schools in the lowest socioeconomic areas in Montreal, Canada, were asked to rate the behavior of each boy in their classroom. The mean and median family income when the boys were age 10 (1988) was between $25,000 and $30,000 (Canadian dollars; $19,000–$23,000 in U.S. dollars) compared with a median income of $44,000 (Canadian dollars) for couples with children in Canada in 1987 (Mitchell, 1991). Eighty seven percent (87%) of the kindergarten teachers agreed to participate, and a total of 1,161 boys were rated. To control for cultural effects, boys were included in the longitudinal study only if both their biological parents were born in Canada and their parents’ mother tongue was French. We also eliminated those who refused to participate and those who could not be traced. These criteria resulted in a homogeneous Caucasian, French-speaking sample of boys (N = 1,037). Informed consent was regularly obtained from both mothers and sons throughout the study.

When they were in kindergarten, 67% of the boys lived with both of their parents, 24% lived only with their mother, 5% lived with their mother and stepfather, and the rest had other family arrangements. The mean age of the parents at birth of the child was 25.4 (SD = 4.8) for mothers and 28.4 for fathers (SD = 5.6). The mean age at the birth of their first child was 23.8 (SD = 4.1) for mothers and 26.4 (SD = 5.1) for fathers.

Measures

Parental and family characteristics. Interviewers with postsecondary education were trained to interview mothers by telephone after the initial teacher assessments in kindergarten. From these data, the following binary indicator variables were constructed: “teenage mother” identifies boys of mothers who began childbearing at age 18 or younger; “low maternal education” indicates boys whose mothers completed 9 years or less of schooling, placing these women in the lower quartile of the parental education distribution for the sampled boys. Nagin and Tremblay (2001) found that these maternal variables predicted trajectory group membership but that companion paternal variables did not predict group membership controlling for the maternal characteristics. “Not intact” identifies natural parents who separated or divorced prior to the boy’s sixth birthday.

Teacher ratings. At the end of the school year, physical aggression, inattention, and prosocial behaviors were rated by the teacher most knowledgeable about the child in kindergarten (age 6), and then from ages 10 to 15 using the Social Behavior Questionnaire (Tremblay et
The teachers were given a list of statements describing children's behavior and asked to rate how often (often = 2, sometimes = 1, never = 0) the child behaved according to the description. Physical aggression at ages 6 and 10–15 was assessed with three items: fights with other children; kicks, bites, hits other children; and bullies or intimidates other children (Cronbach's $\alpha = 0.78–0.87; M = 0.84$). Inattentiveness in kindergarten was assessed with two items: inattentive and has poor concentration (Cronbach's $\alpha = 0.74$). Prosocial behavior in kindergarten was assessed with 10 items: tries to stop quarrels; invites bystanders to join in; tries to help someone hurt; helps pick up things someone else dropped; praises work of less able children; shows sympathy toward someone who made a mistake; helps children having difficulty with a task; helps children who are sick; comforts crying or upset child; and helps clean up mess made by someone else (Cronbach's $\alpha = 0.92$). Following Nagin and Tremblay (2001), the inattention and prosociality scales were dichotomized to distinguish individuals in the upper quartile of the sample distribution of each scale.

Self-reported violent delinquency. Questions concerning the boys' violent delinquent behavior over the past 12 months from ages 12 and 15 were used to create a physically violent behavior scale that included fist fighting, gang fighting, carrying/using a deadly weapon, threatening or attacking someone and throwing an object at someone. The internal consistency index (Cronbach's alpha) for that scale ranged from 0.76 to 0.81 ($M = 0.79$).

Grade retention. Indicators of grade retention were extracted from official records on whether the boy was in an age appropriate classroom at the end of the school year. First-time grade retention, the focus of this analysis, was defined as the first occasion on which the boy was behind grade level. Thus, for example, if a boy was in less than grade 5 at the age 11 and in all previous years he had been at grade level, we infer that he had been retained for the time first after completing the grade 4 school year.

In this analysis we focus on first-time grade retention. We do not attempt to estimate the impact of any subsequent experience with retention. Also, we do not distinguish boys who were retained but kept in the regular school system from their counterparts who were transferred to a special school for children with behavioral problems. While there are good reasons for suspecting that these two groups might respond differently to grade retention, subsidiary analyses revealed no such difference. Thus, in the interest of parsimony we combine these two conceptually distinct groups.

**Intellectual assessment.** Verbal IQ was assessed at age 13 with the Sentence Completion Test (Lorge & Thorndike, 1950). IQ is generally relatively stable from kindergarten to adolescence (Sattler, 1988). As such, the age 13 assessment should give an estimate of the relative ranking of the boys on cognitive performance when they were in kindergarten. A correlation of .67 was found when the age 13 Sentence Completion Test score for a subsample of subjects ($N = 80$) was correlated with an IQ assessment at age 10 using the vocabulary and block design subtest of the WISC-R. “Low IQ” identifies boys in the lower quartile of the sample verbal IQ distribution.

**Data analysis procedure**

The present analysis demonstrates an extension of the group-based statistical model used to estimate the trajectories displayed in Figure 1. The extension provides the capacity for testing whether a turning point like grade retention alters the trajectory itself while controlling for enduring individual characteristics that may affect both the individual’s trajectory and the likelihood of experiencing the transition.

As a prelude to describing this extension, we first discuss the basic form of the statistical model used to estimate the trajectories in Figure 1. Like other approaches for statistically modeling developmental trajectories such as hierarchical and latent growth curve modeling, the semiparametric, group-based method uses a polynomial function to model
the relationship between an attribute (e.g., physical aggression) and age (Nagin, 1999; Nagin & Tremblay, 1999). The function takes the form,

$$y_{ij}^* = \beta_0 + \beta_1 \text{age}_i + \beta_2 \text{age}^2_i + \epsilon,$$

where $y_{ij}^*$ is a latent variable characterizing the level of the attribute for participant $i$ at time $t$ given membership in group $j$, $\text{age}_i$ is participant $i$’s age at time $t$, $\text{age}^2_i$ is the square of subject $i$’s age at time $t$, and $\epsilon$ is a disturbance assumed to be normally distributed with zero mean and constant variance $\sigma^2$. The model’s coefficients, $\beta_0$, $\beta_1$, and $\beta_2$, determine the shape of the trajectory and are subscripted by $j$ to denote that the coefficients are free to vary across groups.

This flexibility is a key feature of the model because it allows for easy identification of population heterogeneity not only in the level of behavior at a given age but also in its development over age. A single peaked trajectory is implied if $\beta_1 > 0$ and $\beta_2 < 0$. If behavior remains constant over age, this trajectory is implied if $\beta_1 = 0$ and $\beta_2 = 0$. Other interesting possibilities include trajectories in which growth is either steadily accelerating or decelerating. The former would be characterized by a trajectory in which both $\beta_1$ and $\beta_2$ are positive and the latter by both being negative. Thus, the method allows for the identification of unusual mixtures of trajectories within the population.

As discussed in Nagin (1999) and Raudenbush (2001), conventional hierarchical modeling (Bryk & Raudenbush, 1987, 1992; Goldstein, 1995) and latent curve analysis (McArdle & Epstein, 1987; Meredith & Tisak, 1990; Muthen, 1989; Willett & Sayer, 1994) methods are ill suited for identifying the sorts of unusual mixtures of trajectories that are commonly postulated in taxonomic theories of development. While these two classes of methods differ in very important respects, they also have important commonalities (MacCallum, Kim, Malarkey, & Kiecolt–Glaser, 1997; Raudenbush, 2001; Willett & Sayer, 1994). For our purposes, one is key: both model the unconditional and conditional population distribution of growth curves based on continuous distribution functions, which makes it very difficult to use them for testing theories positing distinctive developmental courses of psychopathology. Raudenbush (2001, p. 513) observes: “It makes no sense to assume that everyone is increasing (or decreasing) in depression... many persons will never be high in depression, others will always be high, while others will become increasingly depressed.”

Because conventional growth curve modeling methods were ill suited to identifying unusual mixtures of trajectories, developmental researchers have commonly resorted to using assignment rules based on subjective categorization criteria to construct categories of developmental trajectories. Although such assignment rules are generally reasonable, there are limitations and pitfalls attendant on their use. One is that the existence of the various developmental trajectories that underlie the taxonomic theory cannot be tested; they must be assumed a priori. A second related pitfall is the risk of simultaneously over- and underfitting the data, thus creating trajectory groups that reflect only random variation and failing to identify unusual, but still real, developmental patterns. Third, ex ante specified rules provide no basis for calibrating the precision of individual classifications to the various groups that comprise the taxonomy.

Trajectory parameters are estimated via maximum likelihood estimation. The essential function of maximum likelihood estimation is to choose estimates of the parameters for which the likelihood of observing the actual data, $Y$, is maximized. In this study, as in many developmental investigations that rely on psychometric scale data, the distribution of $y_{ij}$ is censored. That is, at any given age, a significant number of participants exhibit none of the problem behaviors measured by the scale, resulting in a cluster of data at the scale minimum. The likelihood function link-
As demonstrated in Roeder et al. (1998) and Nagin (1999) controls for the impact of each of these factors on trajectory \( \pi_j \) can be introduced as follows: let \( x_i \) denote a vector of factors measuring individual, familial, or environmental factors that are potentially associated with group membership and \( \pi_j(x_i) \) denote the probability of membership in group \( j \) given \( x_i \). For a two-group model, the logit model is a natural candidate for modeling group membership probability as a function of \( x_i \). For this special case, we need only estimate \( \pi_j(x_i) \) for one group, say group \( 1 \), because

\[
\pi_2(x_i) = 1 - \pi_1(x_i)
\]

where

\[
\pi_1(x_i) = \frac{e^{\theta_0}}{1 + e^{\theta_0}}.
\]

For the more general case, in which there are more than the two-groups, the logit model generalizes to the multinomial logit model (Maddala, 1983):

\[
\pi_j(x_i) = \frac{e^{\theta_j}}{\sum_j e^{\theta_j}},
\]

where the parameters of the multinomial logit model \( \theta_j \) capture the impact of the covariates of interest \( x_i \) on the probability of group membership. Without loss of generality, \( \theta \) for one “contrast” group can be set to zero. The coefficient estimates for the remaining groups should be interpreted as measuring the impact of covariates on group membership relative to the contrast group.

This brings us to the model extension required for the analyses reported here. To capture the influence of grade retention on each group’s trajectory, Equation 1 is expanded to include indicators of whether and when individuals first experienced grade retention:

\[
y_{ij}^* = \beta_0 + \beta_1 \text{age}_t + \beta_2 \text{age}^2_t + \alpha_1 \text{fail 6-10} + \alpha_2 \text{fail 11-12} + \alpha_3 \text{fail 13-15} + \epsilon,
\]
where fail6_10, fail 11_12, and fail 13_15 are binary variables that are specified to allow estimation of separate estimates of the impact of grade retention for three age intervals of approximately equal length: ages 6–10, 11–12, and 13–15. Specifically, fail6_10 distinguishes boys who first failed between ages 6 to 10. For these boys it equals 0 for all ages prior to the retention and equals 1 for all ages after the retention. For boys who never failed or failed after age 10 this variable equals 0 for all ages. Fail11_12 distinguished boys who first failed at ages 11 or 12. For these boys it equals 0 for all ages prior to the retention and equals 1 for all ages after the retention. For boys who never failed or failed outside the age 11–12 window this variable equals 0 for all ages. Fail13_15 distinguishes boys who first failed at ages 13–15. For these boys it equals 0 for all ages prior to the retention and equals 1 for all ages after the retention. For boys who never failed or failed outside the age 13–15 window this variable equals 0 for all ages.

To illustrate the specification of these indicator variables consider two hypothetical cases: a boy who is first retained at age 10 and a boy who is first retained at age 14. For the first example, fail6_10 equals to 0 at age 6 but from age 10 onward equals 1. Fail11_12 and fail13_15 equal zero throughout the measurement period from age 6 to age 15 because each is respectively reserved for capturing the ongoing effects of failure at ages 11 and 12 and ages 13 thru 15. For the second example fail13_15 equal zero until age 13 at which time it equals 1 for that and all subsequent ages. Fail6_10 and fail11_12 equal zero at all ages because the first retention occurs during the age 13 to 15 window. With this model specification three age-dependent estimates of the impact on physical aggression of being held back in school are generated. Thus, for a four group model, a total of 12 impact estimates are generated.

Figure 2 illustrates the output of the model for a hypothetical trajectory group. The baseline trajectory depicts the behavior in the absence of grade retention. Also, included in the figure are illustrative examples of the possible effects of failing at ages 9, 11, and 13. For this hypothetical trajectory group, problem behaviors increase if the transition occurs at age 9 but decrease if it occurs at age 11 and has no material impact if it occurs at age 13.

Figure 2 also serves to illustrate another aspect of model interpretation. If the trajectory groups were identifiable based on some measured characteristic of the individual, separate regression models of the form of Equation 2 could be estimated for each such group simply by sorting the data based on that characteristic. We make this point because the model’s estimated coefficients can be interpreted just like those of a conventional regression: they measure the change in the response variable associated with changes in the explanatory variables.

Figure 3 provides a schematic diagram of the total model structure. The top of the figure lists psychosocial characteristics of the individual measured at or prior to the outset of the measurement series. Grade retention is a potential contemporaneous influence on physical aggression that is measured by a time varying covariate. Its impact is calibrated by its influence on the trajectory itself.

Results

As prelude to describing the results concerning the impact of retention on physical aggression, we first report findings on the psychosocial predictors of trajectory group membership and grade retention. Table 1 reports summary statistics on distinguishing characteristics of trajectory group members. To construct this table, individuals were assigned to the trajectory groups shown in Figure 1 based on their posterior probabilities of group membership. These probabilities estimate the probability of the individual belonging to each trajectory group given his history of physical aggression. Individuals were assigned to the group with the largest posterior probability estimate.
For example, consider an individual who persistently received high physical aggression ratings by teachers. For this individual, the posterior probability estimate of his belonging to the low trajectory group would be near zero whereas the estimate of his belonging to the chronic group would be high.

The tabulations show large differences across groups with respect to both grade retention rates and the group membership predictors. By age 15 nearly all of the boys assigned to the chronic trajectory had been held back whereas the retention rate for the low group was far smaller, about 25%. The prevalence rates for each of the predictor variables also differed substantially between the chronic and low groups. The teenage and low maternal education, household not intact at age 6, low IQ, high prosociality, and high inattention variables all significantly distinguished trajectory group membership. Furthermore, except for the prosociality factor, their prevalence increased sequentially from the low group to the chronic group. For the prosociality variable, the pattern was reversed. All these patterns accord with expectations (Nagin & Tremblay, 2001). Also, it is important for the purposes of this paper that each of these psychosocial variables is significantly correlated with grade retention by age 15. For example, the propor-
Table 1. Parental and child characteristics by trajectory group

<table>
<thead>
<tr>
<th>Variables</th>
<th>Moderate Declining</th>
<th>High Declining</th>
<th>Chronic</th>
<th>χ² (df = 3)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>191</td>
<td>518</td>
<td>298</td>
<td>30</td>
<td>NA</td>
</tr>
<tr>
<td>Child variables (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below grade level by age 15</td>
<td>27.2</td>
<td>41.3</td>
<td>75.5</td>
<td>96.7</td>
<td>171.0</td>
</tr>
<tr>
<td>Low IQ</td>
<td>19.9</td>
<td>22.6</td>
<td>36.5</td>
<td>43.3</td>
<td>27.19</td>
</tr>
<tr>
<td>High prosocial</td>
<td>28.8</td>
<td>25.9</td>
<td>12.1</td>
<td>10.0</td>
<td>31.73</td>
</tr>
<tr>
<td>High inattention</td>
<td>7.9</td>
<td>16.0</td>
<td>32.6</td>
<td>40.0</td>
<td>59.29</td>
</tr>
<tr>
<td>Parent variables (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenage mother</td>
<td>12.6</td>
<td>20.7</td>
<td>28.9</td>
<td>53.3</td>
<td>33.28</td>
</tr>
<tr>
<td>Low maternal education</td>
<td>23.6</td>
<td>28.2</td>
<td>40.9</td>
<td>66.7</td>
<td>35.74</td>
</tr>
<tr>
<td>Family not intact at age 6</td>
<td>15.1</td>
<td>26.6</td>
<td>38.3</td>
<td>43.3</td>
<td>35.82</td>
</tr>
</tbody>
</table>

Note: The p values are on a joint test of significance.

3. Calculation of this odd ratio requires that the levels of the other predictor variables be specified. We set them equal to the sample average. Also, the odds ratio is a nonlinear function of the multinomial logit parameter estimates used in specifying the mixture distribution. Because conventional methods for computing confidence intervals do not apply, we used a parametric bootstrap to compute the confidence intervals (Efron, 1979).
physical aggression. As discussed below, this has important implications for remediation of the impact of grade retention on violence.

Table 3 reports parameter estimates and z scores for the 12 estimates of the influence of grade retention on classroom physical aggression. None of the impact estimates for the never and chronic trajectories were individually significant. For both groups we also tested the joint significance of the three separate age dependent estimates of the effect of grade retention. Specifically, for this test the null hypothesis is that all three grade-retention effects are zero (i.e., \( \alpha_1 = \alpha_2 = \alpha_3 = 0 \) in Equation 2) and the alternative hypothesis is that at least 1 is nonzero. Based on the likelihood ratio test, the null hypothesis can not be rejected for even a level of significance of .10.

A possible explanation for our failure to find statistically significant impact estimates for the chronic group is a lack of power. This group is small, estimated to comprise less than 5% of the population, and based on the posterior-probability-sorting rule, only 30 individuals were assigned to this group. We, thus, estimated a model in which we did not attempt to estimate three age-graded impacts of grade retention. Instead we estimated a single grade retention impact coefficient that applied to all ages. The result was the same: the impact estimate was statistically insignificant and small in magnitude. We also repeated this same exercise for the low group and still found no evidence of increased physical aggression. In contrast, highly significant impacts were found for the moderate and high declining trajectories. Combined, these two groups comprise about 80% of our sampled population. All of the point estimates were positive and five of six were highly significant. Table 3 also includes the predicted change in physical aggression. The calculations reported under this heading estimate the change in the measured physical aggression index that are implied by the coefficient estimates in the adjoining column. Because the model was constructed to account for censoring due to data clusters at the scale minimum and maximum, a technical adjustment is required to translate the coefficient values into an estimate of the expected change in the measured index of physical aggression. 4 As a benchmark for calibrating the

4. The expected latent value (\( y^* \)) is

<table>
<thead>
<tr>
<th>Variable</th>
<th>With School Failure in Model</th>
<th>Without School Failure in Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low maternal education</td>
<td>1.9</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>(1.1, 3.4)</td>
<td>(1.3, 4.5)</td>
</tr>
<tr>
<td>Teenage mother</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>(1.2, 6.7)</td>
<td>(1.3, 5.9)</td>
</tr>
<tr>
<td>Family not intact at age 6</td>
<td>2.5</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>(1.5, 4.2)</td>
<td>(1.7, 5.2)</td>
</tr>
<tr>
<td>Low IQ</td>
<td>1.9</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>(1.1, 3.7)</td>
<td>(1.4, 3.6)</td>
</tr>
<tr>
<td>High inattention</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>(1.6, 6.8)</td>
<td>(1.9, 6.8)</td>
</tr>
<tr>
<td>High prosocial</td>
<td>.7</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>(.3, 2.0)</td>
<td>(.2, 1.6)</td>
</tr>
</tbody>
</table>

Note: The values are the odds ratio of high versus low trajectory group membership by parental and child characteristics. Low aggression combines the low and moderate declining groups and high aggression combines the high delinquency groups.
size of these impact estimates, the average measured level of the physical aggression index across all subjects and years is about .8. For the moderate and high declining groups, the average is, respectively, .5 and 1.8. Compared to these sample averages the impact estimates for the five statistically significant impact estimates are quite sizable in magnitude.

Figures 4 and 5 provide graphic depictions of these retention impact estimates for the moderate and high declining trajectory groups. In both figures, four trajectories are shown. One is for a boy who experiences no grade retention. This trajectory is based solely on the quadratic relationship in age that defines the expected trajectory without retention. A second trajectory depicts the impact of retention at age 6. The estimated change in aggression is based on the age 6–10 impact estimate. The third and fourth trajectories depict the change in the baseline trajectories attending retention at ages 11 and 13, respectively. These changes were based on the ages 11–12 and ages 13–15 impact estimates, respectively. For the moderate declining group, Figure 4 shows that the size of grade retention impact is about the same for ages 6–10 and 11–12 but decreases for ages 13–15, whereas the counterpart trajectories for the high declining group shown in Figure 5 do not show any timing effects. The impact estimates by age range were virtually identical for this group which suggests that timing has no consequential influence on the deleterious impact of grade retention for this group. Tests of the equality of the magnitude of these impact estimates (i.e., null: $\alpha_j = \alpha_k = \alpha_l$, alternative at least one not equal to the other two), again based on the likelihood ratio, confirm this impression. The tests showed significant differences for the moderate declining trajectory group but not for the high declining group.

Our analysis is based on teacher reports of school-based violence. We also conducted a supplementary analysis to investigate whether our finding that grade retention exacerbated the physical aggression of the moderate and high declining trajectory groups extended to violence outside the school environment. This analysis was based on the self-reported violent delinquency of these two groups. Although some of the reported events may have pertained to violence in school, it is likely that the great proportion of the acts, which involve behaviors such as gang fighting, occurred outside of school.

The analysis examined whether boys who experienced first-time grade retention from age 13 to age 15 had a larger change in self-reported violence from ages 12 to 15 than that of boys who maintained their status as never having been retained. The analysis was conducted as follows: the model was reestimated without the predictor variables but with the grade retention variables. Individuals most likely belonging to the high and moderate declining groups were then identified based on their posterior probabilities of group membership. The model without the early psychosocial risk factors was employed for this purpose because we wanted group assignments based on actual behavior only. As discussed below, we otherwise controlled for these psychosocial variables. For each group, we examined only the subset of individuals who prior to age 13 had never been held back.

The analysis was designed to test whether the individuals who first failed from ages 13
Table 3. Impact of school failure on physical aggression by age and trajectory group

<table>
<thead>
<tr>
<th>Group</th>
<th>School Failure Age (years)</th>
<th>Coefficient (z Score)</th>
<th>Predicted Change in Physical Aggression</th>
<th>Signif. Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>6–10</td>
<td>8.77 (0.08)</td>
<td>.05</td>
<td>No a</td>
</tr>
<tr>
<td></td>
<td>11–12</td>
<td>9.23 (0.09)</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13–15</td>
<td>−0.64 (−0.00)</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>Moderate declining</td>
<td>6–10</td>
<td>1.24 (5.75)</td>
<td>.80</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>11–12</td>
<td>1.67 (6.47)</td>
<td>.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13–15</td>
<td>0.45 (1.07)</td>
<td>.08</td>
<td></td>
</tr>
<tr>
<td>High declining</td>
<td>6–10</td>
<td>1.2 (4.84)</td>
<td>.96</td>
<td>No c</td>
</tr>
<tr>
<td></td>
<td>11–12</td>
<td>1.21 (4.32)</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13–15</td>
<td>1.44 (2.34)</td>
<td>.98</td>
<td></td>
</tr>
<tr>
<td>Chronic</td>
<td>6–10</td>
<td>−0.10 (−0.17)</td>
<td>−.07</td>
<td>No b</td>
</tr>
<tr>
<td></td>
<td>11–12</td>
<td>0.72 (0.90)</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13–15</td>
<td>−1.59 (−1.13)</td>
<td>−1.24</td>
<td></td>
</tr>
</tbody>
</table>

aThe impact estimates by age differ significantly from each other.
bBut all are not significantly different than 0.
cAll are significantly different than 0.

Figure 4. The school failure impacts for a moderate declining trajectory.

To 15 had a larger change in their self-reported violence over this age interval than their counterparts who continued their nonretained status over this period. For the moderate declining group, the mean increase in the self-reported violence score for individuals who continued their nonretention status was .18, whereas the average increase for individuals who experienced first-time retention was 1.32. A z test of the difference in these two change scores shows a highly significant increase in self-reported violence for the firsttime retention group compared to their counterparts who maintained their nonretention status (z = 3.14). A counterpart test for the high declining group showed a similar pattern: a higher
Life course turning points

For the chronic group, this conclusion supports Moffitt’s (1993) suggestion that patterns of behavior in this group are so deeply ingrained that their behavior is difficult to influence, whether for good or bad. The findings for the low group may appear tautological, but they are not. If indeed grade retention increased the physical aggression of children with no history of such behavior, we would have expected to find evidence of such a jump, but we did not.

Although developmental history of violence clearly seems to condition the individual’s response to grade retention, the evidence on whether the timing of retention matters is not so clear. For the high declining group, the exacerbating effects of grade retention appear unrelated to timing, whereas for the moderate declining group, timing does seem to matter. Specifically, for the moderate declining group, we found evidence that grade retention aggra-vated classroom physical aggression, but only when it occurred prior to age 13. From ages 13 to 15 we found no evidence of an effect on violence in the classroom. However, we did find evidence that self-reported violence, presumably outside the classroom, did increase following retention after age 12.

We offer several related explanations for this seeming discrepancy. One is that by age 12, classroom physical aggression within the moderate declining group has declined to a negligible level. Like the never group, a pattern of nonviolent behavior within the school setting might be sufficiently ingrained by age 13 that grade retention no longer evokes a violent response in this context. Instead, the impact is manifested outside of school. A second explanation is that teachers are less acquainted with students after age 12 because of limited contact due to course scheduling. This makes them less reliable as informants. Third, as children grow older, antisocial activity may become more covert (Loeber, 1991) and, as a result, less detectable in the classroom. This underscores the importance of self-report measurement in the investigation of antisocial activities in adolescence.

However, this modest inconsistency in our results for the moderate declining group should not overshadow the finding that this propor-

Figure 5. The school failure impacts for a high declining trajectory.
A third conclusion implied by our findings is that there is heterogeneity in response to turning point events. This conclusion relates to a growing movement in the econometrics and statistics literature that interprets regression coefficient estimates as measuring the “treatment” effect of the covariate on the model’s response variable (Heckman, 1997; Manski, 1995; Rosenbaum, 1984; Rosenbaum & Rubin, 1983). By framing the inference problem in the language of experimental design, this terminology serves to emphasize that regression coefficients are measuring population averages. Unless the responses of all population members are homogenous, individual response will vary around this mean. Our finding that the treatment effect of grade retention depends on trajectory group and to a lesser extent on age is an empirical example of differential response across a population.

However, even within trajectory and age groups, there is likely to be variation in individual response. As reviewed in the opening sections, there are many theoretical accounts of the processes that might tie grade retention to either heightened or reduced physical aggression. A priori, there is no reason to expect that any one theory will explain the behavior of all people. We emphasize this point because much could be learned by closely studying the processes that underlay turning points for the purpose of better understanding how these processes interact with trajectory and age in modulating behavioral responses. For example, such research would be valuable in sorting out the contributions of labelling, social control, and social learning accounts of the linkage between grade retention and violence in a life course context.

Our study also had the methodological objective of demonstrating an approach for statistically embedding turning points into a life course trajectory. This demonstration was performed in the context of an important substantive and policy problem, the impact of grade retention on physical aggression. However, grade retention is only one example of a turning point event. It would be valuable to use this method to study other potentially important developmental turning point events. In this regard, we especially recommend studies
Life course turning points

of turning points for which there are good theoretical arguments for either positive or negative impact. An example from criminology is the effect of sending an adjudicated juvenile to a residential training facility (Manski & Nagin, 1998). We also recommend the study of events whose impact is likely to be heavily conditioned by age. A good example is the effect of childbearing onset on the well-being trajectories for both mother and child.

The key finding here is that grade retention influences the life course of individuals with specific risk characteristics. The question of when and under what circumstances a child should be held back in school remains a contentious issue. This study focuses on one relevant consideration in a complex predicament. Our finding that retention, on average, increases the physical aggression of two groups of children comprising 80% of the population is one important reason for caution about the efficacy of this educational intervention. Although the preponderance of evidence does not support the effectiveness of grade retention as it is typically implemented, it remains on the rise because many children are not mastering the standard criteria for adequate learning performance that allows progress to the next academic level (Mantzicopoulos & Morrison, 1992; Shepard, 1997). Although the effect on violence is only one factor in a panoply of competing considerations, at a minimum our results underscore the importance of providing programs to monitor and counsel children who have been held back in school.

There are also limitations to this study. Our analysis assumes that grade retention is exogenous controlling for trajectory group and early psychosocial characteristics. While in our judgement this assumption is reasonable, it would be very desirable to generalize the group-based trajectory methodology to allow for endogenous treatment assignment. Such an extension could build from a large econometric literature on this topic. In addition, because our analysis was restricted to boys of one ethnicity and nationality, replication for other ethnic groups in other countries is important to testing the generalizability of the findings. It would be especially pertinent to replicate this study with a sample of females. It is also important that we examine the effect of retention on developmental trajectories of academic achievement and other domains of performance and behavior. Only in this way will it be possible to assemble a complete picture of the developmental consequences of grade retention.

References


Nagin, D. (1999). Analyzing developmental trajectories...