A Dynamic Model of School Effects on Students' Academic Achievement*

Andrew Halpern-Manners
University of Minnesota

Abstract

Prior research on contextual effects nearly always measures "context" at a single point in time (e.g., children's school conditions in grade 10 or their family socioeconomic status at age 16). I argue that this measurement strategy fails to account for (1) age-based variation in children's sensitivity to their surroundings; (2) differential effects stemming from differences in the length of children's exposures; and (3) moves between contexts and changes over time within them. To evaluate the implications of this argument, I specify and test a more dynamic model of school effects on young people's academic performance. Drawing on nationally-representative longitudinal data and recent advances in growth mixture modeling, I identify a series of qualitatively distinct trajectories of school exposure that extend across a substantial portion of respondents' childhood and adolescent years. I then use these trajectories as predictors in models of math and reading achievement.

Despite relatively complex theories about the process of human development and the influence of childhood environments, researchers in the social sciences have tended to rely on crude measurement techniques to characterize

*Paper prepared for presentation at the annual meeting of the Population Association of America, San Francisco, CA, May 2012. Please do not cite or quote without permission. Correspondence can be directed to Andrew Halpern-Manners, Department of Sociology, University of Minnesota, 909 Social Sciences, 267 19th Avenue South, Minneapolis, MN 55455 or via email at andy@pop.umn.edu.

children's social surroundings. A good example is the work of Mayer (1991), who used data from a large-scale longitudinal survey to estimate the relationship between the socioeconomic and racial/ethnic composition of a school's student body and the likelihood that a young person drops out or gives birth to a child prior to graduation. Like any number of similar studies, the key school-level covariates in her analysis reflected the socio-demographic makeup of the schools students attended in the tenth grade, measures that she described as suitable proxies "for the mix of the schools the student attended in grades 1 through 9 as well as in grades 11 and 12" (Mayer 1991:326).

In this paper, I argue that this measurement strategy, and the thinking behind it, are misguided in at least two important respects. First, if the qualities of young people's environments vary in meaningful ways over the course of their early life (either because they move between contexts or because their social contexts change around them), then one-time assessments of those environments may distort their lived experiences, compromising the ability of researchers' to evaluate the importance of different contextual characteristics. Second, if the effects associated with childhood settings are conditional on the developmental stage of the individual or the overall duration of their exposure, the findings that emerge from traditional analyses may be incomplete (or worse, inaccurate), failing to reflect differential effects stemming from individuals' time-indexed circumstances.

I explore these possibilities through an examination of school effects on young people's academic performance. Using nationally representative longitudinal data and recent advances in finite mixture modeling, I identify a distinct set of school context trajectories that extend across a significant portion of individuals' childhood and adolescent years. As I elaborate below, these trajectories simultaneously incorporate information on the timing and duration of individuals' exposures, and can be used to predict important distal outcomes like student achievement. Together, these features allow me to (1) characterize students' full history of school exposure, including any changes over time in the characteristics of their surroundings; and (2) test whether the impact of an exposure depends on when it occurs and/or how long it lasts.

Data and methods

My analysis of school effects on young people's educational achievement—which I expect to complete by December of 2011—will be based on data drawn from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K). Administered by the National Center for Education Statistics (NCES), the ECLS-K is a randomly sampled, nationally representative study of U.S. children who were enrolled in kindergarten in the fall of 1998. Members of the study were assessed in reading, mathematics, and general knowledge skills at seven points in time between 1998 and 2007, making the ECLS-K an ideal data source for analyzing academic achievement. Of the students who participated in the base year, 7,803 remained in the panel in 2007; these cases will form the basis for my analytic sample.[1]

To each student's record, I will append annual school characteristics data acquired from the NCES's Common Core of Data (CCD) and/or Private School Universe Survey (PSS).[2] These matches can be performed using school identifiers available in the restricted-use ECLS-K file, which I obtained access to in December, 2010. Following prior research, the specific contextual

[1]Cases without full information on my outcome measures (n = 33) will be dropped from the analysis. Other missing data will be multiply imputed.

[2]Because the PSS is administered on a biennial basis, I plan to use a linear interpolant to fill in data for "off years."

variables that I intend to model are per pupil expenditures (expressed in constant 2007 dollars); pupil-teacher ratio; the percentage of the students who are from a minority group; the percentage of students who are from a racial/ethnic group other than the respondent's own; and the percentage of students receiving free or reduced-price lunch. With the exception of the expenditures variable (which is only available for school districts), these indicators are all measured annually at the school level.

Research design

After performing the necessary matches, my analysis will proceed in two stages. In the first stage, I will use maximum likelihood latent class growth models to summarize respondents' full histories of school exposures. A specialized application of finite mixture models, this method allows researchers to distill the tremendous complexity of longitudinal data into a manageable number of trajectory groups, or "latent longitudinal strata" (Haviland and Nagin 2005). Each strata is characterized by a unique patterning of exposure. As a result, it is possible to model situations in which the level of exposure to a particular contextual resource (e.g., small classes or high per pupil expenditures) is stable for some individuals, but increasing, decreasing, or changing in other complicated ways over time for others.

Once each of the latent class growth models has been fit, I will classify children into the trajectory group that most closely approximates their actual history of school exposure, both in terms of timing and duration. These classifications will be made on the basis of individuals' posterior probabilities of group membership, which measure the likelihood that a respondent with a specific sequence of measurements belongs to a given trajectory group (Bauer and Curran 2003; Nylund et al. 2007; Jones and Nagin 2007). In addition to providing an objective rule for making trajectory group assignments, these probabilities can also be used to judge the overall precision of the model. Following convention, I will require that the mean assignment probability within each group be 0.7 or greater (Nagin 2005).

In the second stage of my analysis, I will use a series of multivariate regression models to estimate associations between individuals' trajectory group assignments and two commonly studied indicators of educational success: reading and math achievement in the 8th grade. This exercise will be instructive in two respects. First, by comparing the results obtained from these models to estimates derived using more traditional point-in-time measures of school context, I will identify instances in which the trajectory-based approach conveys information about school effects that would otherwise be missed. Second, by comparing the coefficients associated with different contextual trajectories, I will be able to evaluate whether the educational effects attributable to school exposures are in fact conditional on their temporal properties.

Timeline to completion

I am currently in the process of fitting school characteristic trajectories using the ECLS-K data. I expect to complete this stage of the analysis by early-October, and will begin to write-up the results shortly thereafter. A full draft of the paper should be ready for circulation by mid-December.

References

Bauer, Daniel J. and Patrick J. Curran. 2003. "Overextraction of Latent Trajectory Classes: Much Ado About Nothing?" *Psychological Methods* 8:384–393.

Haviland, Amelia M. and Daniel S. Nagin. 2005. "Causal Inferences with Group Based Trajectory Models." *Psychometrika* 70:557–578.

Jones, Bobby L. and Daniel S. Nagin. 2007. "Advances in Group-Based Trajectory Modeling and a SAS Procedure for Estimating Them." *Sociological Methods and Research* 35:542–571.

Mayer, Susan E. 1991. "How Much Does a High School's Racial and Socioeconomic Mix Affect Graduation and Teenage Fertility Rates?" In *The Urban Underclass*, edited by Christopher Jencks and Paul Peterson, pp. 321–341. Washington, D.C.: Brookings Institution.

Nagin, Daniel S. 2005. *Group-Based Modeling of Development*. Cambridge: Harvard University Press.

Nylund, Karen L., Tihomir Asparaouhov, and Bengt Muthén. 2007. "Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study." *Structural Equation Modeling* 14:535–569.