Peer Contextual Effects or Selection?
Exposure to High Poverty Classrooms and Test Score Growth

Douglas Lee Lauen
Assistant Professor
Carolina Institute of Public Policy
Department of Public Policy
University of North Carolina at Chapel Hill

S. Michael Gaddis
Department of Sociology
University of North Carolina at Chapel Hill

August 2011 Draft
Do not cite without author's permission

Keywords: contextual effects, peer effects, academic achievement, growth model, poverty

Corresponding Author:
Douglas Lee Lauen
Department of Public Policy
University of North Carolina at Chapel Hill
Abernethy Hall, CB #3435
Chapel Hill, NC 27599-3435
919-843-5010
dlauen@unc.edu

Acknowledgements: The authors gratefully acknowledge the support of the Spencer Foundation and the North Carolina Education Research Data Center for providing the data. We thank Kyle Crowder, Patrick Curran, Mike Foster, Eric Grodsky, Guang Guo, Ashu Handa, Roz Mickelson, Mike Shanahan, Chris Wiesen, and especially Stephen R. Cole for their assistance and many useful comments.

Scholars have spent decades researching and debating the influence of school and neighborhood context on academic achievement, aspirations and attitudes (Alexander and Eckland 1975; Crosnoe 2009; Felmlee and Eder 1983; Rumberger and Willms 1992; Wilson 1959). The scholarly consensus is that high SES schools and neighborhoods positively affect individual academic outcomes (Brooks-Gunn, Duncan, Klebanov, and Sealand 1993; Entwisle, Alexander, and Olson 1994; Willms 1986), whereas high poverty schools and neighborhoods negatively affect academic outcomes (Crane 1991; Harding 2003; South, Baumer, and Lutz 2003). For example, Coleman and colleagues, in their seminal Equality of Educational Opportunity report, argued that peer effects were strong predictors of academic achievement: "the social composition of the student body is more highly related to achievement, independent of the student's own social background, than is any other school factor" (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York 1966: 325). Social science evidence on contextual effects has informed social science theory as well as educational policy in the United States, which for the past four decades has sought to mix students by racial background and, more recently, by poverty status (Bazelon 2008; Grant 2009; Kahlenberg 2001). The relevance of contextual effects research is demonstrated by the prominent role such research played in the recent social science statement submitted as an amicus curiae brief in a Supreme Court case about school desegregation.[1]

Despite the scholarly consensus, very little contextual effects evidence shows that *changes* in context affects *changes* in outcomes rather than simply the *level* of outcomes at one point in time. Therefore, scholars often attempt to make causal claims with correlational

---

[1] Statement of American Social Scientists of Research on School Desegregation to the U.S. Supreme Court in "Parents v. Seattle School District" and "Meredith v. Jefferson County" Urban Review. V. 40. No 1.

evidence. Selection bias can give rise to what Hauser (1970) termed the “contextual fallacy”: “…the contextual method rests on the arbitrary identification of residual group differences in the dependent variable with correlated aspects of group composition on an independent variable…The only way to eliminate such correlations is to assign individuals randomly to groups, and this is impossible with observational data” (p. 660). Recent work in sociology (Crosnoe 2009; Harding 2003) and in economics (Hanushek, Kain, and Rivkin 2009; Hoxby and Weingarth 2005; Solon, Page, and Duncan 2000) attempts to reduce bias in contextual effects through propensity score matching and weighting, comparison of sibling and neighbor correlations, fixed effects, instrumental variables, and natural experiments. Experimental evidence on the effect of changes in school and neighborhood context and academic achievement has emerged from the Moving to Opportunity program (Kling, Liebman, and Katz 2007; Orr, Feins, Jacob, Beercroft, Sanbonmatsu, Katz, Liebman, and Kling 2003; Sanbonmatsu, Kling, Duncan, and Brooks-Gunn 2006). Some of this recent work raises important questions about whether causal inferences about contextual effects are warranted (Mouw 2006). Finally, of the few longitudinal studies of the effects of contextual effects, most do not account for time-dependent confounding, which can arise when treatment exposure and confounders are measured on a number of occasions. Time-dependent confounders, which predict both future treatment and future outcome, conditional on past treatment, present a challenge to estimating unbiased treatment effects. For example, in estimating the effect of poverty context on child outcomes, one may wish to control for intermediate outcomes such as educational experiences while in school (such as assignment to gifted and remedial programs or being retained in grade). If these intermediate outcomes then predict both future treatment and outcome, standard methods – controlling for these factors, omitting them, controlling for baseline values – can produce biased estimates (Hong and Raudenbush 2008; Robins, Hernan, and Brumback 2000). Methods for addressing treatment effect bias from time-dependent confounding have been developed in epidemiology by Robins and colleagues (Cole and Hernán 2008; Hernan, Brumback, and Robins 2000; Robins 1999; Robins, Hernan, and Brumback 2000). Recent work using these methods has demonstrated negative effects of exposure to neighborhood concentrated disadvantage on verbal ability (Sampson, Sharkey, and Raudenbush 2008).

This study uses panel data to estimate the effect of classroom peer poverty rates on elementary and middle school students' test score status and growth. This panel includes interval metric and vertically equated mathematics test scores and variation across time in classroom-level peer poverty rates from a complete cohort of public school children in grades three through eight in the state of North Carolina from 2001 to 2006. The study contributes to contextual effects research by carefully specifying and accounting for bias from omitted and mismeasured time-invariant student and family background characteristics. We report effects of peer poverty based on two measures: attending a high poverty classroom (i.e., one in the top quartile the peer poverty distribution) and cumulative exposure to a high poverty classroom. We first present cross-sectional multilevel estimates of the association between peer poverty and math test score. These estimates reproduce the negative effects reported in previous research with cross-sectional designs. The strength of the cross-sectional association increases with grade level. By eighth grade, these estimates are particularly large, which is suggestive evidence that growth trajectories could widen over time. We then present growth model estimates, which show virtually no effect of exposure to peer poverty on children's test score growth trajectories. To address endogenous self-selection based on fixed unobservables, we present student fixed effects estimates, which remove between-student confounding (Allison 2009). This approach controls for time-invariant unmeasured and mismeasured aspects of student and family background that may predict both family choice of neighborhood and school and test score achievement. We then estimate marginal structural models with inverse probability of treatment weighting to address time-dependent confounding (Hong and Raudenbush 2008; Robins, Hernan, and Brumback 2000). Both of these alternative specifications produce predicted test score growth trajectories virtually identical to the growth model trajectories, which suggests that our estimates are robust to two different threats to validity. The null effect reported by this study does not mean that all children's life course outcomes are insensitive to peer poverty, but it raises important doubts about the causal status of the effect of peer poverty on student test score growth among children and young adolescents, an implication which we discuss in our conclusion.

THEORY AND EVIDENCE ABOUT CONTEXTUAL EFFECTS

Drawing upon the theory and evidence from the contextual effects literature on school and neighborhood effects, we suggest four potential theoretical explanations specific to the effect of peer poverty on student achievement growth for children and young adolescents. First, peer poverty may have a *negative effect* on student achievement growth due to *institutional* mechanisms: low parental involvement in schooling, lower quality teachers, teachers with lower expectations, and less rigorous curriculum. Second, peer poverty may have a *negative* effect due to *contagion* mechanisms: the downward leveling norms of predominantly low achieving peers. Third, peer poverty may have a *positive effect* due to *relative deprivation* mechanisms: the lack of competitive pressure and a lower average comparison group. Fourth, peer poverty may have *no effect* on student achievement growth once student background is properly controlled, which could point to a *selection* mechanism, i.e., that the apparent effect of context is due to the selection of families into schools and classrooms based on factors that are also correlated with test score growth and peer poverty level.

In their seminal review of school and neighborhood effects, Jencks and Mayer (1990) posited three hypotheses of the effect of affluent peers on youth outcomes: affluent peers are an advantage, affluent peers are a disadvantage, and affluent peers have no effects on student outcomes. The advantages of affluent peers flow from three sources: 1) a contagion (or epidemic) model in which positive values, attitudes, and behaviors are spread from affluent peers through peer group norms, 2) a collective socialization model in which adults other than family members serve as positive role models, and 3) an institutional model that stresses the additional resources flowing into affluent neighborhoods, such as higher levels of parent involvement and better teaching (Harris 2010; Jencks and Mayer 1990; Willms 2010). Research also points to the inverse of the affluence effect: the epidemic or contagion model predicts a negative effect on aspirations and achievement because high-poverty peers "infect" other students with negative school values, attitudes, and behaviors, and students in high poverty schools also have reduced access to beneficial resources, such as high-quality teachers (Crane 1991; Harding 2003; Harris 2010; Jencks and Mayer 1990; South, Baumer, and Lutz 2003). Among these three theories, education scholars most consistently advance institutional explanations of contextual effects. For example, in high poverty classrooms, teachers may lower their expectations, reduce course requirements, and adjust their instruction to a lower academic level, thus depressing achievement of students in such classrooms (Barr and Dreeben 1983; Lee, Bryk, and Smith 1993; Sedlak, Wheeler, Pullin, and Cusick 1986).

Cross-sectional contextual effects research finds a positive association between socially desirable youth outcomes and average school and neighborhood socioeconomic status (SES). For example, studies find positive effects of school mean parental education on standardized test scores (Entwisle, Alexander, and Olson 1994) and 4-year college enrollment (Choi, Raley, Muller, and Riegle-Crumb 2008), positive effects of school mean SES on grades and attainment (Willms 1986), and negative effects of the school mean poverty rate on academic self-esteem, educational aspirations and expectations, and standardized test scores (Battistich, Solomon, Kim, Watson, and Schaps 1995). Neighborhood effects research finds positive effects of high poverty neighborhoods on teenage pregnancy and high school drop-out rates (Crane 1991; Harding 2003), negative effects of early childhood neighborhood poverty on educational attainment measured in adulthood (Entwisle, Alexander, and Olson 2005), and negative effects of neighborhood deprivation on educational attainment in Scotland (Garner and Raudenbush 1991). Similarly, low levels of neighborhood poverty have been associated with positive effects on educational attainment (Duncan 1994), positive effects on standardized test scores (Entwisle, Alexander, and Olson 1994), positive effects on IQ, and negative effects on high school dropout rates (Brooks-Gunn, Duncan, Klebanov, and Sealand 1993).

There is also evidence to support the hypothesis that affluent peers and neighbors can have *negative* effects on youth outcomes. Scholars posit that relative deprivation, sometimes referred to as the "frog pond effect," discourages and depresses the aspirations, achievement, and attainment of students in more affluent schools (Attewell 2001; Bachman and O'Malley 1986; Crosnoe 2009; Davis 1966; Jencks and Mayer 1990; Marsh 1987; Marsh and Parker 1984). Though it may be advantageous to associate with affluent neighbors and peers, high achieving peers may harm aspirations, grades, curricular placement, and other academic outcomes, especially if students must compete for scarce resources. For example, Davis (1966) investigated whether the theory of relative deprivation explained college student career and graduate school application decisions. His results indicate that school mean achievement may have a negative effect on career aspirations, suggesting that students in more competitive environments may remove themselves from contention for high status careers and graduate schools. A more recent study finds that students in elite public high schools suffer a competitive disadvantage in entering elite colleges due to the importance of class rank in the college admissions process (Attewell 2001). This disadvantage may produce an organizational adaptation to triage resources in favor of the top students. Therefore, students in high, but not the highest quantiles of class rank, may receive worse grades and take less advanced courses than they would if they had attended a less elite public high school (ibid). Other studies with findings consistent with the relative deprivation hypothesis include Meyer (1970), Alexander & Eckland (1975), Alwin & Otto (1977), and multiple papers by Marsh and colleagues (Marsh 1987; Marsh and Hau 2003; Marsh, Kong, and Hau 2000; Marsh and Parker 1984).

On the other hand, peers may have little or no influence on individual outcomes. Youth may associate with peers like themselves no matter what the social context of higher level grouping units such as classrooms or schools (Jencks and Mayer 1990; Mouw 2006). School policy, in fact, may facilitate social homophily through tracking (Gamoran 1987; Harris 2010; Kubitschek and Hallinan 1998; Lucas 1999). For example, poor families who move to more affluent neighborhoods may experience no change in access to affluent peers if schools place poor students in classes with predominantly poor peers. Another study of high school students finds very minimal and mostly non-significant effects of school mean SES on a variety of test score outcomes (Gamoran 1987). In this case, the author controls for types of coursework and tracking variables and concludes that within-school differences in opportunities are more important than, and perhaps explanations for, school effects. It may also be the case that institutional, contagion, and relative deprivation mechanisms operate simultaneously, pulling student outcomes in opposite directions. Some contextual effects research has accounted for this possibility by including both school mean SES and school average ability in the same models (Alwin and Otto 1977; Meyer 1970), finding that these effects work in opposite directions to create a null combined effect.

Finally, contextual effects of peer poverty and affluence may simply reflect self-selection or omitted variable bias (Evans, Oates, and Schwab 1992; Hauser 1970; Leventhal and Brooks-Gunn 2000; Quigley and Raphael 2008). Important omitted and mismeasured family and student background characteristics may be causal determinants of test score achievement and how individuals sort into neighborhoods and schools. Controlling for these factors may greatly reduce the unadjusted difference in outcomes between students from high and low poverty contexts. For instance, Alexander and colleagues investigate the nature of school effects and find that controlling for individual SES reduces the effect of school mean SES on college plans to near zero (Alexander, Fennessey, McDill, and D'Amico 1979). Their conclusion is that “the school SES influences are shown to result to a considerable degree simply from SES differences in the kinds of students attending various schools” (235).

Longitudinal Research Designs of Contextual Effects

Much of the research discussed thus far employs cross-sectional designs, which ignore the cumulative nature of students’ educational development. A point-in-time study captures both the effect of prior educational experiences and student and family background and the effect of schooling in a focal year. Reviews of the literature note the importance of controlling for exogenous factors (i.e., those that do not depend on type of neighborhood/school) and call for more longitudinal designs (Duncan and Raudenbush 1999; Galster, Marcotte, Mandell, Wolman, and Augustine 2007; Harris 2010; Jencks and Mayer 1990; Saporito and Sohoni 2007). Two recent studies examine the effect of school SES composition on test score gain in high school using NELS, a nationally representative database. Using a multi-level growth model, Rumberger and Palardy (2005) find that the predictive power of school SES on student test score growth is as strong as family SES, and has particularly strong effects on science test score growth. The authors also find that the effect of school SES is explained by teacher expectations, the amount of homework students do, course taking, and student perceptions of school safety. Palardy (2008) uses a multilevel latent growth curve model and finds that, adjusting for student background characteristics, students in high SES schools enter high school with test scores 3.3 grade levels higher and adjusted learning growth rates 1.1 grade levels higher than students in low SES schools (p. 37). These studies represent a significant advance over prior research. They treat test score growth rather than levels as the outcome of interest. Although both studies use an impressive array of control variables to adjust for *observable* differences in student populations that could confound the school SES effect, the designs of these studies do not permit ruling out bias from the sorting of students into schools based on *unobservables*. They also fail to account for the problem of time-dependent confounding, which could arise if student's school SES is a function of lagged values of school SES and lagged values of the outcome. To our knowledge only one study, from the neighborhood effects literature, addresses the effect of poverty context on children's cognitive outcomes. Sampson, Sharkey, and Raudenbush (2008) examine the effect of changes in neighborhood concentrated disadvantage on children's verbal ability across three waves of African American families in the Project on Human Development in Chicago Neighborhoods study. To address the problem of time-varying confounding, this study estimates a marginal structural model with inverse probability of treatment weighting and reports that the effect of neighborhood concentrated disadvantage on children's verbal ability is large and negative, equivalent to missing a year of school (ibid).

Neighborhood Relocation Experiments

Although no experiment instituted to date allows for direct examination of school contextual effects, evaluations of two housing relocation programs, Gautreaux and Moving to Opportunity, provide further evidence about the impact of changes in context. Results from the Gautreaux program indicated positive effects for the families that moved to the suburbs compared to this comparison group including better academic outcomes for children and better employment outcomes for mothers (Rubinowitz and Rosenbam 2000). Although the results from the Gatreaux program support the hypothesis that reducing neighborhood poverty has positive effects, "critics were not mollified because assignment to comparison groups was nonrandom" (Clampet-Lundquist and Massey 2008:110). The Moving to Opportunity (MTO) experiment, improved upon the design of the Gatreaux program by randomly assigning participants to three groups: a control group that was not offered a voucher to move, a treatment group that was provided a Section 8 voucher and allowed to move without restrictions, and another treatment group that was provided a rental assistance voucher but allowed to move only to a census tract with less than 10% poverty. Early results indicated a number of benefits for the treatment groups including positive academic, behavioral, and health outcomes (see DeLuca and Dayton 2009 for a review of this research). However, later follow-up studies found that some of these positive results dissipated; in some dimensions individuals in the control group fared better than those in treatment (Kling, Ludwig, and Katz 2005). Children in the treatment group showed no real academic improvement and were in only marginally better schools than before the switch (Orr et al. 2003; Sanbonmatsu, Kling, Duncan, and Brooks-Gunn 2006). Recent debates on

MTO's applicability to the discussion of generalized neighborhood effects and problems of lingering selection bias raise further questions about measuring and determining contextual effects (Clampet-Lundquist and Massey 2008; Ludwig, Liebman, Kling, Duncan, Katz, Kessler, and Sanbonmatsu 2008; Sampson 2008). In particular, at least one study has discussed the potential problem of self-selection within the experiment because a sizable percentage of the families assigned to both treatment groups did not use their vouchers and these families were significantly different in a number of important dimensions from the families who did use their vouchers (Feins and Shroder 2005). Although families were randomly assigned to these three groups, those in the treatment groups were not required to move. Additionally, as Sampson (2008) points out, MTO is not a general test of neighborhood effects but rather an examination of the effect of changing neighborhood contexts in one direction on one dimension for a limited portion of the population who are extremely poor and overwhelmingly minority.

In summary, existing research on school contextual effects rests primarily on a base of cross-sectional designs of correlational evidence. Two studies of school contextual effects employ longitudinal designs, but ignore the problem of unobserved heterogeneity and time-dependent confounding. One study from the neighborhood effects literature uses appropriate techniques to address time-dependent confounding and reports a negative effect of neighborhood poverty on children’s verbal ability. Housing relocation studies provide evidence about changes in neighborhood, which also involve changes in school context, but suffer from limitations of generalizability to non-poor and non-minority populations and leave unexamined the effects of increases in peer poverty.

The present study makes a number of contributions to existing research. First, we use a longitudinal design to estimate growth curve trajectories over six years (grades three through eight) that relate the effect of changes in peer poverty to changes in students' math achievement. Second, we address selection bias by including student fixed effects. A large literature in economics and a growing literature in sociology (e.g. England, Allison, and Wu 2007; Jacobs and Carmichael 2001; Jacobs and Tope 2007; Kocak and Carroll 2008; Mouw 2003; Schneiberg, King, and Smith 2008) uses fixed effects methods to control for time-invariant unobserved heterogeneity.[2] These models, which require treatment variation within units over time, remove confounding bias that can emerge from omitted observable, mismeasured, or unobservable time-invariant student or group characteristics (Allison 2009; Halaby 2004; Mouw 2006; Wooldridge 2003). In the present context, this technique accounts for important student-level confounders such as low birth weight, early childhood education, and genetic factors, as well as family-level confounders such as parental IQ and class background. Third, following Sampson, Sharkey, and Raudenbush (2008), we account for time-dependent confounding by presenting growth curve trajectories produced from marginal structural models with inverse probability of treatment weighting. Fourth, unlike most prior school contextual research, we measure peer poverty at the classroom level rather than the school level, which, due to the non-random sorting of students to classrooms and middle school tracking based on achievement level, may produce less valid estimates of peer poverty effects. Fifth, we measure peer poverty two ways: attending a high poverty classroom (i.e., in the top quartile of the peer poverty distribution), and cumulative exposure to a high poverty classroom, which more accurately reflects the time-varying exposure to context over a youth’s life course. Sixth, we examine the effects of both increases and decreases in peer poverty among a diverse population of students enrolled in the North Carolina

[2] Here fixed effects refers to the panel data technique of using differencing or including indicator variables to control for unit-specific effects, not fixing random effects to zero or fixed (versus random) coefficients in random effects model.

public school system (a population that includes in large numbers whites, blacks, Hispanics, non-poor and poor students).[3] We focus on elementary and middle school aged student test score growth for two reasons: 1) the effects of peer poverty on younger students is relatively understudied, and 2) the effect of peer poverty has been shown to be stronger for cognitive and achievement outcomes than for behavioral and health outcomes (Duncan and Brooks-Gunn 1997).

DATA

This project uses test score and related data for one cohort of public school students in North Carolina beginning in grade three in 2001 through grade eight in 2006. North Carolina is a particularly appropriate state for this analysis because its tests are vertically equated, interval scaled, and consistently administered over this time period. The scores are produced from a three-parameter logistic IRT model and are scored on a developmental scale to allow computation of growth across grade levels.[4] The sample includes more than 500,000 student-year observations, beginning with about 100,000 unique third graders in 2001. By 2006, we observe about 75% of the original sample as being enrolled in a public school in North Carolina.[5] We focus on math rather than reading because educational research generally finds larger school and

---

[3] Relative to national statistics, blacks are overrepresented in North Carolina public elementary and secondary schools (31.5% vs. 17.2% nationally) and Hispanics are underrepresented (8.4% vs. 19.8% nationally). The percentage of whites in the North Carolina system closely mirrors the national percentage (56.6% vs. 57.1%). Data from the 2007 Digest of Education Statistics, table 40.

[4] The test is anchored at grade three and re-centered on grade five. The math test was rescaled in 2006. To compute growth scores for the state's accountability system the North Carolina Department of Public Instruction (NCDPI) conducted equating studies to permit conversion of scores across time. These studies, which used equipercentile equating, produced concordance tables to convert old scores to the new metric for the purposes of the state's accountability calculations. This study used these concordance tables to convert scores to a consistent metric.

[5] We retain all students in the cohort regardless of grade retention or promotion status. Students become censored from the sample due to leaving the public school system for in-state private schools and schools in different states. Due to the age of the sample (third through eighth graders), we suspect that very few are school drop outs, but we have no way of verifying drop out status with the data available for this study.

program effects on math scores than on reading scores (Murnane 1975).[6] Math scores for students in grades three through eight range from 303 to 388, with an average of 350.8 and a standard deviation of 11.9 (table 1). The average within-grade standard deviation of math scores is 9.2 (not shown). By the end of third grade, the average student math score is 339; by the end of 8th grade it is 360, suggesting a linear growth rate of about 4.2. This average masks the relatively large increases in the elementary grades (6-7 points per grade), relatively small increases in middle school grades (3 points per grade), and the stagnation of growth rates between 5th and 6th grade (2 points), a time which for most students corresponds to the elementary to middle school transition. Because initial status in this study is grade three, we subtract three from grade level to produce a regression intercept corresponding to average third grade achievement (i.e., initial status). To define high poverty classroom, we begin by standardizing the mean level of a student's classroom peers' free/reduced lunch status[7] by grade. To be consistent with prior research (Sampson, Sharkey, and Raudenbush 2008) we dichotomize this variable into a variable coded 1 if a student is in the top quartile of peer poverty and 0 if a student is in the bottom three quartiles of peer poverty.[8] Classroom is defined as the group of students with whom the student took their math test in each year.[9] We also derive an alternate measure of peer poverty designed to better capture cumulative effects, which we call *cumulative exposure* to a high poverty classroom. This time-varying variable measures the proportion of

[6] The choice to present effects on math rather than reading does not affect our conclusions. See appendix table A4 and figure A1.

[7] Unfortunately, a broader measure of socio-economic status is not available in the administrative data used for this study. Parental education, based on student self-report, is available. Free/reduced lunch status has the advantage of requiring parent income verification. Using parental education instead of free/reduced lunch status does not, however, change our conclusions (see appendix table A3).

[8] Using an absolute definition of peer poverty, coded 1 if 75% or more of a student's classroom peers are classified as free or reduced lunch and 0 otherwise, rather than a relative one, does not affect our conclusions about the effect of peer poverty exposure on test score growth (see appendix table A2).

[9] Classrooms with five or fewer students (less than two percent of the student-year observations) were dropped from the analysis.

years up to and including the current year a student has attended a high poverty classroom. Thus, a student can be coded only 0 or 1 during third grade but can be coded 0, 0.5, or 1 during fourth grade. An eighth grader coded as 0 was never exposed to a high poverty classroom, while one coded as 1 was always exposed to high poverty classrooms. Descriptive statistics in table 1 indicate that, on average, students in our cohort spend 24% of their third through eighth grade years in high poverty classrooms. 51% of 8th graders were never exposed. 5% of 8th graders were always exposed. Operationalizing peer poverty as a cumulative exposure measure provides a stronger test of contextual effects since we would expect the effect of the contrast between never exposed and always exposed to be larger than the contrast between high and low poverty classroom. For ease of exposition, below we will refer generically to the construct of peer poverty to encompass both high poverty classroom and cumulative exposure to high poverty classroom, distinguishing between the two when needed.

Peer poverty is time-varying rather than a fixed because 1) students can be assigned to classrooms with varying poverty composition over time, 2) students change schools due to residential changes and school choice, and 3) students make structural school enrollment changes (i.e., those arising from policy-induced school mobility due to how grade configurations are structured, chiefly changing from an elementary to a middle school, rather than family choices). Measuring peer poverty at the classroom level rather than the school level permits within-school variation in peer poverty to contribute to estimates. There is considerable variation in peer poverty both within and between schools. School average peer poverty rates range from 0% to 100%, with an average of 50% and a standard deviation of 24%. About 75% of total variation in peer poverty rates lies between elementary schools, while 25% of variation is between classrooms within schools. Perhaps due to early tracking, the portion of variation that lies between classrooms in middle schools is larger, at 40%, leaving 60% between schools.

We control for race/ethnicity, gender, ever retained (all time invariant), and parental education, status as gifted, special education, limited English proficient, school transitions, and poverty status (all time varying). We include family poverty (free/reduced lunch eligibility) as a time-varying covariate because student free and reduced lunch eligibility changes from year to year due to changes in family income. For the population used in this study, the family poverty level of the student changes at least once for about 15% of the students. School moves are separated into structural and non-structural based on whether a school switch was mandated by school district policy (i.e., primarily the division between elementary and middle school grades). We impute missing values for covariates at time *t* by assigning the subject-specific panel average. For example, if a student has a missing value in their panel of the family poverty indicator, we impute the average of that student's family poverty indicators across their other panels. For the dependent variable, math test score, we drop subjects whose panels are comprised of less than half non-missing scores and then impute with the grade level average of students who were ever missing, since students who were ever missing had lower test scores than kids who were never missing. A table of means before and after imputation for analysis variables is shown in appendix table A1.

METHODS

Cross-Sectional Model

To reproduce cross-sectional estimates commonly reported in previous research, we begin by presenting point-in-time estimates of the association of peer poverty on student achievement from a multilevel model (students nested within classrooms). We model math achievement, *A*, for student *i* in classroom *j* as a function of peer poverty and **X**, a vector of time varying and time invariant student covariates which includes student's own family poverty status:

(1)

In (1), we include a random intercept for each classroom, , and estimate (1) by grade level to examine whether the effect of peer poverty varies by grade. We hypothesize that association should increase with age, which could portend a widening of the gap between high and low poverty classroom exposure over time.

Growth Model

Using test score data that are interval scaled and vertically equated to allow for growth modeling, we estimate a non-linear growth model with a random intercept. Researchers in sociology, psychology, education, and criminology often use multilevel methods for growth modeling because it allows for unbalanced panels, permits a wide range of covariance structures, accounts for within-subject inter correlation, and often uses empirical Bayes estimation, which weights estimates by their reliability (the ratio of the true score variance to the observed score variance) (Bryk and Raudenbush 1987; Singer and Willett 2003). We formulate our quadratic growth model as:

(2)

This random intercept growth model regresses a math achievement test score, A, at time *t* for student *i* on grade level, grade squared, a classroom peer poverty indicator, interactions between grade and grade squared the peer poverty indicator, a vector of time-varying covariates, , and a vector of time-invariant covariates, **X**. The interpretation of these coefficients is as follows: β_0 is student initial status (3rd grade, in our case), β_1 is the instantaneous average growth rate for each increase in grade level, β_2 is the curvature parameter (this is negative in our models, so it represents the rate of deceleration with each increase in grade level), β_3 is the effect of peer poverty on initial status of achievement, β_4 is the change in β_1 from an increase in peer poverty, and β_5 is the change in β_2 from an increase in peer poverty. Random intercept parameter estimates on the variables of interest are virtually identical to those produced by a random coefficients model with random effects for grade and grade^2 (comparison of results shown in appendix table A5). We therefore present only random intercept models in our results below.

Parameters estimated with this method are unbiased and efficient assuming that given the covariates, the random effects and the student-level residual, , are normally distributed with zero mean, are independent of one another, with the random effect independent across subjects and independent across subjects and occasions.[10] The growth model produces an unbiased estimate of the effect of peer poverty on test score growth if peer poverty is uncorrelated with the random effects, if the variables in are exogenous, and if family background is adequately controlled and well measured. Omitted variable bias could produce inconsistent parameter estimates, which could threaten the validity of this model. For example, if the model used to estimate the effect of peer poverty on student test score growth omits adequate measures of parenting skill and if more skilled parents are less likely to send their children to high poverty schools and more likely to have high achieving students, then the estimate of peer poverty on student test score achievement would be downwardly biased. Although multilevel models can

[10] We specify growth models as two-level models, occasions within subjects, rather than occasions nested within subjects cross-nested in classrooms due to computational limitations, our focus on the parameter estimates rather than the random effects, and the fact that with population level data efficient estimates of standard errors are a secondary concern. Moreover, simulation evidence suggests that ignoring cross-nesting is likely to affect the variance components and not the parameter estimates (see Luo, Wen and Oi-man Kwok. 2009. "The Impacts of Ignoring a Crossed Factor in Analyzing Cross-Classified Data." *Multivariate Behavioral Research* 44:182-212.).

increase efficiency due to the use of both within and between variance, such models provide no solution for this type of confounding bias. If the between-student effects of peer poverty are large relative to the within-student effects, it is possible that the omission of student and family background characteristics biases our estimates of peer poverty contextual effects. In thinking about bias, it is helpful to return to our explanations of peer poverty effects: contagion, relative deprivation, collective socialization, and institutions. Peer effects can emerge either because students affect each other or because adults in schools affect students. The former pertains to contagion and relative deprivation explanations; the latter to an institutional or collective socialization explanation. In either case, the validity of inferences about contagion or institutional effects hinges on removing the confounding effects of student and family background.

Student Fixed Effects Model

We use a student fixed effects model to control for fixed unobservables such as innate ability and early childhood experiences that might confound the effect of peer poverty on test score growth. The fixed effects formulation uses each student as his/her own baseline, which holds constant all observable, unobservable, and mismeasured time-invariant student and family background characteristics. This approach eliminates all time-invariant between-student confounding in the peer poverty effect and produces consistent parameter estimates when there is no within-student confounding of the peer poverty effect (i.e., that the peer poverty effect is uncorrelated with time-varying unmeasured student characteristics). The student fixed effects model is specified as:

(3)

Here we treat the subject-specific intercept as a fixed unknown parameter to be estimated, with representing the deviation of subject *i*'s intercept from the mean intercept with . This model is often estimated by "demeaning" both sides of the equation by the subject's panel mean, which removes between-student confounding by using only within-subject variation to estimate parameters. Omitted from equation (3) are time invariant covariates because these have no within-subject variance and are therefore not estimable with this approach (though their effects are subsumed into the subject-specific intercept).

In studies of peer achievement effects, but not peer poverty effects, the direction of causality may be quite difficult to determine because student achievement at time *t* and peer achievement at time *t* are simultaneously determined (Manski 1993). Determining the causal direction between peer poverty and achievement is more straightforward. We posit that peer poverty level affects student achievement initial status and growth and that student achievement at time *t* does not affect peer poverty at time *t*. This seems like a reasonable assumption given that student academic performance has no bearing on their parents' earning power.

The student fixed effects approach requires within-student variation on peer poverty to identify parameters and is relatively inefficient relative to the random effects models. Due its large sample size and the six-year panels within it, however, our data are well suited to this approach. We identify the peer poverty effect from year-to-year variation in the poverty composition of students' classrooms. This changes due to school mobility and due to variations in the poverty compositions of student's assigned classrooms as they progress through grade levels in the same school. Because peer poverty rates vary more between schools than within schools, school movers are somewhat more likely to experience a change in peer poverty than students who remain in the same school. Nearly the entire sample makes some sort of school move during their panel: 85% of students make a structural move, 35% of students make a non-structural, and 91% of students make either a structural move or a non-structural move or both. The evidence suggests that across time variation exists to analyze for both school stayers and school movers, but that a larger portion of the variation that is analyzed appears to come from movers.

Marginal Structural Model

Both the multilevel and fixed effects models outlined above are vulnerable to the threat of time-varying confounding, which arises when there is a time varying variable that is affected by prior treatment and is associated with subsequent treatment and the outcome. For example, consider the causal diagram in figure 1. In this diagram, T0 and T1 represent the treatment, peer poverty, at baseline and time 1, X0 is a time varying confounder (which could be achievement, assignment to gifted or special education, or being retained in grade), Y1 is the outcome at time 1, and U is an unobservable that affects both X0 and Y1. X0 is a time-varying confounder because it predicts future treatment, T1, and is associated with future outcome, Y1, via U (Robins, Hernan, and Brumback 2000). Because X0 is affected by prior treatment (i.e., endogenous), standard models will produce biased estimates of the effect of T0 and T1 on Y1.

Time-varying confounding presents a dilemma: X0 is a confounder for later treatment and thus must be controlled, but may also be affected by earlier treatment and thus cannot be controlled (Robins, Hernan, and Brumback 2000). Because X0 is a collider (it is an effect of both T0 and U), controlling for X0 introduces collider stratification bias (Cole, Platt, Schisterman, Chu, Westreich, Richardson, and Poole 2010; Greenland 2003; Hernán, Hernández-Díaz, and Robins 2004). It has been shown that under the assumption of no unmeasured confounders or sequential strong ignorability (that treatment assignment is conditionally independent of the current and future potential outcomes given the measured past), marginal structural modeling (MSM) fit using inverse-probability-of-treatment (IPT) weighting can account for time-varying confounding and produce asymptotically consistent estimates of treatment effects in longitudinal analysis. This approach involves first computing IPT weights from each subject's probability of having their own treatment history and second estimating an IPT-weighted regression model. To construct stabilized[11] IPT weights, we compute:

(4)

where *t* indexes time, is treatment actually received (peer poverty exposure), ***X*** is a vector of time-invariant and time-dependent confounds, and *G* is grade level, and variables subscripted with a 0 represent baseline values. The denominator is, informally, a student's conditional probability of receiving her own observed treatment up to time *t*, given past treatment and covariate history and grade level. The numerator is, informally, a student's conditional probability of receiving her own observed treatment up to time *t*, given past treatment, baseline covariates, and grade level. Following standard practice in epidemiology, our MSM models also adjust for possible bias due to selective attrition. We compute a stabilized censoring weight as:

(5)

where *t*, *Z*, ***X***, and *G* are defined above, and is an indicator for student became censored at time t (i.e., it was his last observation in his panel). To adjust for both the inverse of the

[11] Following standard practice we compute stabilized weights which have lower variance than non-stabilized weights.

probability of treatment and censoring, the weight used in the MSM models is .
Our MSM is a weighted version of the growth model shown in equation (2).

RESULTS

We begin by presenting non-parametric and parametric plots to show changes in test score growth over time for students with varying levels of peer poverty exposure. We then discuss cross-sectional estimates of both our measures of peer poverty which show substantial associations between our two measures of peer poverty and student test score, especially in the middle school grades. Following this, we present growth model estimates of classroom poverty effects and cumulative exposure, and alternative specifications with student fixed effects and inverse-probability-of-treatment weighting. We produce growth curve trajectories from all three of these specifications to provide visual evidence of the quantitative results.

Figure 2 displays box plots of the distribution of math test scores by grade level and peer poverty exposure. Panel A shows differences between students in high- and non-high poverty classrooms and panel B shows differences based on the two extremes of cumulative exposure: always exposed and never exposed to a high poverty classroom. The plots show a general upward trend in scores and a reduction in the inter-quartile range for both groups as students increase in grade level. The median score for third grade students in high poverty classrooms lies just above the 25th percentile of third grade students in low poverty classrooms. By eighth grade, however, the median score for students in high poverty classrooms is just below the 25th percentile of students in low poverty classrooms, suggesting a small widening of the test score gap. The widening of the gap is more noticeable among eighth graders always and never exposed, with the 75th percentile of the always exposed roughly equal to the 25th percentile of the never exposed. Parametric plots from OLS models regressing math test score on grade,

$grade^2$, peer poverty, peer povertyXgrade, and peer povertyX$grade^2$ (with no control variables), show more clearly the non-linear growth of math test score achievement and the widening of the eighth grade test score gap between those always and never exposed to high poverty peers (figure 3). In numeric terms, the 3rd grade gap between these two groups is about 5.5 scale points and the 8th grade gap is about 11.5 points (see figure 3b). By contrast, the analogous figures for point in time (rather than cumulative) exposure are 5.5 and 6.5 points (see figure 3a). In summary, this descriptive analysis shows that the extent of widening depends on the measure of peer poverty exposure and that contrasting the two extremes of the cumulative measure produces the larger change in the gap.

Cross-Sectional Estimates of Classroom Poverty and Cumulative Exposure

We produce cross-sectional estimates from students-within-classrooms multilevel random intercept models, which control for (but do not display) race, gender, parental education, family poverty status, gifted, special education, limited English proficiency status, grade retention, and structural and non-structural school mobility (see equation (1)). We estimate models separately by grade level and in panels A and B of table 2 and we separately estimate the effects of two measures of peer poverty exposure: high poverty classroom and cumulative exposure to a high poverty classroom (i.e., table 2 displays summary results from 12 separate regressions). Net of controls, the third grade cross-sectional effect of attending a high poverty classroom is -.869 scale points, or -.081σ. Between fifth and sixth grade, this effect jumps from -.133σ to -.233σ. By eighth grade this estimate has grown to -.279σ. The fact that the middle school math scores of students in high poverty classrooms are about one-quarter of a σ lower than the scores of students in high poverty classrooms may reflect the differentiation of math curriculum in middle school (Dauber, Alexander, and Entwisle 1996; Hallinan 1992) and the growing influence of peers for young adolescents (Buchmann and Dalton 2002; Furman 1982; Veronneau and Dishion 2011). As expected, the effects of cumulative exposure to peer poverty produce larger point estimates. Recall that a one-unit change in the cumulate exposure variable represents the difference between subjects never exposed and always exposed to peer poverty at time *t*. Not surprisingly, the size of the third grade effect in panel B is approximately equal to the size of the third grade effect in panel A (-.085σ versus -.081σ). By eighth grade, however, the effect of cumulative exposure is -.425σ, much larger than the eighth grade effect of high poverty class (-.279σ). In summary, we find substantively large cross-sectional associations of high poverty classroom and cumulative exposure to high poverty classrooms that increase with age and become especially large by eighth grade, suggesting that test score trajectories may widen over time for students exposed to higher and lower peer poverty classrooms. As we have argued above, however, the most appropriate model to determine whether the gap between poor and less poor classrooms is to estimate a growth model, not a cross-sectional one. A growth model produces regression adjusted trajectories, which provide much more convincing evidence of the effect of context on student cognitive growth than a point-in-time estimate.

Growth Model Estimates of High Poverty Classroom

Table 3 displays coefficients from time-nested-within-student random intercept growth models (see equation 2, above). The fully unconditional model (1) decomposes math test score variation into between-student (   ) and within-student (   ) components. This model produces an estimate of intraclass correlation, ρ, of .48, which means that 48% of the variance in math test scores lies between students. Including grade level and the square of grade level (model 2), reduces the within-student variation in test score by about 50% and has very little effect on the between-student variance component,   . Because all variables in the models shown in table 3 except for the growth parameters are grand mean centered, and the grade level instantaneous growth parameter is coded 0 to 5, the constant represents initial status for the average 3rd grade student. In column (2), the average third grade student has a score of 339.2, an instantaneous growth rate of 7.229, and a negative curvature parameter, -0.653, which indicates that students' rate of change in test score growth declines over time. Including fixed and time-varying student characteristics (model 3), reduces between-student variation by about 18% (relative to the fully unconditional model). Most of the coefficients on these control variables conform to expectations, with negative differences in initial status for minorities, and poor, special education, limited English proficient, non-structural school movers, and retained students, and positive differences for academically gifted, and students with highly educated parents. Net of grade level, students making structural moves suffer test score decrements, which suggests that the event of changing schools between elementary and middle school has a negative effect on students that is analytically distinguishable from the effect of age (most, but not all, students change schools between 5th and 6th grade in North Carolina). Including the poverty rate of the student’s grade level peers (model 4) explains virtually no additional variation over and above the variation explained by the variables included in model 3. High poverty class has a negative effect on initial status, producing a .325 decrease in 3rd grade math test score. This represents 0.027σ of the math test score. As noted above, the linear growth rate in math is 4.2. The peer poverty main effect, therefore, represents about 8% of a year’s growth. In model (5), the peer poverty main effect increases from -.325 to -.997, but the interactions of the main peer poverty effect and the growth parameters, .0798 and -.144, while statistically distinguishable from zero, combine to produce only small changes in growth rates and fail to produce a substantial incremental reduction in within-student variance. This model predicts that at baseline, -.997 scale points separate students with in higher and lower poverty classrooms. The corresponding figure at 8th grade is -.607, indicating a small closing of the test score gap between students with high poverty and low poverty peers (9% of a standard deviation at baseline and 6% of a standard deviation by 8th grade).

Fixed Effect and MSM Growth Model Estimates of High Poverty Classroom

Growth model estimates are unbiased assuming all confounders are controlled. A student fixed effects model controls for fixed pre-baseline unobservables such as innate ability and early childhood experiences that might confound the effect of peer poverty on test score growth. By using only within-student variation in peer poverty, this approach eliminates all time-invariant between-student confounding and produces unbiased parameter estimates when all time-varying confounders are controlled. The strength of the student fixed effects model is adjustment for time-invariant unobservables. A weakness of both the growth model and the student fixed effects specification is that neither appropriately adjusts for time-dependent confounding. A marginal structural model (MSM) with inverse-probability-of-treatment weighting (IPTW) is designed to address time-dependent confounding. Table 4 presents estimates from the primary coefficients of interest from the random effects growth model shown in table 3, model 5 (labeled "RE" in table 4), a student fixed effects specification (see equation 3), and an MSM with IPTW (the random effects growth model estimated with the weights computed by equations 4 and 5), which each address two different threats to validity.

The estimates from these two alternative specifications essentially reproduce the growth estimates, suggesting that the latter estimates are not biased by fixed unobservables or time-dependent confounding. The student fixed effects model predicts that at baseline, the classroom poverty gap is -.655 (6%σ) and by eighth grade is -.285 (3%σ). The MSM model predicts a gap of -.764 at baseline (7%σ) and a gap of -.309 by eighth grade (3%σ). Therefore, each model predicts a slight narrowing of the gap, although the magnitudes of the effects are small.

Growth Model Estimates of Cumulative Exposure to Classroom Poverty

To address the concern that our estimates presented thus far could potentially underestimate the effect of classroom context by ignoring the cumulative nature of such effects, in table 5 we present growth model estimates with cumulative exposure to high classroom poverty as the explanatory variable. The coefficients on the main effect of cumulative exposure differ somewhat, ranging from -1.57 in the random effects growth model to -.503 in the fixed effects model, though the coefficients on the growth parameters are all quite similar, which indicates that trajectories assume similar shapes. The random effects growth model predicts a -1.57 point gap (14%σ) in third grade and a -1.60 point gap (15%σ) in eighth grade; the student fixed effects specification predicts a -.503 point gap (5%σ) in third grade and a -.038 point (.3%σ) in eighth grade; the MSM model predicts a -.865 point gap (8%σ) in third grade and a -.53 point gap (5%σ) in eighth grade. Therefore, these models predict either an unchanging gap, or one that closes slightly over time.

Growth Trajectories

Due to the apparently modest changes in growth curve trajectories and the statistical significance of all key estimates due to the large number of subjects in the study, perhaps visual, rather than numeric, evidence of the trajectories predicted by each model is more compelling. From the estimated coefficients of the intercepts, grade, grade^2, peer poverty measure, gradeXpeer poverty, and grade^{2}Xpeer poverty shown in tables 4 and 5, we plot predicted test score growth trajectories for exposed and unexposed students from grades three to eight in figure 4, holding all control variables at their means. This figure shows plots for students exposed and unexposed to classroom peer poverty that are virtually indistinguishable. In one case, figure D, the cumulative exposure trajectories estimated from the growth model, the lines are distinguishable but grow at approximately the same rate. Therefore, while we find quite strong cross-sectional associations between peer poverty and student test score, we find that exposure to high classroom poverty does not affect math test score growth in elementary and middle school aged students.

CONCLUSION

Much ink has been spilled searching for evidence of contextual effects on youth outcomes. What do we add to this rich literature? This study is designed to test the hypothesis that classroom contexts with high levels of peer poverty harm test score growth. This study moves beyond a conception of contextual effects as correlations estimated on young people at one point in time to context shaping youth development over time. Our findings suggest that previous research based on cross-sectional designs which tend to report negative effects of peer poverty on student achievement may be misleading. With cross-sectional evidence, we establish that exposure to high poverty classrooms is negatively associated with math test score, with the strength of the association becoming quite large as students increase in grade level. The growth model evidence presented, however, shows essentially no differences in the growth trajectories between those with poor and non-poor peers. Models with student fixed effects, which controls for time-invariant student background unobservables, and with inverse-probability-of-treatment weighting, a method to properly adjust for observable time-dependent confounding, also produce null effects of peer poverty rate on test score growth. That exposure to classroom poverty has a strong association with test score in the cross-section, but has no effect on achievement growth, strongly suggests that selection bias is present in the cross-sectional estimates reported in the many other studies based on point-in-time designs. The question of whether the effect of school poverty is causal or simply a function of either omitted variable bias or endogenous self-selection is a critical conceptual and empirical matter for both the theory of school effects and policies that seek to integrate students by socio-economic background (Duncan and Raudenbush 1999).

This study has important implications for both research and public policy. These findings suggest that standard estimates and prevailing theories about social influence among pre- and early-adolescents may not hold for test score achievement, one of the most important educational outcomes. This suggests that simply mixing students by poverty level and not altering important institutional resources such as high quality instruction and teacher expectations may not have the intended effect of increasing achievement because achievement and achievement growth are not a function of peers but of student and family background. The policy goal of mixing students by race-ethnicity or social background has been a mainstay in educational policy since the Brown vs. Board decision. Since the 1980s, school desegregation orders have been vacated by an increasingly conservative judiciary. The changing legal landscape has contributed to a resegregation of American schools (Orfield, Eaton, and Desegregation 1996; Reardon and Yun 2005; Rumberger and Palardy 2005) These trends are likely to continue given that in the 2007 case of *Parents Involved v. Seattle School District No. 1,* the Supreme Court ruled that school districts may not use race in assigning students or granting transfers to achieve or maintain school integration. In response to increases in school racial segregation and the Supreme Court's prohibition on the use of race in making school assignments, some advocate for integrating students based on socio-economic background (Kahlenberg 2001), which is constitutionally permissible. Kahlenberg (2001) argues that the best way to ensure the presence of high standards, highly qualified teachers, and less crowded classes is to ensure a critical mass of middle class families to advocate for these resources. Various forms of SES integration have been implemented in more than 50 districts across the U.S., including Lacrosse, WI; Wake County, NC; Cambridge, MA; and San Francisco, CA. The findings of the present study suggest that simply mixing students by social background may not have the intended effects, unless such mixing can garner increased resources and support for proven teaching practices that can increase student achievement in impoverished contexts.

There are some limitations to this study that point the way for future work on peer poverty effects. Although North Carolina is racially and economically diverse, the study covers only the public school students from one state, which limits the generalizability of our findings. Using population-level administrative data, we have pursued an identification strategy that privileges reduction of bias over national representativeness. The external validity of these results will hinge on cross-state replications using administrative data, and large nationally representative surveys with rich contextual information and interval metric test scores designed to measure growth over time. The estimates in this study are based on observational data, not an experiment, so we cannot claim them to be causal estimates. We must be careful to stress that we use a research design that reduces, but probably does not entirely eliminate, bias from unobservables. For example, our inability to account for time-varying student or school unobservables could prove these estimates to be biased. Future work should carefully theorize and measure time-varying factors that predict test score growth. We cannot empirically examine whether changes in peer poverty correspond to substantial differences in micro-level interaction between students and students and teachers. Within-school sorting processes – tracking, social homophily, and inequitable access to resources – may limit the actual changes in context students experience from changes in peer poverty rates due to school transitions, although our use of classroom-level data is a benefit of our study relative to those measuring contextual effects at the school or census tract level. Finally, test scores may be mostly impervious to the influence of peers, institutional resource allocation, and socialization processes. Other outcomes such as pregnancy, drug use, school completion, and college attendance, may be more amenable to these factors than a test score, which is a discrete cognitive task rather than a behavioral event.

Much of the existing research base on contextual effects has examined the experiences and outcomes of high school students. This study represents one of the first sociological examinations of contextual effects among elementary and middle school students. Despite this contribution, it may be that by third grade, the earliest time point in this study, early childhood experiences have largely determined a student’s potential for test score growth. If test score gaps among socioeconomic groups are essentially stable by third grade and variations in school quality have little effect on these gaps over time (Heckman 2006), then policies to mix students by social background may be of limited utility. Still, research has found neighborhood effects on birth weight and other early childhood development experiences (e.g. Chase-Lansdale, Gordon, Brooks-Gunn, and Klebanov 1997; Masi, Hawkley, Piotrowski, and Pickett 2007; Morenoff 2003). Therefore, a promising avenue for future research may be the investigation of younger children's sensitivity to impoverished contexts, preferably with research designs that permit accounting for unobserved family or individual heterogeneity and time-dependent confounding.

REFERENCES

Alexander, Karl and Bruce K. Eckland. 1975. "Contextual Effects in High-School Attainment Process." *American Sociological Review* 40:402-416.

Alexander, Karl L., James Fennessey, Edward L. McDill, and Ronald J. D'Amico. 1979. "School SES Influences - Composition or Context." *Sociology of Education* 52:222-237.

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.

Alwin, Duane F. and Luther B. Otto. 1977. "High-School Context Effects on Aspirations." *Sociology of Education* 50:259-273.

Attewell, Paul. 2001. "The Winner-Take-All High School: Organizational Adaptations to Educational Stratification." *Sociology of Education* 74:267-295.

Bachman, Jerald G. and Patrick M. O'Malley. 1986. "Self-Concepts, Self-Esteem, and Educational-Experiences - The Frog Pond Revisited (Again)." *Journal of Personality and Social Psychology* 50:35-46.

Barr, Rebecca and Robert Dreeben. 1983. *How Schools Work*. Chicago, IL: University of Chicago Press.

Battistich, Victor, Daniel Solomon, Dong-il Kim, Marilyn Watson, and Eric Schaps. 1995. "Schools as Communities, Poverty Levels of Student Populations, and Students Attitudes, Motives, and Performance - a Multilevel Analysis." *American Educational Research Journal* 32:627-658.

Bazelon, Emily 2008. "The Next Kind of Integration." *New York Times Magazine*, July 20, 2008.

Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela Kato Klebanov, and Naomi Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99:353-395.

Bryk, Anthony S. and Stephan W. Raudenbush. 1987. "Application of Hierarchical Linear-Models to Assessing Change." *Psychological Bulletin* 101:147-158.

Buchmann, Claudia and Ben Dalton. 2002. "Interpersonal Influences and Educational Aspirations in 12 Countries: The Importance of Institutional Context." *Sociology of Education* 75:99-122.

Chase-Lansdale, P. Lindsay, Rachel A. Gordon, Jeanne Brooks-Gunn, and Pamela Kato Klebanov. 1997. "Neighborhood and Family Influences on the Intelllectual and Behavioral Competence of Preschool and Early School-Age Children." Pp. 79-118 in *Neighborhood Poverty: Context and Consequences for Children*, vol. 1, edited by J. Brooks-Gunn, G. J. Duncan, and J. Aber. New York, NY: Russell Sage Foundation.

Choi, Kate H., R. Kelly Raley, Chandra Muller, and Catherine Riegle-Crumb. 2008. "Class Composition: Socioeconomic Characteristics of Coursemates and College Enrollment." *Social Science Quarterly* 89:846-866.

Clampet-Lundquist, Susan and Douglas S. Massey. 2008. "Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment." *American Journal of Sociology* 114:107-43.

Cole, S.R. and M.A. Hernán. 2008. "Constructing Inverse Probability Weights for Marginal Structural Models." *American Journal of Epidemiology* 168:656.

Cole, S.R., R.W. Platt, E.F. Schisterman, H. Chu, D. Westreich, D. Richardson, and C. Poole. 2010. "Illustrating Bias due to Conditioning on a Collider." *International Journal of Epidemiology* 39:417.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander J. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. "Equality of Educational Opportunity." USGPO, Washington.

Crane, Jonathan. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology* 96:1226-1259.

Crosnoe, Robert. 2009. "Low-Income Students and the Socioeconomic Composition of Public High Schools." *American Sociological Review* 74:709-730.

Dauber, Susan L., Karl L. Alexander, and Doris R. Entwisle. 1996. "Tracking and Transitions through the Middle Grades: Channeling Educational Trajectories." *Sociology of Education* 69:290-307.

Davis, James A. 1966. "Campus as a Frog Pond - Application of Theory of Relative Deprivation to Career Decisions of College Men." *American Journal of Sociology* 72:17-31.

DeLuca, Stefanie and Elizabeth Dayton. 2009. "Switching Social Contexts: The Effects of Housing Mobility and School Choice Programs on Youth Outcomes." *Annual Review of Sociology* 35:457-91.

Duncan, Greg J. 1994. "Families and Neighbors as Sources of Disadvantage in the Schooling Decisions of White and Black Adolescents." *American Journal of Education* 103:20-53.

Duncan, Greg J. and Jeanne Brooks-Gunn. 1997. *Consequences of Growing up Poor*. New York: Russell Sage Foundation.

Duncan, Greg J. and Stephen W. Raudenbush. 1999. "Assessing the Effects of Context in Studies of Child and Youth Development." *Educational Psychologist* 34:29-41.

England, Paula, Paul D. Allison, and Yuxiao Wu. 2007. "Does Bad Pay Cause Occupations to Feminize, Does Feminization Reduce Pay, and How Can We Tell with Longitudinal Data?" *Social Science Research* 36:1237-1256.

Entwisle, Doris R., Karl L. Alexander, and Linda S. Olson. 1994. "The Gender Gap in Math: Its Possible Origins in Neighborhood Effects." *American Sociological Review* 59:822-838.

—. 2005. "First Grade and Educational Attainment by Age 22: A New Story." *American Journal of Sociology* 110:1458-1502.

Evans, William N., Wallace E. Oates, and Robert M. Schwab. 1992. "Measuring Peer Group Effects - A Study of Teenage Behavior." *Journal of Political Economy* 100:966-91.

Feins, Judith D. and Mark D. Shroder. 2005. "Moving to Opportunity: The Demonstration's Design and its Effects on Mobility." *Urban Studies* 42:1275-1299.

Felmlee, Diane and Donna Eder. 1983. "Contextual Effects in the Classroom - The Impact of Ability Groups on Student Attention." *Sociology of Education* 56:77-87.

Furman, Wyndol. 1982. "Children's Friendships." Pp. 327-339 in *Review of Human Development*, edited by T. M. Field. New York, NY: Wiley.

Galster, George, Dave E. Marcotte, Marv Mandell, Hal Wolman, and Nancy Augustine. 2007. "The Influence of Neighborhood Poverty during Childhood on Fertility, Education, and Earnings Outcomes." *Housing Studies* 22:723-751.

Gamoran, Adam. 1987. "The Stratification of High-School Learning Opportunities." *Sociology of Education* 60:135-155.

Garner, Catherine L. and Stephen W. Raudenbush. 1991. "Neighborhood Effects on Educational Attainment: A Multilevel Analysis." *Sociology of Education* 64:251-262.

Grant, Gerald. 2009. *Hope and Despair in the American City*. Cambridge, MA: Harvard University Press.

Greenland, S. 2003. "Quantifying biases in causal models: classical confounding vs collider-stratification bias." *Epidemiology* 14:300.
Halaby, Charles N. 2004. "Panel Models in Sociological Research: Theory into Practice." *Annual Review of Sociology* 30:507-544.
Hallinan, Maureen T. 1992. "The Organization of Students for Instruction in Middle School." *Sociology of Education* 65:114-127.
Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2009. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." *Journal of Labor Economics* 27:349-383.
Harding, David J. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping out and Teenage Pregnancy." *American Journal of Sociology* 109:676-719.
Harris, Douglas N. 2010. "How Do School Peers Influence Student Educational Outcomes? Theory and Evidence from Economics and Other Social Sciences." *Teachers College Record* 112.
Hauser, Robert M. 1970. "Context and Consex - Cautionary Tale." *American Journal of Sociology* 75:645-&.
Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312:1900-1902.
Hernán, M.A., S. Hernández-Díaz, and J.M. Robins. 2004. "A structural approach to selection bias." *Epidemiology* 15:615.
Hernan, Miguel Angel, Babette Brumback, and James M. Robins. 2000. "Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-positive Men." *Epidemiology* 11:561-570.
Hong, Guanglei and Stephen W. Raudenbush. 2008. "Causal Inference for Time-Varying Instructional Treatments." *Journal of Educational and Behavioral Statistics* 33:333-362.
Hoxby, Caroline M. and Gretchen Weingarth. 2005. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." *Working Paper*.
Jacobs, David and Jason T. Carmichael. 2001. "The Politics of Punishment across Time and Space: A Pooled Time-Series Analysis of Imprisonment Rates." *Social Forces* 80:61-89.
Jacobs, David and Daniel Tope. 2007. "The Politics of Resentment in the Post-Civil Rights Era: Minority Threat, Homicide, and Ideological Voting in Congress." *American Journal of Sociology* 112:1458-1494.
Jencks, Christopher and Susan E. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." Pp. 111-186 in *Inner-City Poverty in the United States*, edited by L. E. Lynn and M. G. H. McGeary. Washington, D.C.: National Academy Press.
Kahlenberg, Richard D. 2001. *All Together Now: Creating Middle-Class Schools Through Public School Choice*. Washington, D.C.: Brookings Institution Press.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75:83-119.
Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Mobility Experiment." *Quarterly Journal of Economics* 120:87-130.
Kocak, Ozgecan and Glenn R. Carroll. 2008. "Growing Church Organizations in Diverse US Communities, 1890-1926." *American Journal of Sociology* 113:1272-1315.

Kubitschek, Warren N. and Maureen T. Hallinan. 1998. "Tracking and Students' Friendships." *Social Psychology Quarterly* 61:1-15.
Lee, Valerie E., Anthony S. Bryk, and Julia B. Smith. 1993. "The Effects of High School Organization on Teachers and Students." Pp. 171-268 in *Review of Research in Education*, vol. 19, edited by L. Darling-Hammond. Washington, DC: American Educational Research Association.
Leventhal, Tama and Jeanne Brooks-Gunn. 2000. "The Neighborhoods They Live in: The Effects of Neighborhood Residence on Child and Adolescent Outcomes." *Psychological Bulletin* 126:309.
Lucas, Samuel R. 1999. *Tracking Inequality*. New York, NY: Teachers College Press.
Ludwig, Jens, Jeffrey B. Liebman, Jeffrey R. Kling, Greg J. Duncan, Lawrence F. Katz, Ronald C. Kessler, and Lisa Sanbonmatsu. 2008. "What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment." *American Journal of Sociology* 114:144-88.
Luo, Wen and Oi-man Kwok. 2009. "The Impacts of Ignoring a Crossed Factor in Analyzing Cross-Classified Data." *Multivariate Behavioral Research* 44:182-212.
Manski, Charles F. 1993. "Identification of Endogenous Social Effects - The Reflection Problem." *Review of Economic Studies* 60:531-542.
Marsh, Herbert W. 1987. "The Big-Fish Little-Pond Effect on Academic Self-Concept." *Journal of Educational Psychology* 79:280-295.
Marsh, Herbert W. and Kit-Tai Hau. 2003. "Big-fish-little-pond effect on academic self-concept - A cross-cultural (26-country) test of the negative effects of academically selective schools." *American Psychologist* 58:364-376.
Marsh, Herbert W., Chit-Kwong Kong, and Kit-Tai Hau. 2000. "Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools." *Journal of Personality and Social Psychology* 78:337-349.
Marsh, Herbert W. and John W. Parker. 1984. "Determinants of Student Self-Concept - Is It Better to Be a Relatively Large Fish in a Small Pond Even If You Dont Learn to Swim as Well." *Journal of Personality and Social Psychology* 47:213-231.
Masi, Christopher M., Louise C. Hawkley, Z. Harry Piotrowski, and Kate E. Pickett. 2007. "Neighborhood Economic Disadvantage, Violent Crime, Group Density, and Pregnancy Outcomes in a Diverse, Urban Population." *Social Science & Medicine* 65:2440-57.
Meyer, John W. 1970. "High School Effects on College Intentions." *American Journal of Sociology* 76:59-&.
Morenoff, Jeffrey D. 2003. "Neighborhood mechanisms and the spatial dynamics of birth weight." *American Journal of Sociology* 108:976-1017.
Mouw, Ted. 2003. "Social capital and finding a job: Do contacts matter?" *American Sociological Review* 68:868-898.
—. 2006. "Estimating the causal effect of social capital: A review of recent research." *Annual Review of Sociology* 32:79-102.
Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner-City Children*. Cambridge, MA: Ballinger.
Orfield, Gary, Susan E. Eaton, and Harvard Project on School Desegregation. 1996. *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education*. New York: New Press : Distributed by W.W. Norton & Company.

Orr, Larry, Judith D. Feins, Rrobin Jacob, Erik Beercroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling. 2003. "Moving to Opportunity: Interim Impacts Evaluation." U..S. Department of Housing and Urban Development, Washington, D.C.

Palardy, Gregory J. 2008. "Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis." *School Effectiveness and School Improvement* 19:21 - 49.

Quigley, John M. and Steven Raphael. 2008. "Neighborhoods, Economic Self-Sufficiency, and the MTO." *Brookings-Wharton Papers on Urban Affairs* 3:1-46.

Reardon, Sean F. and John T. Yun. 2005. "Integrating Neighborhoods, Segregating Schools: The Retreat from School Desegregation in the South, 1990-2000." Pp. 51-69 in *School Resegregation: Must the South Turn Back?*, edited by J. C. Boger and G. Orfield. Chapel Hill: University of North Carolina Press.

Robins, James M. 1999. "Association, Causation, and Marginal Structural Models." *Synthese* 121:151-179.

Robins, James M., Miguel Angel Hernan, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11:550-560.

Rubinowitz, Leonard S. and James E. Rosenbam. 2000. *Crossing the Class and Color Lines: From Public Housing to White Suburbia*. Chicago, IL: University of Chicago Press.

Rumberger, Russell W. and Gregory J. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *The Teachers College Record* 107:1999-2045.

Rumberger, Russell W. and J. Douglas Willms. 1992. "The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools." *Educational Evaluation and Policy Analysis* 14:377-396.

Sampson, Robert J. 2008. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure." *American Journal of Sociology* 118:189-231.

Sampson, Robert J., Patrick Sharkey, and Stephen W. Raudenbush. 2008. "Durable Effects of Concentrated Disadvantage on Verbal Ability among African-American Children." *Proceedings of the National Academy of Sciences of the United States of America* 105:845-852.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment." *Journal of Human Resources* 41:649-691.

Saporito, Salvatore and Deenesh Sohoni. 2007. "Mapping Educational Inequality: Concentrations of Poverty among Poor and Minority Students in Public Schools." *Social Forces* 85:1227.

Schneiberg, Marc, Marissa King, and Thomas Smith. 2008. "Social movements and organizational form: Cooperative alternatives to corporations in the American insurance, dairy, and grain industries." *American Sociological Review* 73:635-667.

Sedlak, Michael W., Christopher W. Wheeler, Diana C. Pullin, and Philip A. Cusick. 1986. *Selling Students Short: Classroom Bargains and Academic Reform in the American Classroom*. New York, NY: Teachers College Press.

Singer, Judith D. and John B. Willett. 2003. *Applied Longitudinal Data Analysis*. New York, NY: Oxford University Press.

Solon, Gary, Marianne E. Page, and Greg J. Duncan. 2000. "Correlations between neighboring children in their subsequent educational attainment." *Review of Economics and Statistics* 82:383-392.
South, Scott J., Eric P. Baumer, and Amy Lutz. 2003. "Interpreting community effects on youth educational attainment." *Youth & Society* 35:3-36.
Veronneau, Marie-Helene and Thomas J. Dishion. 2011. "Middle School Friendships and Academic Achievement in Early Adolescence: A Longitudinal Analysis." *Journal of Early Adolescence* 31:99-124.
Willms, J. Douglas. 1986. "Social Class Segregation and Its Relationship to Pupils' Examination Results in Scotland." *American Sociological Review* 51:224-241.
—. 2010. "School Composition and Contextual Effects on Student Outcomes." *Teachers College Record* 112.
Wilson, Alan B. 1959. "Residential Segregation of Social-Classes and Aspirations of High-School Boys." *American Sociological Review* 24:836-845.
Wooldridge, Jeffrey M. 2003. *Introductory Econometrics: A Modern Approach*. Cincinnati, OH: South-Western College Publishing.