Lasting Consequences of the Summer Learning Gap

Karl L. Alexander
Johns Hopkins University

Doris R. Entwisle
Johns Hopkins University

Linda Steffel Olson
Johns Hopkins University

Prior research has demonstrated that summer learning rooted in family and community influences widens the achievement gap across social lines, while schooling offsets those family and community influences. In this article, we examine the long-term educational consequences of summer learning differences by family socioeconomic level. Using data from the Baltimore Beginning School Study youth panel, we decompose achievement scores at the start of high school into their developmental precursors, back to the time of school entry in 1st grade. We find that cumulative achievement gains over the first nine years of children's schooling mainly reflect school-year learning, whereas the high SES–low SES achievement gap at 9th grade mainly traces to differential summer learning over the elementary years. These early out-of-school summer learning differences, in turn, substantially account for achievement-related differences by family SES in high school track placements (college preparatory or not), high school noncompletion, and four-year college attendance. We discuss implications for understanding the bases of educational stratification, as well as educational policy and practice.

Comparisons of school-year and summer learning inform fundamental questions of educational stratification and help parse school, family, and community influences on children's academic development. With children "in" their homes, schools, and communities during the school year, but just "in" their homes and communities over the summer months, the academic calendar approximates a natural experiment that affords leverage for isolating the distinctive role of schooling in children's cognitive development. This was the great insight exploited by Barbara Heyns in her 1978 book *Summer Learning,* which established that achievement gaps by family SES (socioeconomic status) and race/ethnicity widen more during the summer months than during the school year.

Although the detailed results of subsequent research on the seasonality of learning do not line up perfectly (see Cooper and colleagues' [1996] meta-analysis for an overview), the patterns documented by Heyns in the 1970s for middle school children in public schools in Atlanta, Georgia appear to have considerable generality. This is especially the case for her conclusions regarding family socioeconomic background, which have been replicated in our Baltimore research on the early elementary years with data from the 1980s (e.g., Entwisle, Alexander, and Olson 1997), in studies conducted in other localities (Murnane 1975; O'Brien 1998), and in national data from earlier (Heyns 1987; Karweit, Ricciuti, and Thompson 1994; Phillips 2000) and more recent periods (Burkam et al. 2004; Downey, von Hippel, and Broh 2004; Reardon 2003).

Direct correspondence to Karl Alexander, Department of Sociology, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218. This analysis was supported by the Spencer Foundation (Grant 20030057) and the W. T. Grant Foundation (Grants No. 9819298 and 95164195).

Using the national Early Child Longitudinal Study data (ECLS) over kindergarten and 1st grade, these recent analyses address a potentially critical source of bias in that they adjust seasonal comparisons for school year beginning and ending dates in relation to when students were actually tested. Other studies of seasonal learning patterns define in-school and summer learning around fall and spring testing dates—for example, the September through June "school year" becomes October through May, if that is the testing schedule. Consequently, the school year is typically understated and summer is overstated, such that an indeterminate portion of school-year learning is allocated to the summer months. While this imprecision probably moderates seasonal differences, ECLS analyses still find that "children from higher-SES families learn more over the summer than do their less-advantaged counterparts" (Burkham et al. 2004:18).

Findings from this literature support two conclusions: 1) prior to high school, the achievement gap by family SES traces substantially to unequal learning opportunities in children's home and community environments; and 2) with learning gains across social lines more nearly equal during the school year, the experience of schooling tends to offset the unequalizing press of children's out-of-school learning environments. Schooling thus appears to play a compensatory role, although we caution that this conclusion holds only for the experience of schooling writ large. It does not imply parity, or even near equivalence, in access to particular school resources or opportunities to learn, which often are quite unequally distributed (e.g., Dougherty 1996).

These insights inform our understanding of the roles played by families, neighborhoods, and schools in cognitive development over the short-term, but do they have consequences for later patterns of educational stratification? It seems reasonable that they would, yet research on the seasonality of learning has yet to inform the question. Rather, studies have been narrowly focused on establishing the seasonal pattern, and to a lesser extent on trying to account for it (e.g., investigating differences in the summer experiences of low-income and upper-income youth). Still, two bodies of evidence suggest there *ought* to be lasting consequences of summer learning differences over the elementary grades—consequences that are likely substantial.

First, achievement scores at any level of schooling predict success at the next level. This holds for high school completion, college attendance, college completion (see Entwisle et al. 1997, table 7.2), and later successes in the labor market (e.g., Kerckhoff, Raudenbush, and Glennie 2001). Second, cognitive achievement scores at the individual level are moderately to highly correlated across time. Most immediately relevant is the patterning of scores from the early elementary grades into middle school and high school. In national data, test scores measured in kindergarten and 1st grade correlate .5 and above with scores at 5th and 10th grade (e.g., Pope, Lehrer, and Stevens 1980; Weller, Schnittjer, and Tuten 1992), while in the present project fall and spring subtests from 1st grade on the California Achievement Test battery on average correlate .54 with their counterpart measures nine years later.

Here, then, is the argument in propositional form: (1) if the achievement gap by family SES during the elementary school years traces substantially to summer learning differences, and (2) if achievement scores are highly correlated across stages of young people's schooling, and (3) if academic placements and attainments at the upper grades are selected on the basis of achievement scores, then (4) summer learning differences during the foundational early grades help explain achievement-dependent outcome differences across social lines in the upper grades, including the transition out of high school and, for some, into college.

Though the argument seems plausible enough, it has yet to be put to the test. Using data from the Baltimore-based Beginning School Study (BSS) youth panel, our analysis examines consequences of seasonal learning differences during the elementary school years for children's later schooling.

SAMPLE AND METHODS

The BSS panel consists of a representative random sample of Baltimore school children whose educational progress has been monitored from 1st grade through age 22. The project began in the fall of 1982, when the study participants (N = 790), randomly selected from 20 public elementary schools within strata defined by school racial composition and socioeconomic level, were starting 1st grade. We use testing data from Baltimore City Public School System (BCPSS) records to track learning patterns, school records and student reports to identify students' high school curriculum placement (college preparatory versus others), and student interview data from an age 22/23 Young Adult Survey (YAS) to determine high school completion and college attendance. Questionnaire data from parents are combined with school record data about parents to rank children's family socioeconomic standing in elementary school. (For more detail on the BSS sampling and research design, see Entwisle and colleagues [1997]; see Table A1 in the Appendix for variable descriptions.)

We analyze scores on the Reading Comprehension subtest of the California Achievement Test (CAT-R) battery from school records over BSS years 1 through 5 (California Achievement Test 1979), fall and spring separately, and from a BSS administration of this same subtest in year 9 (analyses using the Math Concepts subtest yield quite similar results). The twice annual testing schedule for the early years allows separate calculation of school-year (fall to spring) and summer (spring to fall) learning gains over the entire elementary school period for children promoted regularly (repeaters are covered through the highest grade attained over these five years). The BCPSS discontinued fall achievement testing in BSS year 7 (the 1988 to 1989 school year), and after year 8 they discontinued use of the CAT battery altogether, but in spring 1991 (the end of 9th grade for children promoted regularly) the BSS did its own administration of the Reading Comprehension and Math Concepts subtests, achieving 75 percent coverage of the original panel.[1] The interval between the end of BSS year 5 and spring of year 9 spans the middle school years and the first year of high school. With no fall scores for those years, gains cannot be calculated separately by season, so overall gains are reported.

We use 11 testing points in the analysis (fall and spring for each of the first five years plus spring of year 9). This is an uncommonly rich set of testing data, but owing to absences, transfers outside the city school system, and other complications, not all children were tested on every occasion. Case coverage when screened on complete testing data is 326 (from 790 originally). Additionally, some positive selection is evident—while fall of 1st grade scores are close (281.7 for the listwise sample and 280.6 for the full sample), by year 9 the listwise group's spring average is .18 SD above the full sample average. However, the 464 excluded cases include many with nearly complete testing records, and some useful testing data are available for just about everyone. For example, 81 percent of cases have observed data for at least 6 of the 11 testing occasions, and 92 percent have data for at least four test scores.[2] To take advantage of this circumstance, we generated an imputed version of the raw data (based on 10 imputations) using multiple imputation methods (e.g., Allison 2002). These methods predict missing scores from the available data (including spring scores over years 6 through 8, which are not used in the substantive analyses), plus race, sex, and family SES background (the continuous version), which are known for all but three cases, and high school track placement. We used STATA software to carry out the imputation procedure.[3] The final imputed data set of test scores and seasonal components includes 787 cases (three cases that lacked data on family SES were dropped). Data were not imputed for high school track and the age 22 educational outcome variables used as dependent variables in the regression models.

The imputed achievement data were derived as the average of the 10 versions generated by the imputation process. We then used these scores (fall and spring over the elementary years and spring of year 9) to calculate the four achievement components used in the analyses: fall of 1st grade score, cumulative school-year gain over the elementary grades, cumulative summer gain over the elementary grades, and total gain over years 6 through 9. Because this

[1] Testing continued for roughly 18 months. Using exact testing dates, a linear interpolation referenced scores back to the spring of year 9.

[2] Missing data range from 6 percent to 31 percent across the 11 tests, averaging 22 percent.

[3] See Royston (2004, 2005) for documentation of the user-written programs Ice and Micombine, which implement multiple implementation in STATA.

is a somewhat unusual application of the multiple imputation methodology, we have carefully checked the robustness of the results it yields. The patterns of interest in analyses based on the pooled imputed data (including patterns of statistical significance) are evident also in each of the 10 replicate data sets separately, as well as the N = 326 full information subsample, which though small and probably somewhat atypical, nevertheless has strong internal validity. Accordingly, we report the results based on the pooled imputed data matrix, with supplementary points of interest from the checks mentioned as results are presented.

DISAGGREGATING YEAR 9 ACHIEVEMENT SCORES

The spring year 9 achievement average for the imputed analytic sample is 547.6 scale score points, with a standard deviation of 80.4 (see Table 1). The developmental foundation for this level of assessed performance early in high school (or, for repeaters, just before) is as follows: 1) a baseline average of 279.8 points from the fall of 1st grade; 2) an average *cumulative school-year gain* of 195.0 points fall to spring during elementary school; 3) an average *cumulative summer gain* of 11.1 points spring to fall over summers during the elementary years; and 4) an average cumulative school-year plus summer gain of 61.7 points over school years 6 though 9 (years for which we are unable to distinguish school-year from summer gains).

These achievement scores are vertically calibrated across years so as to approximate a continuous metric, but the distribution lacks a meaningful zero point (the fall of 1st grade *floor,* or lowest possible score, is 133). However, winter gains over the elementary school years account for 195.0 points of the 267.7 point increment from the baseline over this nine-year period, by far the largest of the three components in Table 1.

Important curricular placement decisions are made at the beginning of high school, and Table 1 shows that achievement assessed at that point mainly reflects skill differentials already in place when children enter 1st grade, as well as skills built up in elementary school. Summer learning in this instance is a small part of the overall picture. But what of *differences across social lines* in year 9 achievement levels? Does the breakdown look the same? The right-most column of Table 1 addresses this question, comparing the learning patterns of children classified as "low" and "high" SES in terms of family background (see Table A1 in the Appendix for measurement detail).

We focus here on comparisons across the SES extremes, but we keep in mind that the BCPSS enrollment is largely low income (half the BSS sample is classified low SES) and that few wealthy families send their children to public schools in low-income, high-poverty school systems. "Extreme," then, is relative to the local context. Still, within the BSS, low SES parents are mainly high school dropouts and high SES parents on average have attained some college, so this is a meaningful contrast, even if truncated relative to family differences nationally. In year 9, the high SES achievement average is 73.2 points above the low SES average (.88 SD, referenced to the standard deviation for high and

Table 1. Reading Comprehension Test Score Decomposition over the First Nine Years of School by Family SES

Reading Comprehension CAT Score Gains, Years 1–9	Total	Family SES: Low SES	Family SES: Mid SES	Family SES: High SES	Gap High-Low
Initial Test Score, Fall 1st Grade	279.81	271.99	277.89	298.47	26.48*
Winter Gain (5 winters)	194.97	191.30	210.19	186.11	–5.19
Summer Gain (4 summers)	11.12	–1.90	4.12	46.58	48.48*
Gain Over Years 6–9	61.69	60.95	60.73	64.34	3.39
Test Score, End Year 9	547.55	522.33	552.94	595.49	73.16*
(N)	(787)	(397)	(204)	(186)	

Note: Significant t-tests for mean differences between Low SES and High SES groups are shown in Gap column.
* $p \leq .05$ (two-tailed tests).

low SES youth combined). About a third of that SES difference, 26.5 points, traces to disparities in place when these children started 1st grade, implicating experiences and family resources that predate school entry.[4] The remainder of the difference is built up over the school years, and Table 1 shows that the largest component, 48.5 points, or about two-thirds of the total, traces to summer learning differences over the elementary years. The low SES group actually gains a bit more during the corresponding school years than does the high group (5.2 points, not a significant difference), but this favorable showing while in school is more than offset by their summer shortfall.

Importantly, and so far as we can tell, this pattern is not an artifact of ceiling limits on high SES children's school-year gains. This was checked on the full information, nonimputed cases, 64 of whom scored at the ceiling on one or more of the spring assessments (28 high SES youth, 15 mid SES, and 21 low SES). Excluding these 64 cases reduces the high SES–low SES gap as would be expected, but the summer differential remains the largest component and winter gains still favor the low SES group by a small, nonsignificant margin. Additionally, as shown elsewhere (Alexander, Entwisle, and Olson 2001), SES differences in summer learning are robust in HLM (hierarchical linear models) within-person growth models over the five school years for which we can distinguish summer learning from school year, with the summer learning difference significant each summer.

The early years of schooling are foundational in that the skills acquired then support all later learning. Our analytic comparisons support this point, insomuch as achievement levels at the start of high school substantially trace back to those early years. Moreover, most of that learning happens when children are in school, so schooling indeed makes a difference for low SES youth and high SES youth alike (for more on this point, see Downey et al. 2004). But with respect specifically to the year 9 achievement gap by SES background, experiences outside school apparently make an even bigger difference, as that gap substantially originates over the years before 1st grade and summer periods during the elementary school years.

SUMMER LEARNING DIFFERENCES AND SCHOOLING OUTCOMES

Do these large cognitive differences that trace back to the period before high school matter in practical terms? The top panel of Table 2 shows attainment outcomes in high school and at selected later benchmarks for the imputed sample, again comparing children classified by family SES. Sixty-two percent of high SES children were enrolled in a college preparatory program in high school versus just 13 percent of the low SES group. There are large differences in high school noncompletion and college attendance as well. Based on information covering the four years after the panel's on-time high school graduation in spring 1994,[5] over a third of the low SES group and just 3 percent of the high group are "permanent dropouts," meaning high school dropouts who at approximately age 22 still lack high school certification of any type. Whereas almost 60 percent of the high SES group attended a four-year college by age 22, just 7 percent of low SES youth did so.

To illustrate how summer learning differences *might be* implicated in the socioeconomic patterning of educational accomplishment in high school and after, the bottom two panels of Table 2 repeat the decomposition exercise from Table 1 across the socioeconomic extremes, but for subsets of youth whose social backgrounds and later experiences align: first, low SES youth in the non-college track and high SES youth in the college track; second, low SES permanent dropouts and high SES youth who attended four-year colleges.

[4] With kindergarten now nearly universal and preschool education common, it cannot be said that 1st grade represents children's first encounter with formal schooling. However, many children still attend half-day kindergartens (45 percent in the late 1990s [West, Denton, and Reaney 2001]), and many kindergartens stress social skills over academic learning. That said, children's achievement levels do improve over the kindergarten year, but at a slower rate than they do in 1st grade (Downey et al. 2004). In the early 1980s, when the BSS began, kindergarten was not yet mandatory in Baltimore.

[5] Owing to grade retention, dropouts, and other circumstances, fewer than half (45 percent) actually finished high school at that time.

Table 2. Reading Comprehension Test Score Decomposition Over the First Nine Years of School by Family SES, High School Track Placement, and Educational Attainment at Age 22

	Family SES			
	Low SES	Mid SES	High SES	Gap
Proportion College-Prep High School Track	.13	.30	.62	.49*
(N)	(320)	(176)	(160)	
Educational Attainment, Age 22				
Proportion Permanent Dropout	.36	.13	.03	–.33*
Proportion High School Graduate/GED	.35	.35	.11	–.24*
Proportion Trade School/Two-Year College	.21	.34	.27	–.06
Proportion Four-Year College	.07	.18	.59	.52*
(N)	(313)	(158)	(159)	

Reading Comprehension CAT Score Gains, Years 1–9	Low SES Non-College Track	High SES College Prep Track	Gap
Initial Test Score, Fall 1st Grade	269.88	310.30	40.42*
Winter Gain (5 winters)	188.20	180.58	–7.62
Summer Gain (4 summers)	–1.85	74.63	76.48*
Gain Over Years 6–9	60.44	67.21	6.77
Test Score, End Year 9	516.67	632.72	116.05*
(N)	(278)	(99)	

Reading Comprehension CAT Score Gains, Years 1–9	Low SES Permanent Dropouts	High SES 4-Year College	Gap
Initial Test Score, Fall 1st Grade	268.06	311.04	42.98*
Winter Gain (5 winters)	183.32	180.19	–3.13
Summer Gain (4 summers)	–11.04	75.53	86.57*
Gain Over Years 6–9	62.93	69.54	6.61
Test Score, End Year 9	503.26	636.30	133.04*
(N)	(114)	(94)	

Note: Significant t-tests for mean differences between Low SES and High SES groups are shown in Gap column.
* $p \leq .05$ (two-tailed tests).

Consider first the college track/high SES–non-college track/low SES comparison. There is a 116.1 point (1.3 SD) difference between the two groups' year 9 achievement averages, more than half of which (76.5 points) traces to summer learning differences carried forward from elementary school. The second largest component is the 40.4 point fall of 1st grade disparity.

The situation is much the same when comparing low SES permanent dropouts against high SES youth who attended four-year colleges, for whom the year 9 achievement difference is 133.0 points (1.4 SD). This huge disparity again traces substantially to the groups' unequal experience of summer learning over the early formative years, which accounts for 86.6 scale points, or 65 percent of the total. The next largest component, at 43.0 points, again is the fall of 1st grade difference.

The groups involved in these comparisons certainly are not distinguished solely by their achievement scores early in high school. For this reason, the comparisons in Table 2 cannot be said to isolate causality. In strictly empirical terms, out-of-school experiences account for the majority of the achievement differences registered in 9th grade, and these achievement differences, in turn, anticipate vastly different high school placements, modes of high school exit, and patterns of postsecondary attendance. More to the point, though, it seems certain that achievement levels at the start of high school *play some role* in the schooling outcomes in Table 2—for example, achievement scores and the competencies they signal are used in making curricular placement decisions; they inform parents', teachers', and counselors' thinking about students' academic prospects; and they are used in a self-referential way to inform a student's own sense of self in the student role.

It is a familiar pattern of educational stratification that disadvantaged social origins anticipate disadvantaged social destinations. Academic skill development, we know, plays a role in cementing that link. Low achievement scores at the start of high school do not auger well for later success, and it is low SES students and those who are socially disadvantaged in other ways who tend to fall toward the low end of the achievement distribution. But comparisons made only in the upper grades obscure the developmental history upon which students' high school records are built. Our results show how out-of-school learning during the elementary grades is linked to the year 9 achievement gap by family SES: a gap that, in turn, separates college track youth from non-college track youth, and that distinguishes those who fall off the path to high school completion from those who attend four year colleges.

REGRESSION ADJUSTED COMPARISONS

It may be the case that the comparisons in Table 2 exaggerate the role of summer differences by focusing on extreme cases. No doubt there is some truth to that. To gauge more formally the descriptive differences just reviewed, Table 3 presents high SES–low SES comparisons adjusted for background attributes. The entries, derived from group-specific logistic regression equations, are predicted probabilities of being in a college preparatory program in high school (estimated for the high SES and the low SES subsamples), of being a permanent dropout (versus any other educational status, estimated for the low SES sample only), and of attending a four-year college (versus any other educational status, estimated for the high SES subsample only). The equations adjust for differences associated with race (a dummy variable distinguishing blacks from whites), sex (females versus males), family SES (the full metric version to control for SES differences within the nominal "low" and "high" classifications), and four variables that make up the components of year 9 achievement: a baseline score from the fall of 1st grade; a measure of cumulative school-year gains over years 1 through 5; a measure of cumulative summer gains over summers 1 through 4; and a measure of total gains, summer and school-year, over years 6 through 9. The estimates in this way disaggregate effects of year 9 achievement scores, with some statistical controls to isolate group differences.

Table 3. Predicted Probabilities of Educational Outcomes from Logistic Regression Models

	Predicted Probability from Logistic Model[a]	Predicted Probability Substituting CAT Means for Opposite SES Group	Predicted Probability Substituting Only Summer CAT Gain for Opposite SES Group
Probability of College-Prep Curriculum for Low SES Sample	.07	.29	.16
Probability of College-Prep Curriculum for High SES Sample	.71	.32	.40
Probability of Permanent Dropout for Low SES Sample	.35	.23	.26
Probability of Attendance at Four-Year College for High SES Sample	.67	.24	.34

[a] Predicted probabilities come from logistic models predicting three different outcomes: permanent dropout, attendance at a four-year college, and enrollment in a college-prep high school curriculum. Each regression controls on demographic factors (gender, race, and a continuous composite measure of family SES). The regressions are estimated separately for low and high SES groups. Predicted probabilities are calculated for a sample member who scores at the mean on each predictor.

As one would expect, year 9 achievement effects themselves are predictive of all three outcomes. To establish the point, using the unimputed data we substituted year 9 achievement scores for the four components in Table 3. For analyses screened on just year 9 scores (N's 514–576), the analytic sample with complete testing data (N = 326), and the smaller high SES–low SES comparison subsamples, with race, sex, and family SES (metric version) controlled, effects of achievement in year 9 are statistically significant and substantively large throughout. Table 3 shows, in summary fashion, the extent to which mean differences across SES levels in the developmental components of year 9 achievement contribute to high SES–low SES differences in college track enrollment at the secondary level, high school noncompletion through age 22, and attendance at a four-year college.

Three predicted probabilities are reported for each outcome. The first is derived by evaluating the logistic equation for the focal group at that group's mean regressor values—high SES youth in the case of college track placement and four-year college enrollment; low SES youth in the case of college track placement and permanent dropout. This simply uses the group's own properties and regression results to generate an expected probability for the event at issue. The second and third estimates are a form of statistical experimentation, in the nature of regression decomposition as applied to least squares regression analyses (e.g., Althauser and Wigler 1972; Iams and Thornton 1975; Jones and Kelley 1984). In these instances, we apply the *other group's* mean achievement values (but not the three background measures) to the focal group's logistic coefficients—for example, applying the low SES group's achievement means to the high SES group's coefficients when predicting the high SES probability of college track enrollment. The first estimate substitutes all four achievement means; the second substitutes just the summer gain component (retaining the focal group's means for the other three components).

This exercise explores a "what if" counterfactual: What if everything else about the group at issue remains the same, but instead of having the achievement averages observed for them, they have the other group's averages? How would that affect, for example, their expected probability of college-prep enrollment? This is clearly an artificial exercise—it's hard to imagine changes in achievement averages of this magnitude with nothing else changing. Still, the calculations usefully highlight interesting properties of the data and allow us to go beyond simple descriptions to controlled comparisons (i.e., outcome differences adjusted for race, sex, and family socioeconomic background in metric form).

With college track enrollment the criterion, there are enough low SES and high SES youths to perform the experiment for both groups (this holds for both the imputed data and the original listwise sample). Based on their respective logistic results and own regressor averages, the estimated probability of enrolling in a college-preparatory high school program is .71 among high SES youth and .07 among low SES youth. Substituting the low SES achievement means into the high SES equation reduces the former group's predicted probability of college track enrollment to .32, or by more than half, as a function of the two groups' very different achievement averages. At the other extreme, substituting the high SES averages in the low SES equation increases college track enrollment prospects four-fold, from .07 to .29.[6]

When we instead substitute *just* the low SES youth's lower average summer gain for the high SES youth's higher average summer gain (and vice versa) in the third column, the estimates are midrange, with the predicted probability for the high SES group increasing to .40 (but still well short of their own estimate of .71) and the low SES group decreasing to .16 (but still double their own estimate of .07). The former reduces the initial .64 point difference in the probability of college track enrollment by roughly 48 percent—(.71 – .07) – (.40 – .07)/(71 – .07); the latter yields a 14 percent reduction.

There are hardly any unimputed high SES permanent dropouts or low SES youth who attended four–year colleges, so the other two sets of calculations in Table 3 are done for one group only. For high SES youth, the *estimated* probability of enrolling in a four-year college is .67. This drops to .24 (a .43 point reduction) when

[6] The logistic model is nonlinear and nothing in the logic or mechanics of this technique obliges symmetric results.

the calculations are based on low SES youth's achievement averages throughout and to .34 (a .33 point reduction) when just the mean value for the summer component is substituted. Both changes are large, and again the summer component alone produces most of the reduction in the predicted probability.[7]

The calculations predicting permanent high school dropout status indicate a smaller role for achievement differences altogether, but a consequential one nonetheless. The predicted probability of permanent dropout status for low SES youth is .35. This improves to .23 when all four high SES achievement means are substituted and to .26 when just the summer mean is used.

DISCUSSION

Our analysis adds an important practical dimension to research on the seasonality of learning, with implications for how the out-of-school institutional contexts of family and community that frame young children's academic development contribute to patterns of educational stratification. Stability in cognitive achievement over the course of young people's schooling is the bridge between summer learning shortfall over the elementary school years and later schooling outcomes. *Since it is low SES youth specifically whose out-of-school learning lags behind, this summer shortfall relative to better-off children contributes to the perpetuation of family advantage and disadvantage across generations.*

Low SES youth, we find, are less likely to find their way to a college-preparatory high school program, partly because their test scores are low at the very time these placements are made. And because their scores are low, they also are more prone to leave school without degrees and less likely to attend a four-year college. In light of these, and no doubt other, serious consequences, the question of why achievement levels at the start of high school are so disparate takes on great importance.

It is well established that there are vast differences across social lines in preschool children's out-of-school learning environments (e.g., Hart and Risley 1995). This helps explain not just why disadvantaged youth start school already far behind in kindergarten or 1st grade (Lee and Burkam 2002), but also why they continue to lag behind later (Farkas and Beron 2004; Phillips, Crouse, and Ralph 1998). Now we see that summer learning differences after children start school follow a like pattern, but what might not have been expected is the extent to which the *continuing press* of school-age children's family and neighborhood environments contributes to the year 9 achievement differential between high and low SES youth: summer shortfall over the five years of elementary school accounts for more than half the difference, a larger component than that built up over the preschool years.[8] And too, these learning differences from the early years that present themselves in 9th grade reverberate to constrain later high school curriculum placements, high school dropout, and college attendance. This lasting legacy of early experience typically is hidden from view.

The BSS is a local study and the analytic sample before imputation is small, just over 300. This is obviously limiting. Yet, the local context is urban and high poverty and thus policy relevant. Though the analytic sample may not be, strictly speaking, representative, these nevertheless are typical urban youth. But more to the point, to our knowledge there is no better data source, *anywhere,* for informing the issues addressed in this article. The national ECLS data, for example, include fall tests in kindergarten and 1st grade only, and so cover just the summer between kindergarten and 1st grade. Other samples used to study learning patterns by season also typically include data for just one summer, and none offer the long-term perspective of the BSS. Our results are best considered suggestive, and certainly the detailed percentages and probabilities reported should not be generalized. That said, BSS findings

[7] Using the imputed data, we get a .03 probability of four-year college attendance for low SES youth. This increases to .13 when calculated using high SES youth's four CAT-R means and to .07 when just the summer mean is substituted.

[8] For perspective on the time line of the black–white achievement gap, see Phillips and colleagues (1998). They estimate that half or more of the gap measured in 12th grade reflects continuity of differences evident at the start of 1st grade.

align well with a now sizeable literature on summer learning differentials, and the links seen in this analysis to later outcomes certainly have surface plausibility. With these caveats understood, we now discuss several implications of the findings presented here.

Surely the point made by David Berliner (2006) in his Invited Presidential Speech at the 2005 American Educational Research Association (AERA) annual meeting is correct: to moderate the achievement gap, the most compelling need is to reduce family and youth poverty. However, there also is a critical role for school reform. Achievement differentials by race/ethnicity and along lines of family advantage/disadvantage over the last 50 years have exhibited more volatility than many seem to realize (e.g., Krueger 1998; Lee 2002). Using National Assessment of Educational Progress (NAEP) testing data, for example, Hauser (1995) estimates that the I.Q. gap separating white and black youths declined by almost a third between 1970 and 1990, while Grissmer and colleagues (1994) conclude that progress during this period was too great to be accounted for by improvements in family life alone. They, and the others mentioned, direct attention to the likely role played by school improvements, including increased funding and class size reduction.

It is unlikely school resources can compensate *wholly* for the limited learning opportunities outside school that hold back many minority and low SES youth. Nevertheless, seasonal comparisons of learning make a compelling case that schooling indeed "makes a difference" in these children's lives, echoing the "differential sensitivity" hypothesis originally advanced in the Equality of Educational Opportunity report (Coleman et al. 1966). But how and when can interventions be most effectively targeted? What is the role of schools in educational stratification? And how should schools be held accountable when achievement scores persist in falling short of expectations? These are large issues, and the realization that much of the problem traces to out-of-school time during the early elementary years has implications for them all.

First, attempting to close the gap after it has opened wide is a rear guard action. Most of the gap increase happens early in elementary school, which is where corrective interventions would be most effective—or even before. To catch up, youth who have fallen behind academically need to make larger than average gains. That is expecting a great deal, perhaps too much, of struggling students. Early interventions to keep the achievement gap from opening wide in the first place should be a high priority, and the earlier the better, with the kinds of preschool compensatory education initiatives that have proven effective (e.g., Ramey, Campbell, and Blair 1998; Schweinhart and Weikart 1998; Reynolds and Temple 1998).

Second, once in school, disadvantaged children need year-round, supplemental programming to counter the continuing press of family and community conditions that hold them back. The school curriculum in the elementary years often is self-consciously pursued at home, as when, for example, parents work with their children on letter and number skills or reading. Parents of means generally did well in school themselves. They understand the skills and behaviors valued there and exemplify them in family life. For their part, poor parents often themselves struggled at school and have low literacy levels, and thus they undoubtedly have difficulties cultivating valued educational skills in their children. While low income, low SES parents generally want the same kinds of enriching experiences for their children as do well-off parents, they often lack the means to provide them (e.g., Chin and Phillips 2004).

Seasonal studies of learning suggest that schooling compensates, to some degree, for a lack of educationally enriching experiences in disadvantaged children's family life—these youth come closer to keeping up with better-off students during the school year than they do during the summer months. But if some school helps, does that mean more school is necessarily better? Summer and after-school programs are the most obvious approaches,[9] but what counts is how that extra learning time is used. Summer schools that incorporate so-called best practice principles have proven effective (Borman and Dowling 2006; Cooper et al. 2000), but to address the achievement gap specifically, programs will need to target dis-

[9] Some have also called for more far-reaching reform of the traditional school-year calendar to eliminate the long summer break (e.g., Cooper 2004; Gandara and Fish 1994; McCabe 2004).

advantaged students specifically. All children can benefit from high quality "universal" programs—preschools for all; summer schools for all—but they will not benefit in equal measure. Families of privilege will tend to find their way to higher quality programs, and their children will be positioned to profit more from programs of like quality (e.g., Cooper et al. 2000). As a result, rather than moderate the achievement gap, across-the-board programming for academic remediation and/or enrichment would likely exacerbate it, making the problem worse rather than better (Ceci and Papierno 2005). This poses a challenge to policy: what to do when two educational goals, each commendable, are in conflict?

Third, the school-year pattern of achievement gain parity (or near parity) across social lines flies in the face of widely held (if often only whispered) assumptions about the learning abilities of poor and minority youth. It also flies in the face of widely held assumptions about the failures of the public schools and school systems burdened by high poverty enrollments. Perhaps these schools and school systems are doing a better job than is generally recognized (e.g., Alexander 1997; Berliner and Biddle 1995; Krueger 1998), with family disadvantages mistaken for school failings (e.g., Rothstein 2002) and the occasional but very real horror story (e.g., Kozol 1991) overgeneralized.

Finally, a seasonal perspective on learning also has implications for school accountability. The No Child Left Behind (NCLB) standard of "adequate yearly progress" is intended to monitor school effectiveness based on annual achievement testing in grades 3 through 8. Schools that fail to meet local NCLB standards in math and reading for two consecutive years are designated "in need of improvement," with increasingly severe correctives required the longer they remain so designated. Certainly schools that chronically fall short need help; however, the punitive cast of NCLB may be misplaced. Indeed, annual assessments confound school-year and summer learning in unknown proportions, and schools that enroll mainly disadvantaged students will be held accountable not just for what happens to their pupils during the school year, but also for their students' summer learning, over which they have no control. If the BSS pattern is at all typical, this will show many schools in a poor light even when their students move ahead during the school year at a rate comparable to that of students in schools deemed to be performing adequately. An accountability system that monitors progress fall to spring, perhaps relative to an expected summer gain baseline (Downey, von Hippel, and Hughes 2005), would be more appropriate for gauging a school's effectiveness. The current arrangement is useful for identifying need, but little more, and certainly not for apportioning blame.

Karl L. Alexander *is John Dewey Professor of Sociology at the Johns Hopkins University. His interests center on problems of educational stratification that can be addressed via organizational, social psychological, and life course perspectives. With Doris Entwisle and Susan Dauber, he recently published a revised updated version of their 1994 evaluation of grade retention,* On the Success of Failure. *He is also working on several projects that examine the early adult transition through the lens of the Beginning School Study.*

Doris R. Entwisle *is Research Professor of Sociology at the Johns Hopkins University. Her main area of interest is the sociology of human development over the life course, with an emphasis on issues of inequality. Her current research concerns the ways in which members of the Beginning School Study panel make the transition to adulthood and how social and financial disadvantage early in life alters the life situations of adults in their early 20s. Her most recent book (with Karl L. Alexander and Linda S. Olson),* Children, Schools, and Inequality, *explores the ways in which social structure shapes the paths that children take through elementary school.*

Linda Steffel Olson *is an associate research scientist in the Department of Sociology at the Johns Hopkins University. She has been associated with the Beginning School Study for the past 19 years. Her interests center on the effects of social structure on schooling outcomes.*

Table A1. Definitions, Means, and Standard Deviations of Variables in Seasonal Gain Analysis

Variable	Mean	(SD)	N	Description
Background Characteristics				
Student Gender	.51	(.50)	787	1 = Female; 0 = Male
Student Race	.55	(.50)	787	1 = Black; 0 = White
Family Socioeconomic Status Composite	–.04	(.80)	787	Average of both parents' educational level, a ranking of mother's and father's occupational status (TSEI2, see Featherman and Stevens 1982), and participation in federal meal subsidy program (all measures converted to z-scores).
Mother's Years of Education	11.7	(2.6)	750	
Father's Years of Education	12.2	(2.7)	529	
Mother's TSEI2 Job Status	31.9	(15.8)	610	
Father's TSEI2 Job Status	33.3	(18.1)	518	
Proportion Meal Subsidy	.66	(.47)	754	
Family SES Composite Trichotomized				Composite variable trichotomized: Low SES = 0, Mid SES = 1, and High SES = 2. The family SES composite was trichotomized into low (50 percent of sample), medium (26 percent), and high (24 percent) SES categories in such a way that the mean educational level for parents was around 10 years for the low group, 12 years for the medium group, and around 15 years for the high SES group. The skew in the data toward the low SES group reflects the socioeconomic makeup of the Baltimore City student population.
Student Achievement, Years 1–9 (Imputed data)				
CAT Reading Score, Fall 1st Grade	279.8	(40.7)	787	California Achievement Test (CAT), Form C, Reading Comprehension, scale scores. BCPS tested students in the fall and spring of each year through year 6, and in spring only in years 7 and 8. In the spring of year 9, the BSS administered the CAT to students (testing continued over an 18 month period), then using exact testing dates, interpolated scores back to the spring of year 9. See Entwisle and colleagues (1997, Appendix), for data on psychometric properties of CAT. Seasonal components of reading achievement, years 1 to 5, were computed by subtracting fall from spring scores (winter gains) and spring from the following fall scores (summer gains). Yearly seasonal gains were then summed for a total Summer Gain and Winter Gain. Years 6 to 9 Gain was calculated by subtracting the spring year 5 score from the spring of year 9 score.
CAT Reading Score, Spring Year 9	547.6	(80.4)	787	
High School and Postsecondary Attainment				
Proportion College-Prep High School Curriculum	.30	(.46)	656	Data on high school curriculum come from BCPS school records. Students attending one of the magnet high schools with strong college-prep tracks were designated as "College Prep." Data on students from non-BCPC schools came from student responses to questions on the high school program.
Educational Attainment Age 22				
Proportion Permanent Dropout	.22	(.42)	630	Data from Young Adult Survey at age 22. Highest level of education completed or enrolled in by age 22. For postsecondary schooling, highest level attended whether or not completed the program.
Proportion Terminal High School	.29	(.45)	630	
Proportion Trade School/Two-Year College	.26	(.44)	630	
Proportion Four-Year College	.23	(.42)	630	

REFERENCES

Alexander, Karl L. 1997. "Public Schools and the Public Good." *Social Forces* 76:1–30.

Alexander, Karl L., Doris R. Entwisle, and Linda S. Olson. 2001. "Schools, Achievement and Inequality: A Seasonal Perspective." *Educational Evaluation and Policy Analysis* 23:171–91.

Allison, Paul D. 2002. *Missing Data.* Thousand Oaks, CA: Sage Publications.

Althauser, Robert P. and Michael Wigler. 1972. "Standardization and Component Analysis." *Sociological Methods and Research* 1:97–135.

Berliner, David. 2006. "Our Impoverished View of Educational Reform." *Teachers College Record* 108:949–95.

Berliner, David C. and Bruce J. Biddle. 1995. *The Manufactured Crisis: Myths, Fraud and the Attack on America's Public Schools.* Reading, MA: Addison-Wesley.

Borman, Geoffrey D. and N. M. Dowling. 2006. "The Longitudinal Achievement Effects of Multi-Year Summer School: Evidence from the Teach Baltimore Randomized Field Trial." *Educational Evaluation and Policy Analysis* 28:25–48.

Burkam, David T., Douglas D. Ready, Valerie E. Lee, and Laura F. LoGerfo. 2004. "Social-Class Differences in Summer Learning Between Kindergarten and First Grade: Model Specification and Estimation." *Sociology of Education* 77:1–31.

California Achievement Test. 1979. *Technical Bulletin 1, Forms C and D, Levels 10–19.* Monterey, CA: McGraw-Hill.

Ceci, Stephen J. and Paul B. Papierno. 2005. "The Rhetoric and Reality of Gap Closing: When the 'Have-Nots' Gain but the 'Haves' Gain Even More." *American Psychologist* 60:149–60.

Chin, Tiffani and Meredith Phillips. 2004. "Social Reproduction and Child-Rearing Practices: Social Class, Children's Agency, and the Summer Activity Gap." *Sociology of Education* 77:185–210.

Coleman, James S., Ernest Q. Campbell, Charles J. Hobson, James McPartland, Alexander Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity.* Washington, DC: U.S. Government Printing Office.

Cooper, Harris. 2004. "Is the School Calendar Dated? Education, Economics, and the Politics of Time." Pp. 3–23 in *Summer Learning, Research, Policies, and Programs,* edited by G. D. Borman and M. Boulay. Mahweh, NJ: Lawrence Erlbaum Associates.

Cooper, Harris, Kelly Charlton, Jeff C. Valentine, and Laura Muhlenbruck. 2000. *Making the Most of Summer School: A Meta-Analytic and Narrative Review.* Monograph Series for the Society for Research in Child Development, vol. 65, no. 1, Serial No. 260. Ann Arbor, MI: Society for Research in Child Development.

Cooper, Harris, Barbara Nye, Kelly Charlton, James Lindsay, and Scott Greathouse. 1996. "The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review." *Review of Educational Research* 66:227–68.

Dougherty, Kevin J. 1996. "Opportunity-to-Learn Standards: A Sociological Critique." *Sociology of Education* Extra issue:40–66.

Downey, Douglas B., Paul T. von Hippel, and Beckett Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review* 69:613–35.

Downey, Douglas B., Paul T. von Hippel, and Melanie Hughes. 2005. "Are 'Failing' Schools Really Failing? Using Seasonal Comparisons to Evaluate School Effectiveness." Presented at the American Sociological Association, Education Section, August, Montreal, Canada.

Entwisle, Doris R., Karl L. Alexander, and Linda S. Olson. 1997. *Children, Schools and Inequality.* Boulder, CO: Westview Press.

Farkas, George and Kurt Beron. 2004. "The Detailed Age Trajectory of Oral Vocabulary Knowledge: Differences by Class and Race." *Social Science Research* 33:464–97.

Featherman, David L. and Gillian Stevens. 1982. "A Revised Socioeconomic Index of Occupational Status: Application in Analysis of Sex Differences in Attainment." Pp. 141–82 in *Social Structure and Behavior: Essays in Honor of William Hamilton Sewell,* edited by R. M. Hauser, D. Mechanic, A. O. Haller, and T. Hauser. New York: Academic Press.

Gandara, P. C. and Judy Fish. 1994. "Year-Round Schooling as an Avenue to Major Structural Reform." *Educational Evaluation and Policy Analysis* 16:67–85.

Grissmer, David W., Shelia Nataraj Kirby, Mark Berends, and Stephanie Williamson. 1994. *Student Achievement and the Changing American Family.* Santa Monica, CA: The Rand Corporation.

Hart, Betty and Todd R. Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children.* Baltimore, MD: Paul H. Brookes Publishing Co.

Hauser, Robert M. 1995. "The Bell Curve." *Contemporary Sociology* 24:149–53.

Heyns, Barbara. 1978. *Summer Learning and the Effects of Schooling.* New York: Academic.

———. 1987. "Schooling and Cognitive Development: Is There a Season for Learning?" *Child Development* 58:1151–60.

Iams, Howard M. and Arland Thornton. 1975. "Decomposition of Differences: A Cautionary Note." *Sociological Methods and Research* 3:341–51.

Jones, F. L. and Jonathan Kelley. 1984. "Decomposing Differences Between Groups: A

Cautionary Note on Measuring Discrimination." *Sociological Methods and Research* 12:323–43.

Karweit, Nancy, Anne Ricciuti, and Bill Thompson. 1994. "Summer Learning Revisited: Achievement Profiles of Prospects' First Grade Cohort." Abt Associates, Washington, DC. Unpublished manuscript.

Kerckhoff, Alan C., Stephen W. Raudenbush, and Elizabeth Glennie. 2001. "Education, Cognitive Skill, and Labor Force Outcomes." *Sociology of Education* 74:1–24.

Kozol, Jonathan. 1991. *Savage Inequalities: Children in America's Schools.* New York: Crown Publishing.

Krueger, Alan B. 1998. "Reassessing the View that American Schools are Broken." *FRBNY Economic Policy Review* March:29–43.

Lee, Jaekyung. 2002. "Racial and Ethnic Achievement Gap Trends: Revising the Progress Toward Equity." *Educational Researcher* 31:3–12.

Lee, Valerie E. and David T. Burkam. 2002. *Inequality at the Starting Gate.* Washington, DC: Economic Policy Institute.

McCabe, Melissa. 2004. "Year-Round Schooling." *Education Week on the Web,* January 27. Education issues A–Z. Retrieved November 20, 2006 (http://www.edweek.org).

Murnane, Richard J. 1975. *The Impact of School Resources on the Learning of Inner City Children.* Cambridge, MA: Ballinger.

O'Brien, Daniel M. 1998. "Family and School Effects on the Cognitive Growth of Minority and Disadvantaged Elementary Students." Presented at the Association for Public Policy Analysis and Management, October 29–31, New York.

Phillips, Meredith. 2000. "Understanding Ethnic Differences in Academic Achievement: Empirical Lessons from National Data." Pp. 103–32 in *Analytic Issues in the Assessment of Student Achievement,* edited by U. S. Department of Education. Washington, DC: U. S. Department of Education, National Center for Education Statistics.

Phillips, Meredith, James Crouse, and John Ralph. 1998. "Does the Black-White Test Score Gap Widen After Children Enter School?" Pp. 229–72 in *The Black-White Test Score Gap,* edited by C. Jencks and M. Phillips. Washington, DC: Brookings.

Pope, Jean, Barry Lehrer, and James Stevens. 1980. "A Multiphasic Reading Screening Procedure." *Journal of Learning Disabilities* 13:98–102.

Ramey, Craig T., Frances A. Campbell, and Clancy Blair. 1998. "Enhancing the Life Course for High-Risk Children: Results from the Abecedarian Project." Pp. 163–83 in *Social Programs That Work,* edited by J. Crane. New York: Russell Sage Foundation.

Reardon, Sean F. 2003. *Sources of Inequality: The Growth of Racial/Ethnic and Socioeconomic Test Score Gaps in Kindergarten and First Grade.* Working Paper No. 03005, Population Research Institute, University Park, PA, Pennsylvania State University.

Reynolds, Arthur J. and Judy A. Temple. 1998. "Extended Early Childhood Intervention and School Achievement: Age Thirteen Findings from the Chicago Longitudinal Study." *Child Development* 69:231–46.

Rothstein, Richard. 2002. *Classes and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap.* Washington, DC: Economic Policy Institute.

Royston, Patrick. 2004. "Multiple Imputation of Missing Values." *Stata Journal* 4:227–41.

———. 2005. "Multiple Imputation of Missing Values: Update." *Stata Journal* 5:88–102.

Schweinhart, Lawrence J. and David P. Weikart. 1998. "High/Scope Perry Preschool Program Effects at Age Twenty-Seven." Pp. 148–83 in *Social Programs That Work,* edited by J. Crane. New York: Russell Sage.

Weller, L. David, Carl J. Schnittjer, and Bertha A. Tuten. 1992. "Predicting Achievement in Grades Three Through Ten Using the Metropolitan Readiness Test." *Journal of Research in Childhood Education* 6:121–29.

West, Jerry, Kristin Denton, and Lizabeth M. Reaney. 2001. *The Kindergarten Year: Findings from the Early Childhood Longitudinal Study.* Technical Report No. NCES 2001-023. Washington, DC: National Center for Educational Statistics.