The Estimation of Summer Learning Rates

Michael S. Hayes
Department of Public Policy and Administration
Rutgers University – Camden

Seth Gershenson[♦]
School of Public Affairs
American University

Abstract

The activities and experiences that contribute to summer learning are potentially important inputs to the education production function, yet are relatively understudied by economists. We propose a method for identifying heterogeneity in summer learning rates when tests are not given on the first and last days of the school year and apply this method to data from the nationally representative Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K). Generally, we find evidence of heterogeneity in summer learning that varies between gain-score and lag-score models of the education production function, and between math and reading achievement. Consistent with previous research, students from low-income households make significantly lower summer reading gains than children from wealthier households. However, we also find evidence of differential rates of summer math development by baseline academic ability, private school attendance, and summer school attendance.

JEL Codes: I21, D13

Keywords: summer learning loss, summer setback, education production function, achievement gap

[♦] Corresponding author – Comments are welcome. Email: gershens@american.edu. Phone: 202-885-2687. Fax: 202-885-2347. American University, School of Public Affairs, 4400 Massachusetts Avenue, NW, Washington DC, 20016-8070. The authors thank attendees of the 2014 AEFP Annual Conference for providing helpful comments. Any remaining errors are our own.

1. Introduction

Improving the quality of publicly provided education, particularly that of socioeconomically disadvantaged students, is a primary goal of state and federal education policy in the U.S., as educational achievement and attainment are thought to improve individual labor-market outcomes and to facilitate intergenerational socioeconomic mobility more generally (e.g., Card & Krueger, 1992; Checchi, 2006; Ellwood & Kane, 2000). To effectively close achievement gaps between students from different socioeconomic and demographic backgrounds, then, it is important that policy makers and educators are aware of the determinants of academic success and the factors that contribute to such achievement gaps.

The activities, individuals, and environments to which children are exposed during summer vacation comprise one potentially important class of inputs in the education production function. Indeed, educational researchers have been interested in the potential detrimental effects of summer vacation on cognitive development for more than a century (Cooper et al., 1996). More recently, the seminal work of Heyns (1978) put forth and tested the hypothesis that higher rates of summer learning loss (SLL) among disadvantaged students might contribute to the stubborn persistence of achievement gaps between students of different demographic and socioeconomic backgrounds.[1] It is therefore important that policy makers and educators understand the causes, consequences, and magnitude of SLL and how SLL varies by students' demographic and socioeconomic backgrounds.

Heyns' (1978) empirical analysis of SLL among sixth and seventh graders in Atlanta spawned a small literature in the sociology of education that empirically investigates differences in SLL by gender, race and ethnicity, and socioeconomic status (SES) (e.g., Burkam et al., 2004;

[1] Heterogeneous summer learning rates have been referred to as summer learning loss, summer setback, and summer slide in the sociology and education literatures.

Downey, von Hippel, & Broh, 2004; Entwisle & Alexander, 1992; Alexander, Entwisle, & Olson, 2001; Quinn, 2014).[2] However, SLL has largely been ignored by labor and education economists and the existing empirical SLL literature yields some contradictory results, perhaps due to variation in the empirical strategies employed in existing studies. The labor and education economists who have broached the topic of SLL have largely done so tangentially, as part of either sensitivity analyses (e.g., Fitzpatrick, Grissmer, & Hastedt, 2011) or broader studies (e.g., Fryer & Levitt, 2004). The current study contributes to this gap in the economics of education literature by presenting a formal approach to estimating summer learning rates, testing for heterogeneity in summer learning rates, and reconciling the sometimes contradictory results in the extant literature.

Specifically, we develop an econometric approach to estimating SLL and then apply the method to nationally representative data from the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K). A variety of specifications and modeling decisions are then considered, in an effort to reconcile our results with those in the existing literature and identify the modeling decisions to which results are sensitive. The paper proceeds as follows: Section 2 summarizes the extant empirical SLL literature. Section 3 describes the ECLS-K data and section 4 develops an econometric model of SLL. Section 5 presents the main results. Section 6 concludes with a discussion of policy implications and directions for future research.

2. Previous Research on Summer Learning Loss (SLL)

Education researchers have long been interested in the impact of summer vacation on academic progress, dating at least to White's (1906) analysis of seven college students' math

[2] Cooper et al. (1996) thoroughly review this literature alongside earlier studies that did not allow for heterogeneity in SLL.

competencies before and after the summer break. Cooper et al. (1996) reviewed 26 studies of SLL that were conducted before 1975, only one of which tested for differences by SES (Hayes & Grether, 1969). Cooper et al. (1996) also reviewed 13 SLL studies conducted between 1975 and 1995 more rigorously and reached four main conclusions.[3] First, the overall average amount of SLL is between 0 and 0.1 test score standard deviations (SD). Second, overall average SLL is close to 0 SD in reading but slightly larger in math. Third, the average amount of math SLL is uniform across SES groups, but reading SLL is greater among low-SES students while high-SES students experience no SLL or even experience summer gains in reading/literacy. Fourth, the overall average amount of SLL is uniform across demographic groups (i.e., race and gender).

The findings of the meta-analysis are not universally accepted, however, as the studies themselves yield some conflicting results. Importantly, the studies covered different summer vacation lengths, which ranged from 92 to 153 days between spring and fall tests, and relied on different samples, testing instruments, time periods, age groups, etc. One concern with these early studies, particularly with the influential studies of SLL in Atlanta and Baltimore by Heyns (1978) and Entwisle and Alexander (1992), respectively, is that the results may not generalize to non-urban schools or urban schools in other parts of the country that have larger Hispanic enrollments (Burkam et al., 2004). In response to this critique several studies have investigated the phenomenon of SLL in the nationally representative ECLS-K. The methods, focus, and conclusions of seven such studies are summarized in table 1.

It is immediately obvious from table 1 that these studies vary widely in estimation method, model specification, focus, and findings. For example, half of the studies used ECLS-K sampling weights and half did not; gain-score, lag-score, and growth models were estimated; the

[3] The meta-analytic sample of Cooper et al. (1996) included the influential studies by Heyns (1978) and Entwisle and Alexander (1992) of Atlanta and Baltimore, respectively.

studies conditioned on different covariates and tested for heterogeneity along different dimensions; studies adjusted for variation in test dates in different ways while one study made no such adjustment; and only two of the seven studies compared school-year and summer learning rates using the same analytic sample of students. While many of these differences legitimately owe to differences in the research question being addressed, the general lack of (reported) systematic sensitivity analyses complicates cross-study comparisons. We contribute to this literature by estimating both school-year and summer learning rates using a single analytic sample and a variety of specifications of the education production function.

3. Data

Data on summer learning, household characteristics, and summer activities are taken from the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K), which was collected by the National Center for Education Statistics (NCES). The full sample of more than 20,000 children from about 1,000 kindergarten programs (i.e., schools) was designed to be nationally representative of the cohort that entered kindergarten in the 1998-99 academic year. Certain subgroups of the population were oversampled, so all analyses are conducted using NCES-provided sampling weights that adjust for the survey's nonrandom sampling frame.

All children in the ECLS-K sample were surveyed in the fall and spring of kindergarten and the spring of first grade. However, the analytic sample is restricted to the 30% random subsample of children who were also surveyed in fall of first grade. This facilitates the estimation of learning that occurred between the spring kindergarten assessment and the fall first grade assessment (i.e., during the summer between kindergarten and first grade). We further restrict the analytic sample by excluding students who experienced a mid-year classroom change, repeated kindergarten, changed schools between kindergarten and first grade, or were missing basic demographic or test score data. School changers are excluded to avoid conflating SLL with shocks to achievement caused by the disruption associated with changing schools, though it is worth noting that including school changers in the analytic sample and conditioning on a binary "school change" indicator yields qualitatively similar results.

The ECLS-K data are well suited for an empirical analysis of SLL for three reasons. First, the ECLS-K is the only nationally representative survey of U.S. students that contains both fall and spring test scores that span the summer vacation. Second, the ECLS-K contains data on students' summer activities, which facilitates analyses of the behaviors and activities that contribute to SLL. These variables are described in detail below. Third, the fall and spring ECLS-K assessments used to calculate SLL covered the same content and had the same (low) stakes, so teachers had no incentive to strategically divert resources or instructional time towards a specific test (Fitzpatrick et al., 2011).

The ECLS-K administered age-appropriate reading and mathematics tests during each wave of the survey. The math examinations tested children's abilities in the following areas: numbers and shapes, relative size, ordinality and sequence, addition and subtraction, and multiplication and division. The reading examinations tested children on letter recognition, beginning sounds, ending sounds, sight words, and words in context. Students did not take identical exams, as the achievement tests used a two-stage assessment approach. Accordingly, the ECLS-K used Item Response Theory (NCES, 2002) to compute scaled test scores using the full set of test items. In all subsequent analyses test scores are standardized by subject and testing period to have mean zero and SD one using all available test scores, as suggested by Ballou (2009). However, alternative scalings (e.g., Quinn, 2014) yield qualitatively similar results.

Importantly, in both the fall and spring semesters, ECLS-K assessments were administered to different students on different days. Differences in test dates across schools, across classrooms within schools, and even across students within classrooms are common in the data, as a relatively small number of ECLS-K administrators individually met with each student to perform the assessment. Fitzpatrick et al. (2011) show that test dates were essentially random and exploit this exogenous variation to estimate the causal effect of instructional days on academic achievement. We propose a similar approach to estimating SLL in section 4.

To avoid conflating SLL with school-year learning that occurred either before the fall first-grade assessment or after the spring kindergarten assessment, the econometric model must acknowledge that assessments were administered on neither the first nor last days of the academic year. While we postpone formal discussion of our test-date adjustment strategy to section 4, note that there are four relevant dates:

A. Spring Kindergarten Assessment Date

B. Kindergarten End Date

C. First Grade Start Date

D. Fall First Grade Assessment Date

Dates *A*, *C*, and *D* are formally reported in the ECLS-K. Unfortunately, the precise date *B* is unobserved. However, the ECLS-K does report the first grade end date, which we use to impute *B*, as the analytic sample is restricted to students who attended the same school for both kindergarten and first grade. While this solution is imperfect, it is unlikely to compromise the econometric analysis, as within-school changes in end dates from one year to the next are likely to be small and exogenously determined by factors such as inclement weather (i.e., snow days added to the end of the year) and scheduling quirks. Indeed, several authors have estimated the effect of instructional days on student achievement by exploiting the exogenous variation in school cancellations associated with inclement weather (e.g., Marcotte & Hansen, 2010).

The number of days between dates *A*, *B*, *C*, and *D*, as well as achievement gains made during kindergarten, first-grade, and the intervening summer are summarized in table 2. Importantly, table 2 shows that unadjusted estimates of SLL that fail to acknowledge that tests were not administered on the last day of kindergarten or the first day of first grade are negative, while modest achievement gains are made during both kindergarten and first grade. The estimated standard deviations of unadjusted summer and school-year learning are sizable and similar in all three time periods, indicating that there is nontrivial variation in both school-year and summer learning. However, these "unadjusted" SLL estimates are potentially misleading, as nearly half of the days between the spring-kindergarten and fall-first grade tests were actually school days. The average summer vacation was about 80 days. Of the 70 school days that transpired between the two tests, about 60% were at the start of first grade before the fall first-grade test and about 40% were at the end of kindergarten after the spring kindergarten test.

Table 2 also summarizes the demographic composition and summer activities of students in the analytic sample. The analytic sample is approximately 74% non-Hispanic white, 11% black, and 10% Hispanic. The remaining 6% is classified as "other race," which includes Asians, Pacific Islanders, Native Americans, and students of mixed race. About 13% of students resided in households below the poverty line. The sample contains equal proportions of males and females. Students with an Individualized Education Plan (IEP) and who do not speak English at home each comprise about 4% of the analytic sample. Over half of students participated in an organized summer activity and about one quarter of students attended summer camp. About 80% of students practiced at least some math over the summer, while nearly all parents read with children at some point during the summer vacation. 16% of students in the analytic sample attended private school and 44% attended a suburban school. About one third of students attended an urban school while the remaining 25% of students attended a rural school.

4. Econometric Model & Estimation

4.1 Estimating Summer Learning Rates

Section 3 makes clear that the timing of the fall and spring ECLS-K assessments complicates the estimation of summer learning rates. Let y^j represent achievement at date j for j = A, B, C, and D. Only y^D and y^A are observed, though the difference between observed test scores can conceptually be decomposed as follows:

$$y^D - y^A = (y^D - y^C) + (y^C - y^B) + (y^B - y^A). \tag{1}$$

The middle term on the RHS of equation (1) constitutes learning that occurs during the summer vacation, which is the object of interest in the current study. We exploit two facts in order to estimate summer learning rates. First, the date (d) is observed for each j. Second, cognitive development in early childhood is a cumulative process that occurs systematically over time (McCoach et al., 2006; Muthen et al., 2003). Accordingly, the RHS of (1) can be approximated by the sum of three general functions of time:

$$y^D - y^A = f(d^D - d^C) + g(d^C - d^B) + h(d^B - d^A) + \varepsilon, \tag{2}$$

where ε is an error term that acknowledges that the functions on the RHS of equation (2) are approximations of the true learning that occurred between dates j and j+1 for each j. Student (i) and school (s) subscripts on the y^j, d^j, and ε in equation (2) are temporarily suppressed. Equation (2) is similar to the main estimating equation in Burkam et al. (2004).

Equation (2) can be estimated by OLS after assuming tractable functional forms of *f*, *g*, and *h*, which could be nonlinear (McCoach et al., 2006).[4] However, we begin our analysis by assuming that *f*, *g*, and *h* are linear in $(d^{j+1} - d^{j})$, as Fitzpatrick et al. (2011) find school-year learning rates in the ECLS-K to be approximately linear. We do allow for different slopes in each of the three time periods, otherwise the RHS of equation (2) would simplify to $f(d^{D} - d^{A}) + \varepsilon$. Moreover, higher-order polynomials and RESET specification tests, which are reported in appendix A, confirm that *f*, *g*, and *h* are approximately linear in calendar days.

The derivative of *g* with respect to $(d^{C} - d^{B})$, which is a scalar when *g* is linear, is the daily rate of summer learning. Whether this parameter can be given a causal interpretation depends primarily on whether or not the ε in equation (2) is correlated with summer vacation length (i.e., $d^{C} - d^{B}$), as Fitzpatrick et al. (2011) have shown that ECLS-K assessment dates d^{D} and d^{A} are essentially random. For example, it could be that parents select into school districts based on academic calendars or that summer vacation length is determined by school resources. While we cannot directly test for this type of endogeneity, we can test for systematic differences in summer vacation length by observable household and school characteristics in two ways. The results of these tests are presented in section 5 and generally provide no evidence of systematic differences in summer vacation lengths based on observables. Nonetheless, regardless of whether these estimates are given a causal interpretation, a contribution of the current study is to provide an accurate description of the distribution of average summer learning rates.

The first test involves regressing the natural log of summer vacation length on a variety of observed student and school characteristics. If summer vacation length is orthogonal to observed student and school characteristics, it is plausible that vacation length is similarly

[4] Alternatively, equation (2) could be estimated non-parametrically (e.g., Eren & Henderson, 2008). Such an analysis is left to future work, as it is beyond the scope of the current study.

orthogonal to unobserved student and school characteristics (Altonji, Elder, & Taber, 2005). The second test amounts to examining the sensitivity of estimates of equation (2) to conditioning on observed student, household, and school characteristics. If conditioning on observables associated with academic achievement significantly changes the estimated coefficient on ($d^C - d^B$), the estimated coefficient on ($d^C - d^B$) is unlikely to have a causal interpretation.[5]

An added benefit of the second exercise is that in addition to potentially increasing the precision of summer learning rate estimates, the coefficients on the controls are interesting in their own right, as they provide suggestive evidence of the sources of heterogeneity in SLL. This evidence is only suggestive, however, as the model cannot distinguish between the contributions of these covariates to achievement gains made during school days and summer days between dates *A* and *D*. Specifically, these coefficients represent intercept shifts (different starting points) and not different slopes (different learning rates). A precise method for estimating heterogeneity in summer learning rates is provided in section 4.2.

4.2 Heterogeneity in Summer Learning Rates

Much of the recent empirical SLL literature has focused on testing for differences in SLL by SES and other observable student and school characteristics. Borman, Benson, and Overman (2005) and Gershenson (2013) provide theoretical discussions of the mechanisms through which SES and summer activities might affect summer learning rates. It is straightforward to model heterogeneity in summer learning rates by generalizing the model presented in equation (2) to allow for a student-specific function *g*. The preferred linear specification then becomes

[5] Note that we cannot condition on school fixed effects because summer vacation length is observed at the school level. Such an analysis would be possible if data were available for two intervening summers.

$$y_{is}^D - y_{is}^A = \lambda\left(d_i^D - d_s^C\right) + \beta_{is}\left(d_s^C - d_s^B\right) + \delta\left(d_s^B - d_i^A\right) + \varepsilon_{is}. \qquad (3)$$

To allow for heterogeneity by observed student and school characteristics, we assume that

$$\beta_{is} = \beta_0 + \beta_1\mathbf{x}_i + \beta_2\mathbf{z}_s \text{ and } \varepsilon_{is} = \gamma_0 + \gamma_1\mathbf{x}_i + \gamma_2\mathbf{z}_s + u_{is}, \qquad (4)$$

where **x** and **z** are vectors of observed student and school characteristics, respectively, and u is an idiosyncratic error term. The coefficients on spring of kindergarten learning (δ) and fall of first grade learning (λ) can be similarly generalized.[6] The model characterized by equations (3) and (4) is similar to one of the multilevel models estimated by Downey et al. (2004).

Importantly, **x** could include lagged achievement (y^A). Including y^A as an additively separable control devolves (2) into a familiar lag-score value-added model (e.g., Sass et al., 2014).[7] Unlike analyses of the education production function in which the goal is to estimate unbiased effects of educational inputs or interventions on achievement (e.g., Todd & Wolpin, 2003), viewed as a descriptive analysis of SLL, the choice between gain-score and lag-score specifications in the current study depends on the descriptive research question of interest (Quinn, 2014; Rubin et al., 2004). Specifically, the gain-score specification (excluding y^A from the RHS of equation 2) yields estimates of "unconditional" summer learning rates, while the lag-score specification (conditioning on y^A) yields estimates of summer learning rates conditional on prior (spring of kindergarten) achievement; that gain-score and lag-score estimates are typically not equivalent is known as Lord's Paradox (Lord, 1967; Quinn, 2014).

Summer learning rates might vary with past achievement for at least two reasons. If it does, y^A belongs in the vector **x** in equation (4). For example, there might be "Matthew Effects" whereby high-achieving students continually learn at faster rates than their lower-achieving

[6] The results yield qualitatively similar results if spring of kindergarten learning (δ) and fall of first grade learning (λ) are also allowed to vary by observed student characteristics.
[7] For example, if $y^D - y^A = \beta y^A$ and $y^D = \alpha y^A$, then $\alpha = \beta + 1$.

peers.[8] Alternatively, convergence in students' test scores might come about due to previously low-achieving students "catching up" by learning at relatively faster rates. Which, if either, of these learning patterns occur is an empirical question addressed in section 5.

5. Results

5.1 Determinants of Summer Vacation Length

Table 3 reports estimates of the summer vacation length regressions described in section 4.1, where the dependent variable is the natural log of summer vacation length. When included individually, the vectors of student characteristics, summer activities, and school characteristics are jointly statistically insignificant at traditional confidence levels, as shown in columns 1 – 3, respectively. The specification reported in column 4, which conditions on all three vectors of covariates, similarly finds that none of the three sets of covariates are jointly statistically significant. Across all four specifications estimated in table 3, only 4 of 34 covariates are even marginally individually statistically significant. Taken together, these results suggest that summer vacation lengths ($d^C - d^B$) are not correlated with observed student or school characteristics and, in the spirit of Altonji et al. (2005), are therefore unlikely to be correlated with unobserved student and school characteristics.

5.2 Unconditional Average Daily Learning Rates

Table 4 reports unconditional OLS estimates of average daily learning rates during three time periods: between kindergarten tests, between first-grade tests, and between spring of kindergarten and fall of first grade tests ($d^D - d^A$). Columns 1 – 4 report estimates for math

[8]The "Matthew effect" occurs when early gains in reading skills lead to future gains in reading skills and gains in other subjects (Stanovich, 1986).

achievement. Columns 1 and 2 report daily school-year learning rates for kindergarten and first grade, respectively, that are similar to the models estimated by Fitzpatrick et al. (2011). Interestingly, column 3 shows that the average math learning rate between dates *A* and *D* during that time is similar in magnitude to the average math learning rate in first grade, despite the fact that almost half of the days between dates *A* and *D* are during the summer vacation. This challenges the general finding of SLL in math, as almost half of the intervening days are actually school days at either the end of kindergarten or start of first grade. Column 4 probes further by estimating the baseline specification presented in equation (2) and finds the learning rate during summer vacation ($d^C - d^B$) to be statistically indistinguishable from zero while both the end of kindergarten ($d^B - d^A$) and the start of first grade ($d^D - d^C$) learning rates are positive and significant. However, the summer learning rate is imprecisely estimated and not significantly different from either the kindergarten or first-grade learning rates. Like the results presented in column 3, these estimates provide only limited support for the existence of SLL in math.

Columns 5 – 8 of table 4 report estimates of the same four specifications for reading achievement. Once again, the average daily reading learning rate between dates *A* and *D* reported in column 7 is similar in magnitude to the school-year learning rates reported in columns 5 and 6. Interestingly, the estimates of the baseline specification reported in column 8 suggest that average summer reading gains are significantly greater than kindergarten and first grade gains.

Table 5 reports unadjusted lag-score estimates of average daily learning rates during the same three time periods: kindergarten, first grade, and between spring of kindergarten and fall of first grade ($d^D - d^A$). The specifications estimated in table 5 condition on lagged (spring-kindergarten) achievement, but are otherwise identical to the baseline gain-score models estimated in table 4. Columns 1 – 4 of table 5 report estimates for math achievement. Columns 1 and 2 report daily school-year math learning rates for kindergarten and first grade, respectively. These estimates are similar in magnitude to the average math learning rates generated by the gain-score specifications estimated in table 4. Like in the gain-score analyses, column 3 of table 5 shows that the average learning rate between dates *A* and *D* is similar in magnitude to the school-year learning rates presented in columns 1 and 2, which challenges the general finding of math SLL.

Column 4 of table 5 probes further by estimating the baseline specification presented in equation (2) and finds the learning rate during summer vacation ($d^C – d^B$) to be statistically indistinguishable from zero while the start of first grade ($d^D – d^C$) learning rate is relatively large, strongly significant, and significantly different from both the summer and end-of-kindergarten learning rates. Again, this highlights the importance of adequately modeling school start, school end, and test dates in analyses of SLL.

Columns 5 – 8 of table 5 contain estimates of the same four lag-score specifications for reading achievement. The daily learning rate between dates *A* and *D* is similar in magnitude to the first-grade learning rate, but only half as large as the kindergarten learning rate. Interestingly, estimates of the baseline specification reported in column 8 suggest that the average summer learning rate is positive and larger than learning rates during both the end of kindergarten and the start of first grade, though neither difference is statistically different from zero at traditional confidence levels. Overall, the lag-score estimates reported in table 5 are qualitatively similar to the analogous gain-score estimates reported in table 4.

5.3 Heterogeneity in Average Daily Learning Rates

Table 6 re-estimates gain-score and lag-score versions of the baseline specification (equation 2), this time conditioning on a rich set of observed student, household, and school characteristics. Columns 1 and 2 of table 6 are comparable to the estimates reported in column 4 of tables 4 and 5, respectively. Columns 3 and 4 of table 6 are comparable to the estimates reported in column 8 of tables 4 and 5. For math and reading achievement in gain-score and lag-score specifications, conditioning on these covariates does not appreciably change the estimated average learning rates, which are remarkably similar to those presented in tables 4 and 5. As discussed in section 4.1, the robustness of these estimates to conditioning on observed student and school covariates provides further evidence that summer vacation lengths are exogenous.

The estimated coefficients on these controls are interesting as well, as they indicate differences by observable characteristics in the achievement gains made between dates *A* and *D*. None of the covariates are statistically significant at traditional confidence levels in the gain-score math regressions. The lag-score math regressions, however, suggest that average learning gains during this time period were about 15% of a test score SD lower for black students than white students and that the children of college-educated mothers outperformed the children of high school graduates by about 14% of a test score SD. These differences are practically significant as well: they are approximately equivalent to the effect of a one SD increase in teacher effectiveness (e.g., Hanushek & Rivkin, 2010). Qualitative differences between the gain-score and lag-score estimates reinforce the importance of specifying the research question and the potential for these models to produce qualitatively different results (Quinn, 2014).

Only the poverty indicator is statistically significant in the reading regressions. This result is consistent with much of the existing SLL literature, and is generally interpreted as low-

SES children experiencing lower rates of summer learning than their more advantaged peers (e.g., Burkam et al., 2004). However, this interpretation is incorrect, as the effect of poverty on achievement in these models is not restricted to summer learning (i.e., gains made between dates *B* and *C*). The same is true for the black and college-educated indicators in the math regressions. Rather, the coefficients on the covariates reported in table 6 capture average difference in achievement gains between observably different students, holding the model's other covariates constant (including the length of summer vacation ($d^C - d^B$), time between the spring kindergarten test and end of kindergarten ($d^B - d^A$), and time between start of first grade and the fall first grade test ($d^D - d^C$)). Importantly, neither specification can determine whether differences by poverty status in reading achievement gains between dates *A* and *D* occurred during summer vacation, during school days after the spring kindergarten test, during school days before the fall first grade tests, or some combination of the three. Instead, as discussed in section 4.2, heterogeneity in average summer learning rates must be identified by estimating interaction models of the form described by equations (3) and (4).

Table 7 reports estimates of gain-score and lag-score versions of the interaction model described by equations (3) and (4) for both math and reading. The math results presented in columns 1 and 2 provide evidence of three sources of heterogeneity in summer math learning rates. First, the lag-score model in column 2 provides evidence of a Matthew Effect: students with high baseline (spring kindergarten) math scores learn math at significantly higher rates during the summer vacation than students who enter the summer with lower baseline math scores. For example, the effect of an additional ten days of summer vacation for a student who scored one SD higher on the spring-kindergarten math test than an observably similar student would be about 8% of a SD larger than for his or her classmate who experienced the same increase in summer vacation. This is a practically significant effect, as 0.08 of a test score SD constitutes about one half of the effect of a one SD increase in teacher effectiveness (e.g., Hanushek & Rivkin, 2010).

An even larger difference in summer learning rates is observed between students who did and did not attend summer school in the lag-score model reported in column 2. Unfortunately, the ECLS-K does not articulate the type of summer school program or the student's reason for attending, which might explain why the difference is only statistically significant in the lag-score model. Finally, both the gain-score and lag-score models show that private school students experience significantly higher rates of math summer learning than their counterparts in traditional public schools. Interestingly, this effect is net of the effect of mothers' education and the ECLS-K's coarse summer activity measures, suggesting that parents who select into private schools are providing stimulating summer environments and activities for their children that are not captured by the ECLS-K's information on household SES and summer activities.

Columns 3 and 4 of table 7 report estimates of the interaction model described by equations (3) and (4) for reading achievement. Consistent with the existing SLL literature, impoverished students experience SLL in reading, relative to more advantaged students, though this difference is only marginally statistically significant. Interestingly, however, the children of mothers who did not complete high school make significantly larger summer reading gains than the children of more educated mothers. We can only speculate as to the source of this counterintuitive result, but in terms of magnitude it nets out the poverty penalty. Similarly, IEP students also tend to make larger summer reading gains than mainstream students, perhaps because students with disabilities benefit from more personalized attention or other non-school environments during the summer vacation. It would be useful for future research to further investigate the summer activities and summer learning of such students.

Together, the results presented in table 7 have at least two implications for the study of SLL. First, the choice between gain-score and lag-score models of the education production function can yield different conclusions regarding the sources and extent of heterogeneity in the rates at which students learn during the summer vacation. As a result, researchers and program evaluators should let the specific research question of interest guide their choice of model. Second, these results highlight the importance of correctly modeling potential heterogeneities in summer learning rates and challenge the conventional wisdom that only summer reading gains are heterogeneous.

6. Conclusion

Summer learning loss (SLL) is a potentially important source of achievement gaps that has implications for education policy regarding school calendars and summer programming, yet is relatively understudied by labor and education economists. This paper develops and presents a formal approach to estimating heterogeneity in SLL and applies the method to nationally representative ECLS-K data. Importantly, we argue that much previous SLL research has incorrectly interpreted and estimated heterogeneity in SLL, and present an alternative method for doing so. Overall, the results suggest that SLL is not a pervasive problem. However, several statistically and practically significant sources of heterogeneity in students' summer learning rates in both math and reading are identified—particularly in low-income students' summer reading gains. Moreover, the results are somewhat sensitive to the choice between gain-score and lag-score models of the education production function, as expected given Lord's Paradox, highlighting the need for thoughtful consideration of the research question of interest (Quinn, 2014). An equally interesting non-finding is that the summer activity measures included in the ECLS-K have very little predictive power of summer learning rates. This could be due to the relative coarseness of these survey instruments, heterogeneity in the quality of the summer activities, or the true unimportance of these particular summer activities. It would be useful for future research to replicate these analyses using different data that have richer, more detailed measures of summer activities.

The findings that summer math gains are higher for at least some students who attend summer school and that summer reading gains are lower for low-income students have implications for education policy and summer programming. For example, the reading results provide support for subsidizing access to effective summer programs for children in low-income households, schools, or neighborhoods, as experimental evidence suggests that the six-week Building Educated Leaders for Life (BELL) program increased the reading skills of low-performing, low-income students in New York City and Boston by the equivalent of about 1 month of schooling (Chaplin & Capizzano, 2006). Experimental analyses of summer reading programs also generally find that low-income students benefit from well-designed summer reading interventions (e.g., Kim & Quinn, 2013). Similarly, it would be useful for future research to further investigate the circumstances in which students benefit from summer school.

References

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. *Educational Evaluation and Policy Analysis*, *23*(2), 171-191.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy, 113*(1), 151-184.

Ballou, D. (2009). Test Scaling and Value-added Measurement. *Education Finance & Policy, 4*(4), 351-383.

Borman, G. D., Benson, J., & Overman, L. T. (2005). Families, schools, and summer learning. The Elementary School Journal, 106(2), 131–150.

Burkam, D. T., Ready, D. D., Lee, V. E., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education*, *77*(1), 1-31.

Carbonaro, W. (2003). Sector Differences in Student Learning: Differences in Achievement Gains Across School Years and During the Summer. *Catholic Education: A Journal of Inquiry and Practice*, *7*(2).

Card, D., & Krueger, A. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100(1), 1-40.

Chaplin, D., & Capizzano, J. (2006). Impacts of a Summer Learning Program: A Random Assignment Study of Building Educated Leaders for Life (BELL). Retrieved March 10, 2014 from http:///www.urban.org/publications/411350.html.

Checchi, D. (2006). *The Economics of Education: Human Capital, Family Background, and Inequality*. New York, New York: Cambridge University Press.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. Review of Education Research, 66(3), 227-268.

Downey, D. B., Von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, *69*(5), 613-635.

Ellwood, D. T., & Kane, T. J. (2000). Who is Getting a College Education? Family Background and the Growing Gaps in Enrollment In S. Danziger & J. Waldfogel (Eds.), *Securing the Future: Investing in Children from Birth to College* (pp. 283-324). New York: Ford Foundation.

Entwisle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review*.

Eren, O., & Henderson, D. J. (2008). The impact of homework on student achievement. *The Econometrics Journal*, *11*(2), 326-348.

Fitzpatrick, M. D., Grissmer, D., & Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review*, *30*(2), 269-279.

Fryer Jr, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.

Gershenson, S. (2013). Do summer time-use gaps vary by socioeconomic status? *American Educational Research Journal*, 50(6): 1219-1248.

Hanushek, E. A., and S. G. Rivkin. 2010. Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100: 267-271.

Hayes, D. P., & Grether, J. (1969). The school year and vacations: When do students learn? Paper presented at the meeting of the Eastern Sociological Association, New York.

Heyns, B. (1978). *Summer learning and the effects of schooling* (pp. 227-268). New York: Academic Press.

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, *83*(3), 386-431.

Lord, F.M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, *68*, 304-305.

Marcotte, D. E., & Hansen, B. (2010). Time for school. *Education Next*, *10*(1), 52-59.

McCoach, D. B., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, *98*(1), 14.

Muthen, B., Khoo, S., Francis, D. J., & Boscardin, C. K. (2003). Analysis of reading skills development from kindergarten through first grade: An application of growth mixture modeling to sequential processes. In N. Duan & S. Reise (Eds.), Multilevel modeling: Methodological issues and applications (pp. 71–89). Mahwah, NJ: Erlbaum.

NCES. 2002. User's Guide to the Kindergarten–First Grade Public Use Data File, NCES-2002-149.

Quinn, D. (2014). Black-white summer learning gaps: Interpreting the variability of estimates across representations. Forthcoming, *Educational Evaluation and Policy Analysis*. doi: 10.3102/0162373714534522

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of educational and behavioral statistics*, *29*, 103-116.

Sass, T. R., Semykina, A., & Harris, D. N. (2014). Value-added models and the measurement of teacher productivity. *Economics of Education Review*, *38*, 9-23.

Stanovich, K. E. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407.

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), F3-F33.

White, W. (1906). Reviews before and after vacation. *American Education*, *10*, 185-188.

Table 1: Previous Analyses of Summer Learning in the ECLS-K

	Weights	Specification & Estimation	Controls	Test Date Adjustment	Same Sample	Covariates of Interest	Results
Burkam et al. (2004)	Yes	OLS estimation of lag score specifications	Yes	Controls for months between: spring-K test and end of K, end of K and start of first grade, and start of first grade and fall-first grade test	Yes	SES & Summer activities	Small SES differences in SL; Negligible effect of summer activities on SL
Carbonaro (2003)	Yes	OLS estimation of gain and lag score specifications	Sometimes	Weeks between tests (sometimes)	No	School type (public, private, Catholic)	Heterogeneity in unadjusted summer reading gains; No differences in summer math or adjusted reading gains
Downey et al. (2004)	No	Multilevel Growth Curve Model; No gains or lags	Sometimes	Days between each test and nearest start/end date of school year	No	Gender, SES, race, unexplained variation	Most SL is unexplainable. Faster learning during school year than summer
Fitzpatrick et al. (2011)	No	OLS estimation of gain score specification	No	Days between tests	No	School days	SL rate is lower than school-year learning rate
Fryer & Levitt (2004)	Yes	Compare adjusted and unadjusted gaps at end of K and start of first grade	Sometimes	No	Yes	Race	No significant racial difference in math or reading SL
McCoach et al. (2006)	No	Multilevel Growth Curve Model; No gains or lags; Student FE	Yes	Months of instruction prior to test; summer dummy	No	School & student SES	Slight heterogeneity in summer reading gains (no analysis of math)
Quinn (2014)	Yes	Several estimation methods; no covariates	No	Yes	Check	Race	Magnitude and direction of black-white gap sensitive to estimation strategy

Notes: Weights refers to the use of sampling weights provided by the ECLS-K. There are multiple ECLS-K weights depending on the wave and unit of analysis; not all studies identify the exact weight used. Same sample refers to whether or not the study conducted all primary analyses on the same analytic sample.

Table 2: Student and School Descriptive Statistics

	Mean	S.D.
Reading Achievement		
Standardized spring K score (y^A)	0.21	0.98
Standardized fall 1st score (y^D)	0.16	0.96
Unadjusted summer gain ($y^D – y^A$)	-0.05	0.45
Unadjusted K school-year gain	0.08	0.63
Unadjusted 1st grade school-year gain	0.08	0.58
Math Achievement		
Standardized spring K score (y^A)	0.26	0.96
Standardized fall 1st score (y^D)	0.24	0.92
Unadjusted summer gain ($y^D – y^A$)	-0.02	0.56
Unadjusted K school-year gain	0.06	0.66
Unadjusted 1st grade school-year gain	0.04	0.58
Calendar Days Between Important Dates		
Spring K test and fall 1st test ($d^D – d^A$)	151.5	20.5
End of K and start of 1st ($d^C – d^B$)	80.8	5.3
Start of 1st and fall 1st Test ($d^D – d^C$)	40.6	14.6
Spring K test and end of K ($d^B – d^A$)	30.2	14.4
Days between K tests	187.1	21.0
Days between first-grade tests	209.5	20.8
Student Characteristics		
White	74.3%	
Black	10.8%	
Hispanic	09.5%	
Other race/ethnicity	5.4%	
Female	50.7%	
Poverty	12.9%	
Does not speak English at home	3.5%	
Has Individualized Education Plan (IEP)	5.1%	
Kindergarten Redshirt	7.1%	
Mom does not have high school diploma	6.7%	
Mom has high school diploma	34.4%	
Mom attended some college	31.2%	
Mom has bachelor’s degree (or more)	27.7%	
Computer at Home	62.8%	
Number of Books at Home	112.9	134.3

Table 2, Continued

Summer Activities		
Organized Summer Activities	54.6%	
Attended Summer School	8.6%	
# of Trips to Library/Bookstore	6.9	7.0
Never Practice Math	18.6%	
Sometimes Practices Math	70.7%	
Practices Math Everyday	10.7%	
Never Reads to Child	2.4%	
Sometimes Reads to Child	51.9%	
Reads to Child Everyday	45.6%	
Attended Summer Camp	24.8%	
Attended Summer Tutoring	2.4%	
Attended Summer Daycare	10.1%	
School Characteristics		
Enrollment	504.0	260.4
School-wide Title I	39.6%	
< First Grade School	0.1%	
Primary School	12.5%	
Elementary School	70.4%	
Combined School	16.3%	
% Minority Students	29.2%	
% Reduced Lunch Students	13.1%	
% Hispanic Students	30.7%	
Attends Private School	16.2%	
Attends Urban School	31.0%	
Attends Suburban School	44.2%	
Attends Rural School	24.7%	
N Children	1,350	
N Schools	100	

Notes: Means and standard deviations (S.D.) are weighted by NCES provided sampling weights to account for unequal probabilities of sample selection. S.D. are only reported for non-binary variables. K = kindergarten and 1st = first grade. Achievement gains are not adjusted for differences in test dates. Primary school has only kindergarten and 1st Grade. Elementary school has K-5. A combined school has additional grade levels beyond 5th grade.

Table 3: Summer Vacation Length Regressions (OLS estimates)

Covariates:	(1) Student	(2) Summer	(3) School	(4) All
Student Covariates	Yes	No	No	Yes
F Statistic	1.44			1.12
(p value)	(0.15)			(0.35)
Summer Activity Covariates	No	Yes	No	Yes
F Statistic		1.63		1.31
(p value)		(0.11)		(0.23)
School Covariates	No	No	Yes	Yes
F Statistic			1.25	1.19
(p value)			(0.26)	(0.30)
Adjusted R^2	0.021	0.013	0.107	0.132

Notes: N = 1,350 (rounded to nearest 50). The dependent variable is the natural log of summer vacation days.

Table 4: Gain-Score Estimates of Average Daily Learning Rates

	Math Achievement				Reading Achievement			
	1	2	3	4	5	6	7	8
Days between K tests	0.004***				0.006***			
	(0.001)				(0.001)			
Days between first-grade tests		0.005***				0.003**		
		(0.001)				(0.001)		
Days between spring K and fall first-grade tests ($d^D - d^A$)			0.005***				0.004***	
			(0.001)				(0.001)	
Start of first grade – fall first-grade test ($d^D - d^C$)				0.007***				0.004***
				(0.002)				(0.001)
Summer Vacation ($d^C - d^B$)				0.004				0.009***
				(0.005)				(0.003)
End of K – spring K test ($d^B - d^A$)				0.002*				0.002***
				(0.001)				(0.001)
Adjusted R^2	0.01	0.03	0.03	0.04	0.04	0.01	0.03	0.03
Tests of Equality								
($d^C - d^B$) = ($d^D - d^C$)								
(p-value)				0.33				0.07*
($d^C - d^B$) = ($d^B - d^A$)								
(p-value)				0.97				0.02**

Notes: N = 1,350 (rounded to nearest 50). Standard errors are robust to clustering at the school level. These specifications contain no controls. The dependent variable is the unadjusted difference between standardized test scores. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Lag-Score Estimates of Average Daily Learning Rates

	Math Achievement				Reading Achievement			
	1	2	3	4	5	6	7	8
Days between K tests	0.004***				0.006***			
	(0.001)				(0.001)			
Days between first-grade tests		0.004***				0.003**		
		(0.001)				(0.001)		
Days between spring K and fall first-grade tests ($d^D - d^A$)			0.004***				0.003***	
			(0.001)				(0.001)	
Start of first grade – fall first-grade test ($d^D - d^C$)				0.008***				0.004***
				(0.002)				(0.001)
Summer Vacation ($d^C - d^B$)				-0.001				0.006**
				(0.004)				(0.003)
End of K – spring K test ($d^B - d^A$)				0.001				0.001
				(0.001)				(0.001)
Adjusted R^2	0.60	0.64	0.68	0.69	0.64	0.66	0.81	0.81
Tests of Equality								
($d^C - d^B$) = ($d^D - d^C$)								
(*p*-value)				0.04**				0.55
($d^C - d^B$) = ($d^B - d^A$)								
(*p*-value)				0.66				0.10

Notes: N = 1,350 (rounded to nearest 50). Standard errors are robust to clustering at the school level. These specifications contain no controls, but do condition on lagged (spring-K) test scores. The dependent variable is the unadjusted difference between standardized test scores. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Predictors of Average Summer Learning

	Math		Reading	
	Gain Score	Lag Score	Gain Score	Lag Score
	(1)	(2)	(3)	(4)
$(d^D - d^C)$	0.007***	0.008***	0.004***	0.005***
	(0.002)	(0.001)	(0.001)	(0.001)
$(d^C - d^B)$	0.002	-0.001	0.009***	0.006**
	(0.005)	(0.005)	(0.003)	(0.003)
$(d^B - d^A)$	0.003**	0.001	0.002*	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Lag score (spring K)		0.735***		0.865***
		(0.025)		(0.014)
White	Omitted			
Black	-0.019	-0.151**	-0.038	-0.075
	(0.073)	(0.058)	(0.050)	(0.049)
Hispanic	-0.070	-0.065	0.006	0.012
	(0.070)	(0.057)	(0.037)	(0.037)
Other race	0.012	-0.035	0.053	0.055
	(0.061)	(0.056)	(0.049)	(0.047)
Female	-0.021	-0.014	-0.017	0.006
	(0.034)	(0.030)	(0.025)	(0.024)
Poverty	0.008	-0.054	-0.071**	-0.100***
	(0.051)	(0.052)	(0.035)	(0.035)
English is not spoken at home	0.088	-0.014	0.091	0.055
	(0.118)	(0.103)	(0.071)	(0.064)
IEP	0.082	-0.080	0.058	-0.019
	(0.094)	(0.080)	(0.056)	(0.049)
Kindergarten redshirt	0.040	0.054	0.037	0.027
	(0.070)	(0.066)	(0.048)	(0.048)
Mother's education				
No high school diploma	0.024	-0.061	0.069	0.027
	(0.067)	(0.066)	(0.055)	(0.049)
High school diploma	Omitted			
Some college	-0.007	0.015	0.005	0.017
	(0.043)	(0.038)	(0.029)	(0.027)
Four-year college degree	0.008	0.140**	-0.052	0.025
	(0.067)	(0.064)	(0.035)	(0.034)
Private school	-0.015	0.028	0.040	0.049
	(0.055)	(0.047)	(0.043)	(0.043)
Suburban school	Omitted			
Urban school	-0.084	-0.080*	-0.012	-0.010
	(0.055)	(0.046)	(0.036)	(0.037)
Rural school	-0.073	-0.073	0.007	0.017
	(0.063)	(0.053)	(0.043)	(0.041)
Adjusted R^2	0.04	0.70	0.04	0.81
Tests of Equality (*p* value)				
$(d^C - d^B) = (d^D - d^C)$	0.33	0.04	0.08*	0.60
$(d^C - d^B) = (d^B - d^A)$	0.88	0.62	0.02**	0.07

Notes: N = 1350. Standard errors are robust to clustering at the school level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 7: Heterogeneity in Average Summer Learning Rates (OLS Gain Score Estimates)

	Math		Reading	
	Gain Score	Lag Score	Gain Score	Lag Score
	(1)	(2)	(3)	(4)
$(d^D - d^C)$	0.006***	0.007***	0.005***	0.005***
	(0.002)	(0.001)	(0.001)	(0.001)
$(d^C - d^B)$ (Summer)	-0.026*	-0.019	0.010	0.010
	(0.015)	(0.012)	(0.009)	(0.009)
$(d^B - d^A)$	0.003***	0.002*	0.002	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Lag-score×Summer		0.008**		-0.001
		(0.004)		(0.003)
Black×Summer	-0.004	0.011	0.009	0.008
	(0.011)	(0.011)	(0.012)	(0.012)
Hispanic×Summer	0.009	0.000	0.002	0.002
	(0.008)	(0.007)	(0.006)	(0.007)
Other race×Summer	-0.003	-0.006	0.017**	0.012
	(0.010)	(0.010)	(0.008)	(0.008)
Female×Summer	0.006	0.005	-0.003	-0.003
	(0.006)	(0.006)	(0.005)	(0.005)
Poverty×Summer	0.002	-0.005	-0.010	-0.012*
	(0.010)	(0.010)	(0.007)	(0.007)
Non-English household×Summer	0.028	0.031	0.013	0.011
	(0.025)	(0.021)	(0.011)	(0.012)
IEP×Summer	0.000	-0.002	0.021**	0.014
	(0.017)	(0.015)	(0.010)	(0.009)
Mom no high school×Summer	0.000	-0.007	0.023**	0.014*
	(0.015)	(0.013)	(0.010)	(0.008)
Mom some college×Summer	0.012*	0.005	-0.003	-0.007
	(0.007)	(0.006)	(0.005)	(0.005)
Mom college×Summer	0.001	-0.005	-0.005	-0.009
	(0.010)	(0.009)	(0.006)	(0.005)
Redshirt×Summer	-0.003	-0.017	0.003	-0.001
	(0.010)	(0.012)	(0.009)	(0.009)
Household has computer×Summer	-0.001	0.002	0.005	0.005
	(0.006)	(0.006)	(0.006)	(0.005)
Household books×Summer	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Org. Summer Activity×Summer	0.001	0.002	-0.005	-0.002
	(0.005)	(0.004)	(0.007)	(0.006)
Summer school×Summer	0.009	0.017**	-0.011	-0.006
	(0.009)	(0.008)	(0.007)	(0.006)
Summer library/bookstore trips×S	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)
Parent never helps with math×S	-0.000	0.003		
	(0.009)	(0.007)		
Parent frequently helps with math×S	0.011	0.014		
	(0.010)	(0.010)		
Parent never reads×Summer			0.004	0.021
			(0.016)	(0.015)

Table 7, Continued

Parent reads every day×Summer			-0.004	-0.004
			(0.005)	(0.005)
Attends summer camp×Summer	-0.001	-0.002	0.003	0.005
	(0.007)	(0.006)	(0.005)	(0.005)
Attends summer tutor×Summer	0.026	0.034	0.003	0.002
	(0.025)	(0.025)	(0.010)	(0.011)
Attends summer day care×Summer	0.008	0.007	0.003	0.003
	(0.008)	(0.007)	(0.007)	(0.007)
School size×Summer	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Title 1 school×Summer	0.006	0.005	-0.007	-0.006
	(0.009)	(0.007)	(0.007)	(0.006)
% Reduced lunch×Summer	-0.019	-0.008	0.030*	0.029
	(0.026)	(0.023)	(0.018)	(0.018)
Private school×Summer	0.050***	0.039**	0.015	0.008
	(0.018)	(0.015)	(0.010)	(0.011)
Urban school×Summer	0.000	-0.003	0.001	0.001
	(0.011)	(0.009)	(0.006)	(0.006)
Rural school×Summer	0.003	0.003	0.011	0.010
	(0.012)	(0.011)	(0.007)	(0.007)
Adjusted R^2	0.04	0.70	0.04	0.81
Joint significance of interactions (F)	2.22	2.12	2.32	2.76
p-value	(0.00)	(0.00)	(0.00)	(0.00)

Notes: N = 1350. Standard errors are robust to clustering at the school level. The variables interacted with summer vacation length (Summer) are included in the model in levels, but these coefficients are not reported in the interest of brevity. The results are qualitatively when the interactions are added to the baseline model one at a time.
*** p<0.01, ** p<0.05, * p<0.1.

Appendix Table A.1: Reading Achievement Specification Tests

	Linear	Quadratic	Cubic	Quartic	RESET Tests	
	1	2	3	4	5	6
$(d^D - d^C)$	0.004***	0.000	0.001	-0.082	0.004***	0.004***
	(0.001)	(0.004)	(0.016)	(0.055)	(0.001)	(0.001)
$(d^C - d^B)$	0.009***	-0.008	-0.239	12.664*	0.010**	0.011***
	(0.003)	(0.041)	(0.560)	(6.470)	(0.004)	(0.004)
$(d^B - d^A)$	0.002*	-0.003	-0.027***	-0.032***	0.002**	0.002**
	(0.001)	(0.004)	(0.007)	(0.008)	(0.001)	(0.001)
$(d^D - d^C)^2$		0.000	0.000	0.003		
		(0.000)	(0.000)	(0.002)		
$(d^C - d^B)^2$		0.000	0.003	-0.242*		
		(0.000)	(0.007)	(0.123)		
$(d^B - d^A)^2$		0.000	0.001***	0.001**		
		(0.000)	(0.000)	(0.000)		
$(d^D - d^C)^3$			0.000	-0.000		
			(0.000)	(0.000)		
$(d^C - d^B)^3$			-0.000	0.002*		
			(0.000)	(0.001)		
$(d^B - d^A)^3$			-0.000***	-0.000		
			(0.000)	(0.000)		
$(d^D - d^C)^4$				0.000		
				(0.000)		
$(d^C - d^B)^4$				-0.000**		
				(0.000)		
$(d^B - d^A)^4$				0.000		
				(0.000)		
$\hat{y}^2$					2.905	1.546
					(2.676)	(1.679)
$\hat{y}^3$					9.189	
					(14.943)	
RESET (F) Statistic					0.67	
(p-value)					(0.51)	
APE						
$(d^D - d^C)$	0.004***	0.004***	0.003***	0.003***	.	.
	(0.001)	(0.001)	(0.001)	(0.001)		
$(d^C - d^B)$	0.009***	0.009***	0.009***	0.009***	.	.
	(0.003)	(0.001)	(0.003)	(0.003)		
$(d^B - d^A)$	0.002*	0.002	0.001	0.001	.	.
	(0.001)	(0.001)	(0.001)	(0.001)		
Adjusted R^2	0.031	0.032	0.041	0.046	0.031	0.031

Notes: N = 1,350 (rounded to nearest 50). Standard errors are robust to clustering at the school level. APE = Average Partial Effect. $\hat{y}$ is the OLS fitted value from the linear specification estimated in column 1. The RESET F statistic tests the joint significance of $\hat{y}^2$ and $\hat{y}^3$. p<0.01, ** p<0.05, * p<0.1.

Appendix Table A.2: Math Achievement Specification Tests

	Linear	Quadratic	Cubic	Quartic	RESET Tests	
	1	2	3	4	5	6
$(d^D - d^C)$	0.007***	0.006	0.025	-0.110	0.007**	0.007***
	(0.002)	(0.006)	(0.026)	(0.106)	(0.003)	(0.002)
$(d^C - d^B)$	0.002	0.049	1.369	27.00***	0.004	0.004
	(0.005)	(0.087)	(1.097)	(5.990)	(0.005)	(0.005)
$(d^B - d^A)$	0.003*	0.001	0.014	0.023***	0.002	0.002*
	(0.001)	(0.004)	(0.011)	(0.008)	(0.001)	(0.001)
$(d^D - d^C)^2$		0.000	-0.000	0.005		
		(0.000)	(0.001)	(0.004)		
$(d^C - d^B)^2$		-0.000	-0.017	-0.505***		
		(0.001)	(0.014)	(0.113)		
$(d^B - d^A)^2$		0.000	-0.000	-0.001**		
		(0.000)	(0.000)	(0.001)		
$(d^D - d^C)^3$			0.000	-0.000		
			(0.000)	(0.000)		
$(d^C - d^B)^3$			0.000	0.004***		
			(0.000)	(0.001)		
$(d^B - d^A)^3$			0.000	0.000		
			(0.000)	(0.000)		
$(d^D - d^C)^4$				0.000		
				(0.000)		
$(d^C - d^B)^4$				-0.000***		
				(0.000)		
$(d^B - d^A)^4$				-0.000		
				(0.000)		
$\hat{y}^2$					0.354	0.456
					(1.120)	(1.157)
$\hat{y}^3$					2.352	
					(8.277)	
RESET (F) Statistic					0.11	
(p-value)					(0.90)	
APE						
$(d^D - d^C)$	0.007***	0.007***	0.008***	0.007***	.	.
	(0.002)	(0.002)	(0.002)	(0.001)		
$(d^C - d^B)$	0.002	0.002	0.001	0.001	.	.
	(0.005)	(0.005)	(0.005)	(0.001)		
$(d^B - d^A)$	0.003*	0.002*	0.003**	0.002*	.	.
	(0.001)	(0.001)	(0.001)	(0.001)		
Adjusted R^2	0.037	0.036	0.039	0.052	0.036	0.037

Notes: N = 1,350 (rounded to nearest 50). Standard errors are robust to clustering at the school level. APE = Average Partial Effect. $\hat{y}$ is the OLS fitted value from the linear specification estimated in column 1. The RESET F statistic tests the joint significance of $\hat{y}^2$ and $\hat{y}^3$. p<0.01, ** p<0.05, * p<0.1.