

DRAFT: COMMENTS WELCOME

QUANTITY OF SCHOOLING AND EDUCATIONAL INEQUALITY:
FULL-DAY KINDERGARTEN IN THE USA

Adam Gamoran and Carolina Milesi
University of Wisconsin-Madison, USA

A paper prepared for presentation at the meetings of the
Research Committee on Social Stratification, International Sociological Association
Tokyo, March 2003

QUANTITY OF SCHOOLING AND EDUCATIONAL INEQUALITY:
FULL-DAY KINDERGARTEN IN THE USA

ABSTRACT

While much sociological research seeks achievement differences in comparisons of different schools or different classrooms to one another, a more consequential comparison is between the presence and absence of schooling. This paper explores a new source of evidence on the effects of exposure to schooling: Full-day versus half-day kindergarten programs in the United States. (Kindergarten is first year of elementary school, for children aged 5-6.) A new national survey of over 20,000 kindergarteners and their principals, teachers, and parents shows that students in full-day programs obtain higher achievement growth in both mathematics and reading. This result holds up to extensive statistical controls. Full-day programs also exhibit more time spent on academic instruction, and instructional differences account for part of the full-day program effects. Full-day kindergarten appears equally beneficial for students of all racial, ethnic, and socioeconomic backgrounds. Offering more full-day programs in areas of concentrated disadvantage would reduce achievement inequality. Universal full-day programs would do little to reduce inequality, but would raise average cognitive skills across the board.

QUANTITY OF SCHOOLING AND EDUCATIONAL INEQUALITY: FULL-DAY KINDERGARTEN IN THE USA

Recent changes in the political and policy environment of American education have raised the stakes for academic performance. Schools must now focus more than ever on standardized test scores, and achievement is more clearly than ever regarded as the primary if not the sole aim of schooling. The federal *No Child Left Behind* (NCLB) act of 2002 holds schools and school districts accountable for student performance, and imposes sanctions if they do not succeed. Many schools and districts have passed on this accountability to students by linking promotion and graduation to test performance. By requiring schools to report average performance for a variety of demographic subgroups, NCLB shines a light on achievement distributions within schools as well as on levels of performance.

Despite nearly 40 years of research, sociological contributions to questions about how schools can raise standards and reduce inequality are modest. Sociologists have had great success in identifying the relations between home background and school performance, but have had less to offer on the links between school conditions and achievement outcomes. Clear evidence shows that student achievement varies within schools far more than between them, so not surprisingly the average characteristics of schools offer little explanation for the wide range of variation in student achievement (e.g., Coleman et al., 1966). Conditions of schooling within schools – the organization of classes into tracks or ability groups, and the experiences of instruction within classes – do a better job of predicting student achievement than average school characteristics (e.g., Gamoran, 1987).

An important limitation of studies that examine differences within as well as between schools is that comparisons of one school or one classroom to another fall within a restricted range of variation: Schools and classes within the same country respond to the same institutional mandate for the structure and processes of schooling, and so they are more alike than different (Meyer, 1977, 1980). Comparisons within this limited range fail to show the full effects of schooling on achievement. A more meaningful comparison would be “school” versus “not-school,” but that is difficult to carry out because in developed countries, virtually everyone goes to school. Still, two types of evidence have shown important school effects when the focus is on exposure to schooling. First, studies of school dropouts show that students who remain in high school exhibit more cognitive growth than students who drop out (e.g. Gamoran, 1987). Second, studies of “summer learning” demonstrate that elementary students from different racial and economic backgrounds make similar progress during the school year, but during the summer, test scores of students from advantaged families continue to rise while those of their less-advantaged counterparts do not (Heyns, 1978; Entwisle, Alexander, and Olson, 1997). This pattern shows how exposure to schooling helps compensate for disadvantages in the home and neighborhood environments of poor and minority students (Gamoran, 1996).

In this paper we provide more evidence of the effects of exposure to schooling by comparing the cognitive growth of kindergartners (5 and 6 year olds, the first year of elementary school) who attend school for a full day, compared to kindergartners who attend for only half of the school day. We offer this comparison to demonstrate the effects of *quantity of schooling* at the onset of elementary school. We further show that

differences between full- and half-day kindergarten programs reflect differences in the *quantity of academic instruction* that is offered in the different programs.

The Quantity and Quality of Schooling in Kindergarten

Most of the attention from sociologists to the expansion of schooling has focused on teenagers and adults staying in school longer and longer (e.g., Shavit and Blossfeld, 1993). Yet another important dimension of expansion is that children are entering formal schooling at earlier and earlier ages (Mare, 1995). In the U.S., kindergarten is nearly universal, and schooling for 3-5 year olds shows signs of being extended onto a mass scale. Over 50% of 3-5 year olds now attend some form of center-based care, with enrollments higher among children from economically advantaged than from disadvantaged families (U.S. Department of Education, 2001). Among children of kindergarten age, about half are enrolled in half-day programs of generally 2-3 hours duration, and about half are enrolled in full-day programs averaging 5-6 hours per weekday (West, Denton, and Germino-Hausken, 2000). Overall, the proportion of 5-6 year olds in school has increased dramatically throughout the century.

Traditional kindergartens were part-time, and combined time for developing social skills through play with school readiness activities such as manipulating objects, counting, and learning letters and sounds (Brosterman, 1997). The shift to full-day programs on a wide scale is a recent phenomenon, and no doubt reflects at least two factors: the increase in the numbers of mothers in the full-time labor force, and the increase in children experiencing full-time day care in the pre-school years (Chmelynski, 1998). Whereas the half-day program seems congruent with a family lifestyle in which one parent stays home while the child is gradually introduced to increasing amounts of

outside care, the full-day kindergarten fits the increasingly common family pattern in which full-time care begins at age 3 or earlier. The introduction of compensatory pre-schools for disadvantaged children (i.e., Head Start) has also probably fueled the growth of full-day kindergarten programs.

Among educators, one often hears that full-day programs offer no academic advantage over half-day programs. According to this claim, both programs offer the same academic activities, and full-day programs use the additional time for more play, for lunch time and for nap time. In short, according to this view, full-day programs offer extended child care, but not more instruction. Many public school districts are reluctant to provide more than a half-day program, leading parents who need full-time care to seek private schools. The common wisdom is belied, however, by a 1993 national survey of kindergarten teachers conducted by the U.S. government. Although higher proportions of full-day than half-day classes offered time for free play (72 percent to 62 percent) and for running, jumping, and climbing (72 percent to 48 percent), teachers of full-day classes were also more likely to use objects for math or science (61 percent to 41 percent), to use worksheets for math or science (20 percent to 9 percent) and to use worksheets for literacy skills (25 percent to 13 percent) (U.S. Department of Education, 2001). A 1998 survey revealed that 89 percent of full-day programs spent an hour or more per day in small-group activities, in contrast to 62 percent of half-day programs (U.S. Department of Education, 2001). Full-day programs may provide additional child care, but they also appear to provide more academic instruction. In pre-school, more hours of high-quality care result in more cognitive growth (NICHD Early Child Care Research Network, 2002), and the same may hold for kindergarten. Consequently, we expect that full-day programs

produce more cognitive growth and thus better prepare children for first grade than half day programs:

Hypothesis 1: Full-day kindergarten promotes more cognitive growth than half-day kindergarten.

Although our prediction runs contrary to a conventional wisdom, recent studies of local school districts support the expectation that full-day kindergarten enhances readiness for first grade. A trial in Philadelphia, for example, showed higher test performance for graduates of full-day kindergarten than for their peers who attended half-day programs (Viadero, 2002). Many districts and states are debating the costs of moving to full-day programs, so a national assessment of the phenomenon can contribute to policy discussions (Chmelynski, 1998; Bowman, 2003).

Other studies that compare the presence and absence of schooling show that schooling is more important for the cognitive growth of minority and economically disadvantaged children than it is for their more privileged peers. In particular, studies of summer learning indicate that the achievement gap widens much more during the summer, when school is not in session, than during the school year, when school is going on (Heyns, 1978; Entwisle, Alexander, and Olson, 1997). We anticipate the same pattern for full-day and half-day kindergarten. Because of differences in the home and neighborhood environments of advantaged and disadvantaged children, we predict that full-day programs will be especially valuable for the disadvantaged:

Hypothesis 2: The cognitive benefits of full-day kindergarten will be especially pronounced for minority and for economically disadvantaged children.

Quality of Academic Instruction in Kindergarten

What aspects of academic instruction are likely to promote achievement in mathematics and reading? Heated debates have raged in both subject areas. In mathematics, educators following the guidelines of the National Council of Teachers of Mathematics (1989, 1991, 2000) emphasize deep conceptual understanding rather than memorizing algorithms. This approach, sometimes called “teaching for understanding,” privileges solving complex problems and engaging students in explaining their solutions (whether correct or incorrect), often using objects to model mathematical ideas, over drilling on mathematical computations, which may be more quickly resolved with electronic calculators (Cohen, McLaughlin, and Talbert, 1993; Burrill, 2001; Gamoran et al., 2003). Critics, including some renowned mathematicians, are skeptical of this “fuzzy math” and call for maintaining the traditional emphasis on drill and practice (Loveless, 1997; Askey, 2001). A recent report by the National Research Council tries to reconcile the opposing positions by calling for a balanced curriculum (Kilpatrick, Swafford, and Findell, 2001).

Reading has also been the subject of a “curriculum war,” between those who favor teaching reading by focusing on words within the context of literature (“whole language”), and those who emphasize letters and sounds as the building blocks of words (“phonics”). Another committee of the National Research Council tried to bring consensus to this issue, concluding that an emphasis on letter-sound relationships within the context of a rich literature environment – a balanced approach – was the most appropriate response (Snow, Burns, and Griffin, 1998). Although not all scholars agree,

the prevailing view is perhaps best summed up by Ravitch (2001, p. 226), who proclaimed: “Both sides should declare victory and go home.”

Based on the conclusions in both subject areas, we expect to find that more academic instruction of all types – teaching for understanding and drill and practice in math, and whole language and phonics in reading – contributes to achievement growth during kindergarten, and that higher levels of all types account for the advantages of full-day over half-day programs:

Hypothesis 3: The achievement advantages of full-day over half-day programs can be attributed to greater amounts of academic instruction.

Instructional approaches that focus on modeling, conceptual understanding, and manipulating objects – as emphasized in teaching for understanding in mathematics – are consistent with the traditional activities of kindergarten, as designed in the 19th century and diffused in the early 20th (Brosterman, 1997). It is not clear, however, whether these activities would pay off on contemporary standardized tests. Indeed, early followers of the kindergarten movement would no doubt abhor narrow tests of reading and numeracy skills, embracing instead the child’s curiosity and ability to reason. Thus, which sorts of academic activities enhance cognitive growth (as measured by standardized tests) is open to question.

Organization of Reading Instruction

Formal instruction in elementary school reading, particularly in the earliest grades, usually occurs in small groups arranged according to children’s reading skills (or “reading readiness”) at the beginning of the school year (Barr, 1975). Long a hallmark of first-grade reading instruction, to the extent that academic instruction is extended to

kindergarten, we may expect to find reading groups used in kindergarten as well. On the one hand, the relatively homogeneous small group offers students an opportunity for closer interaction with the teacher and a more fine-grained focus on the student's skill level than is possible in a whole-class format. On the other hand, reading groups are a form of differentiation that may be associated with inequality, if high groups cover more academic material at faster pace than low groups (Barr and Dreeben, 1983; Rowan and Miracle, 1983; Gamoran, 1986). No quantitative studies that we are aware of have examined the effects of reading groups in kindergarten. Rist (1970) provided a well-known ethnographic account of a kindergarten class in which students appeared to be sorted for instruction by their family economic backgrounds, with advantaged students receiving most of the teacher's attention and disadvantaged students virtually ignored. This case suggests that grouping is a source of inequality. Gamoran (1986) observed a growing gap in reading achievement between high and low groups over the course of first grade, a pattern that was attributable to instructional differentiation. Slavin's (1987) review of elementary-school grouping uncovered only one study of within-class grouping for reading, and this yielded positive results for high, middle, and low groups compared to ungrouped classes, as did several similar studies of within-class grouping for math. Slavin speculated that the shortage of comparisons of within-class grouping to whole-class instruction in reading occurred because reading groups are "so widespread in elementary schools that formation of ungrouped control groups is difficult to arrange" (p. 317). The kindergarten year may offer more prospects for comparison, particularly with the rise of full-day classes that focus more on academic instruction. We anticipate that

reading groups promote achievement in reading, at least among students assigned to middle and high groups:

Hypothesis 4: More reading group time is another condition that contributes to the achievement advantage of full-day over half-day kindergarten programs.

Hypothesis 5: Although reading groups promote higher achievement on average, they also magnify inequality between strong and weak readers.

Data and Methods

To extend research on the quantity of schooling to kindergarten, we use data from the Early Childhood Longitudinal Study-Kindergarten Class of 1998-1999 (ECLS-K). Starting in the fall of 1998, the study follows a U.S. nationally representative sample of approximately 22,000 children from kindergarten through fifth grade. The children come from about 1,000 kindergarten programs in public and private schools offering full and half-day kindergarten programs. The sample consists of children from different racial-ethnic and socioeconomic backgrounds, including an oversample of Asian children, private kindergartens, and private school kindergartners.

Children's cognitive performance was assessed in the fall and spring of kindergarten year through an untimed one-on-one computer-assisted personal interview. The reading component of the ECLS-K assessment was only administered to children who were proficient in English. If the child was not proficient in English (English proficiency is determined at each round of data collection), but spoke Spanish, s/he was administered a Spanish translation of the math section of the direct assessment (5 percent of the total sample, and 30 percent of Hispanic children in the sample were assessed in Spanish). Children from other non-English home language backgrounds who were not

proficient in English were excluded from the assessment (approximately 2 percent of the sample).

The child's parents or guardians, teachers, and school administrators were also interviewed. During the fall and spring of kindergarten year parents or guardians provided information on children's school experience and socioeconomic background. School administrators completed a self-administered questionnaire in the spring of kindergarten where they provided information on the physical, organizational, and fiscal characteristics of their schools, and on the schools' learning environment and programs.

The child's teacher filled out three self-administered questionnaires in the fall of kindergarten and two in the spring of the same year (an additional questionnaire was given in the spring to teachers who were new to the study). Two of these questionnaires (one in the fall and one in the spring of kindergarten) are child-specific, that is, the teacher evaluates each sampled child on several noncognitive and cognitive dimensions, such as the reading ability group in which the child is placed with respect to the class. The other questionnaires pertained to the teacher's own background, teaching practices, experience and the classroom setting(s) where s/he teaches. Some questions in these instruments refer to the specific class(es) the teacher teaches (morning, afternoon, or all-day class). Class size, racial composition, and class behavior are among this type of question. Another set of questions is asked regardless of the type of class teachers teach. Instructional practices are among these questions, meaning that teachers are assumed to apply the same pedagogy in all classroom contexts, and in all instructional groups within each class. Moreover, given that most of the instructional practices questions are asked in the spring of kindergarten, it is necessary to assume that teachers' accounts are consistent

over the school year, and that the unobserved variation of classroom activities throughout the year is not simultaneously correlated with student achievement.

Variables and Measures

The dependent variable in this study – math and reading student achievement – is measured by IRT (item response theory) scores obtained by each kindergartner in the spring semester. IRT scores obtained in the fall of kindergarten in the corresponding subject are included as statistical controls.

Quantity and quality of academic instruction are the main explanatory variables in the analysis. Quantity of schooling is measured as a categorical variable. The model includes an indicator of full-day program, with half-day (morning or afternoon) kindergarten programs as the reference category. Quality of instruction is measured by two sets of teaching practices. First, three variables address class organization and teachers' time allocation. They correspond to the amount of time per day (in a "typical day") teachers use teacher-directed whole class activities, teacher-directed small group activities, and teacher-directed individual activities.¹ These questions, taken from the spring teacher questionnaire, were not subject-specific, so they were introduced in both the reading and math models.

The second set of instruction practices corresponds to those debated in the reading and math "curriculum wars." Representing the extreme positions within these "wars" (Whole Language versus Phonics in reading, and Teaching for Understanding versus Drill in math) two scales of instructional activities were constructed in each subject, based on teachers' reports of their practices. The spring teacher questionnaire contains a

¹ The variables' response categories are: 1=no time; 2=half hour or less; 3=about one hour; 4=about two hours; and 5=three hours or more.

23-item question regarding the frequency of teacher's practices in reading and language arts, and a 17-item question regarding instructional practices in math.² An exploratory factor analysis was performed using reading and math items separately. Two factors (with eigenvalues greater than one) were retained in each subject. The scales of instructional activities were constructed by summing up the responses to the items with higher loadings on each of the factors. The Whole Language scale sums up the following four items: write with encouragement to use invented spellings if needed, read books children have chosen for themselves, compose and write stories or reports, and write stories in journal. The four items included in the Phonics scale are: work on phonics, work on learning the names of the letters, practice writing the letters of the alphabet, and work in a reading workbook or on a worksheet. Cronbach's alpha for these scales is 0.77 for Whole Language and 0.44 for Phonics. The Teaching for Understanding scale contains the following three items: work with counting manipulatives to learn basic operations, solve math problems in small groups or with a partner, and work on math problems that reflect real-life situations. The Drill scale includes: do math worksheets, do math problems from children's textbooks, and complete math problems on chalkboard. Cronbach's alpha in this case is 0.63 for Teaching for Understanding and 0.59 for the Drill scale.

Other important characteristics of classes included in the model are the percentage of black and Hispanic children in the classroom (as reported by the teacher), an indicator of private kindergarten, as compared with kindergarten programs in public schools, teacher's rating of class behavior – a 1 to 5 scale, with 5 indicating well-behaved classes,

² The categories of response in all the items are: 1=never; 2=once a month or less; 3=two to three times a month; 4=one to two times a week; 5=three to four times a week; 6=daily.

present in the fall teacher questionnaire – and an indicator for small class size. The indicator of small class was computed as the first quintile of the distribution of class sizes (at the class level), which corresponds to classes that have 17 or fewer students. This same threshold was used by the Tennessee’s Project STAR, where small classes had between 13 and 17 students (Finn and Achilles 1999).

As a way of controlling for the nonrandom assignment of students to different classes we introduced measures of student’s family socioeconomic background, race, gender, and age. Family SES is a continuous variable that ranges from -4.75 to 2.75 . It is computed as an average composite of up to five measures: father/male guardian’s education and occupation, mother/female guardian’s education and occupation, and household income. Each of these variables was standardized and approximately 10 percent of each was imputed using hot-deck imputation. Three dummy variables indicating an Asian, black, or Hispanic student refer to the child’s race and ethnicity. Also included is an indicator for male students, and children’s age at kindergarten entrance (measured in months). In the models that predict reading achievement, we added the child’s location within the distribution of reading ability groups in the class (as reported by the teacher in the spring of kindergarten), categorized as being in the high, middle, or low ability group.³ At the class level, we introduced a dummy indicating whether the class had no reading ability groups.

³ The child is considered to be in the “high ability group” if s/he is reported to be in the first or second highest group in a class with five or four groups, or if s/he is in the highest group in a class with three or two ability groups. Children located in the “middle ability group” correspond to those in the third highest group in a class of five groups, or those in the second highest group in a class of three groups. Children are in the “low ability group” when they are in the lowest or second lowest group in a class of five or four groups, or if they are in the lowest group in class of three or two groups.

Methods

We use hierarchical linear models (HLM) to examine the effect of quantity and quality of academic instruction on reading and math student achievement (Raudenbush and Bryk, 2002). We estimated 2-level models of students within classes. The sample design of ECLS-K however, consists of students clustered within schools. One way of accounting for the dependency of observations within schools would be to add a third level of analysis pertaining to schools. This was not possible because the number of classes per school was too small (5.1 on average) and because there were too many schools with only 1 or 2 sampled kindergartens (38 percent of the schools), making it difficult to distinguish between the school and class levels. Although full- vs. half-day program is generally a school-level variable – i.e., the kindergarten classes in a school are either full-day or half-day – in a small number of cases, a school offered both a full-day and a half-day program. This led us to consider full-day kindergarten as a class-level variable. For these reasons, models are estimated at two levels of students within classes, and the school level variables of public vs. private and full- vs. half-day are assigned to the class level.

We constructed class-level data files by matching information from students and teachers. Students whose teachers failed to respond to the survey, and teachers who had no responding students, were lost. Out of the 3832 possible classes, 3549 had matched student and teacher information, meaning that 92.6 percent of original classes were retained.⁴ Out of the 3549 “matched classes” with 18173 “matched students”, the HLM analyses was carried out with approximately two thirds of the student and class cases,

⁴ Of these 3549 classes, 962 were morning classes, 716 were afternoon classes, and 1871 were all-day classes.

with an average of 7.75 sampled children per class. Models that predicted reading achievement were estimated with 10,488 students and 2,275 classes. Math models include 11,544 students and 2,440 classes. The larger number of cases retained in the case of math achievement is due to the fact that Hispanic children who were orally proficient in Spanish at kindergarten entrance were assessed in math, but not in reading.

Results

We present the results in two sections. First, we offer descriptive results that answer the question, who attends full-day kindergarten? This section also compares achievement levels of children in full- and half-day programs. Next, we provide multivariate results that answer the main research questions: What are the effects of full-day kindergarten? Do the effects vary for different students? Are the effects attributable to different amounts and types of academic instruction?

Descriptive Results

Table 1 presents descriptive statistics for full-day and half-day classes, and for the students who attend them. Means are adjusted for the ECLS-K complex sample design. In mathematics, fall achievement averages were virtually identical for students in full- and half-day programs, but students in full-day programs gained about three-quarters of a point more than students in half-day programs. In reading, fall achievement was .32 points higher in full-day classes, and spring achievement was 1.42 points higher, a difference in growth of 1.1 points. Thus, children in full-day programs exhibited more achievement growth in both subjects. Whether this growth can be attributed to the programs and what occurred there, or to other confounding conditions, remains to be seen.

More black and fewer Hispanic students participated in full-day than in half-day programs, and students in half-day programs displayed higher family socioeconomic levels on average. Students in full-day programs were also slightly older than their counterparts in half-day classes. With regard to reading ability groups, 46 percent of the students in full-day classes were located in some ability group within the class. The use of reading ability groups in half-day classes was less frequent, with 35 percent of the students located in some group. Among classes containing any reading groups, the distribution of students to groups was similar between half and full-day classes: almost half of the students were placed in high ability groups, a little over a third were in low ability groups, and the rest (approximately a sixth) were in middle ability groups.

Differences also appear at the class level. One fifth of full-day classes contained 17 or fewer students, compared to approximately one fourth of half-day classes, and on average, full-day classes were larger than half-day ones. Teachers' ratings of class behavior were more positive in half-day than in full-day programs, a difference that may be related to the smaller size of half-day classes (Rice, 1999).

In ECLS-K, the average half-day class lasted 3.1 hours per day. The average full-day class lasted more than two times this amount: 6.4 hours per day. How did full-day classes spend their "extra time"? The descriptives show that full-day programs spent more time on academic instruction than half-day programs. Full-day classes presented a higher frequency of all types of academic practices, that is, teachers in these classes devoted more time to both "teaching for understanding" and "drill" activities in math, more time to both "whole language" and "phonics" activities in reading and language arts, and also used more frequently instructional activities directed either to the whole class, to

small groups, and to individual students. So, contrary to some conventional wisdom, at least part of the “extra” time that full-day classes have is spent on more academic instruction—not just in more play, lunch and nap time. Multivariate analyses will reveal whether more instruction results in higher achievement.

The percentage of private-school classes is higher among full-day programs (15 percent) than among half-day programs (9 percent). Means computed at the school level yield a similar pattern: while 56 percent of private schools offered at least one full-day kindergarten class, 47 percent of public schools did so. Public schools offered a similar proportion of half and full-day kindergarten classes, but private schools offered 2.5 times more full-day than half-day classes.

Table 2 presents descriptive statistics for variables used in the reading and math HLM models. The number of students and classes is higher in the analysis of math, than in the analysis of reading achievement. This difference is due to the higher participation of Hispanic children in the math assessment, compared to the reading test. Because children who were only proficient in Spanish at the beginning of the kindergarten year took the math, but not the reading test, it is necessary to interpret differently the results for each subject. In reading, generalization of results is restricted to children who are English-proficient at kindergarten entrance. In math, results apply more generally to children who are proficient either in English or Spanish at the beginning of kindergarten year.

Multivariate Results

Unconditional models designed to partition the variance into student and class-level components reveal that 72 percent of the variance in spring math achievement and

in spring reading achievement lies among students within classes. The remaining 28 percent of the variance in both subjects lies at the class (or school) level. Table 3 provides variance component statistics for the unconditional (variance components) model and all succeeding models.

Preliminary analyses in both subjects showed no significant variation across classes in the effects of race and ethnicity, but significant variation in the effects of SES. Consequently, our models allow SES to vary across classes, while all other student-level effects are fixed across classes. Student SES is centered around the class means, and class mean SES is included as a predictor of average achievement. Other continuous variables in the student-level model (fall achievement and age) are centered around their grand means, and dummy variables (gender, race/ethnicity, ability group location) are not centered. Thus, the student-level intercepts in these models indicate the achievement of a white female student of average age and SES and in the middle ability group.

Mathematics. Table 4 shows effects on spring math achievement. The first column of results shows the effects of full-day versus half-day kindergarten along with a variety of control variables: two other policy variables (class size and private versus public), three compositional variables (percent black and Hispanic, and average class SES), and individual-level controls for gender, race, ethnicity, SES, age, and fall achievement. The results show, perhaps surprisingly in light of other research, no effects of small classes or private schools, but a one-point advantage for students achievement in full-day as compared with those in half-day kindergarten programs.⁵ This finding

⁵ Other specifications of private school (e.g., Catholic, other religious, and non-religious private versus public) also revealed no achievement differences. Similarly, other specifications of class size (a linear indicator, a dummy for classes of 15 students or smaller, and a dummy for the smallest versus the largest classes) revealed no class size effects.

supports Hypothesis 1. The gap is slightly larger than the raw difference in achievement growth, suggesting that students in half-day programs are advantaged in ways that are taken into account by the control variables. Although full-day kindergarten promotes higher achievement on average, it gives no evidence of an interaction with student SES; its effects are the same for all students, as far as we can tell from this model.⁶

Coefficients for black and Hispanic are negatively associated with achievement at both the student and the class levels, and student and class coefficients for SES indicate positive effects, as do the coefficients for age and prior achievement. There are no gender effects on kindergarten math, and no differences between Asian and white students in net math performance. The three policy variables account for an additional 6 percent of the variance in average math achievement over a baseline model that includes all the control variables (compare Uoj in Model 1 to Uoj in the Baseline Model in Table 3: $4.205 - 3.937 / 4.205 = .064$). About 4 percent of the variance across classes in student SES effects is attributable to the policy variables in Model 1, but none of the predictors are significant. Thus, we find support for Hypothesis 1, but not for Hypothesis 2.

Model 2 introduces instructional variables at the class level. The coefficient for full-day kindergarten is reduced from 1.06 to .732, indicating that about 31 percent of the full-day effect is attributable to more academic instruction in full-day programs. What sorts of instruction promote achievement? The results show that *both* teaching for understanding and drill-and-practice activities promote higher achievement on average, as do classes in which students behave better. In short, according to these results, those seeking a balance of instructional approaches are correct, and the main difference is not

⁶ Despite the lack of significant variation in the effects of race and ethnicity across classes, we tested for interactions by full-day kindergarten, and found no significant effects.

between teaching for understanding and drills, but between more instruction and less.

We find partial support for Hypothesis 3, in that more instruction helps account for the full-day kindergarten effect, but substantial effects remain unexplained in our model.

The results further indicate no significant predictors of the student SES effect, and no additional variance explained in the effects of student SES across classes. For average math achievement, the inclusion of instructional variables accounts for an additional 10 percent of the variance beyond the baseline model (see Table 3).

Reading. In reading, we estimated two additional models to show the effects of reading ability groups along with those of full-day kindergarten and instruction. This was important because, as we learned earlier, full-day classes use reading groups more often than half-day classes. The first column of results in Table 5 shows, as in mathematics, that full-day kindergarten promotes higher achievement. The coefficient, at 1.49 points, appears larger than the effect in math, but in standard deviation units it is only slightly larger (about 12 percent of a standard deviation in math, and 15 percent in reading). Also similar to the findings in math, the policy variables fail to predict variation across classes in the effects of student SES. This model explains about 6 percent of the variation in average reading achievement, and less than 2 percent of the SES effect (see Table 3). As in mathematics, our results support Hypothesis 1, but not Hypothesis 2.

Model 2 for reading adds a dummy variable for the use of reading groups, at the class level. These results show substantially *lower* achievement in classes that did not use reading groups for instruction; moreover the coefficient for full-day classes drops to 1.38, suggesting that the one reason full-day programs produce higher achievement is that they provide time for reading groups. In Model 3, we find that the effects of reading groups

for a given student depend on where the student is located in the hierarchy. The coefficient of $-.629$ for “no reading groups” indicates that having no reading groups is worse than being in the middle level of a grouped classroom (the omitted category). However, the coefficient of -1.566 for students in low groups suggest that these students scored significantly lower than their middle-group counterparts, and indeed, being in the low group is *worse* than having no groups at all ($-1.556 - -.629 = -.927$). Meanwhile, students in the top reading groups obtain the highest achievement; the coefficient of 2.003 reflects their advantage over similar students in similar classes assigned to middle groups, and their advantage over similar students in classes with no reading groups is even larger ($2.003 - -.629 = 2.632$). Thus, both Hypotheses 4 and 5 are supported. Note that the use of grouping in Model 3 is *positively* associated with the student SES effect; that is, the effects of SES are greater in classes that use reading ability groups. Thus, reading groups appear to enhance average achievement, but they also magnify the gap between high- and low-SES students.

Model 4 adds the effects of instructional variables. The results show, as in mathematics, that all sorts of instruction boost achievement: whole language and phonics activities, whole-class and small-group settings, are all associated with higher achievement on average. The coefficient for full-day kindergarten is now $.681$, less than half its original size. Thus, Hypothesis 3 is partially supported in reading as in mathematics: although much is attributable to a greater quantity of instruction, a significant advantage for full-day programs remains unexplained by our models. In the prediction of between-class variation in student SES effects, in addition to the effect of reading groups, Model 4 indicates that more whole-language instruction is associated

with larger SES inequality, while better class behavior is related to smaller SES gaps. Overall, this model accounts for about 12 percent of the variance in average class achievement, and 21 percent of the variance in student SES effects, compared to a baseline model including only the control variables (see Table 3).

Discussion and Conclusions

To summarize the main results, we found more achievement growth in full-day than in half-day kindergarten classes, confirming Hypothesis 1. This effect did not interact with race, ethnicity, or socioeconomic status, so Hypothesis 2 was disconfirmed. Hypotheses 3 was partially supported: we observed more academic instruction in full-day classes, and this accounted for 31-54 percent of the full-day advantage, but much remains unexplained. Our results also supported Hypotheses 4 and 5, in that full-day classes devoted more time to reading groups, which promoted achievement growth, except for students assigned to low groups, who gained less than students in classes with no groups at all.

These results raise new questions. First, why did we fail to observe interactions of full-day kindergarten by social background? The benefits of full-day kindergarten extend equally to students of different racial and economic groups. This suggests that what the half-day students are doing outside of kindergarten – staying home or attending day care, playing, reading, or watching tv – is equally valuable or harmful on average for all demographic subgroups, so that the absence of full-day kindergarten and its associated instructional activities hurts all students the same. This interpretation contradicts the standard interpretation for differential learning over the summer (advantaged students

engage in cognitively stimulating activities more than disadvantaged students), but it may be that summer activities are qualitatively different than after-school activities.

A second new question is, what may account for the benefits of full-day kindergarten beyond the observed differences in instructional organization and activities? A likely candidate is that there are other aspects of instruction that our survey did not measure, and these also favored full-day programs. In addition, teacher reports of instruction are less than perfectly reliable, and it is possible that more reliable indicators, such as those from classroom observations, would account for the full-day effects more completely. Yet another possibility is that there is unobserved selectivity into the different types of classes. If so, this could account not only for the residual (unexplained) effects of full-day programs, but for all of the effects. We think this explanation is the least likely, for two reasons. First, our equations include a rich set of control variables for pre-existing conditions. These controls included not only a pre-test, but family background variables based on responses from parents. Differences among students should be referenced by the background variables, particularly by the pre-tests. Second, it is difficult to discern a pattern of unobserved selectivity that would be consistent with the observed variables. For example, full-day classes include students of higher SES, but black students are strongly overrepresented in the full-day classes, while Hispanic students are slightly underrepresented. Full-day classes are more often found in private schools, but our results showed no benefits of private schools for student achievement (and the coefficients were negative). Consequently, we think unobserved selectivity is an unlikely explanation for our results.

How important is the advantage of full-day kindergarten? Our original plan was to compare its effects to those of a prominent intervention, class size reduction. That plan was foiled when the ECLS kindergarten data failed to show any benefits to small classes (a matter we take up at length in another paper, see Milesi and Gamoran, 2003). Observed racial and ethnic gaps offer another standard of comparison. In mathematics, the full-day advantage is about 80 percent of the black-white gap in achievement growth, and more than double the difference in growth between Hispanics and non-Hispanic whites. In reading, the full-day advantage is larger than the black-white difference in achievement growth. That is, the average black student in a full-day program gains more than the average white in a half-day program. Thus, introducing full-day kindergarten in areas of concentrations of disadvantaged and minority students would help to mitigate achievement inequality, in some cases not only wiping out differential growth but actually reducing the overall gap.

If full-day kindergarten follows the path of class-size reform, what seems good for disadvantaged students will shortly be extended to everyone. According to our results, universal full-day kindergarten would raise achievement for all groups, but it would do little to reduce inequality. If anything, full-day programs for everyone would exacerbate existing inequalities, because blacks currently have more access to full-day programs than whites, a situation that at present prevents inequality from being as great as it otherwise might be.

As a final note, we cannot help but comment on the irony of studying the effects of kindergarten in a production function analysis. Nothing could be further from the original intentions of kindergarten programs! Instead of emphasizing a warm, nurturing

environment primarily designed to stimulate children's curiosity and their ability to interact with their physical and social environments, our paper fits its times by focusing on a narrow range of cognitive skills that can be demonstrated on tests. For whatever advantage we gain in test scores, what have we given up?

References

- Askey, R. (2001). Good intentions are not enough. Pp. 163-183 in T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* Washington, DC: Brookings.
- Barr, R. (1975). How children are taught to read: Grouping and pacing. *School Review*, 83, 479-498.
- Barr, R., and Dreeben, R. (1983). *How schools work*. Chicago: The University of Chicago Press.
- Bowman, D. (2003). Gov. Napolitano looks to business for help on all-day kindergarten. *Education Week*, January 22, p. 16.
- Brosterman, N. (1997). *Inventing kindergarten*. New York: Harry N. Abrams, Inc.
- Burrill, G. (2001). Mathematics education: The future and the past create a context for today's issues. Pp. 25-41 in T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* Washington, DC: Brookings.
- Chmelynski, C. (1998). All-day kindergarten on the rise. *Education Digest*, 64 (1), 32-34.
- Cohen, D. K., McLaughlin, M. W., and Talbert, J. E. (Eds.) (1993). *Teaching for understanding: Challenges for policy and practice*. San Francisco: Jossey-Bass.
- Coleman, J. S., Campbell, E. Q., Hobson, C., McPartland, J., Mood, A., Weinfield, F., and York, R. *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office, 1966.
- Entwisle, D., Alexander, K. L., and Olson, L. S. 1997. *Children, schools, and inequality*. Boulder, CO: Westview.
- Finn, J., and Achilles, C. M. (1999). "Tennessee's class size study: Findings, implications, and misconceptions." *Educational Evaluation and Policy Analysis*, 21, 97-109.
- Gamoran, A. 1986. Instructional and institutional effects of ability grouping. *Sociology of Education* 59: 185-198.
- Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education*, 60, 135-155.

- Gamoran, A. (1996). Effects of schooling on children and families. Pp. 107-114 in A. Booth and J. F. Dunn (Eds.), *Family-school links: How do they affect educational outcomes?* Hillsdale, NJ: Erlbaum.
- Gamoran, A., Anderson, C. W., Quiroz, P. A., Secada, W. G., Williams, T., and Ashmann, S. (2003). *Transforming teaching in math and science: How schools and districts can support change*. New York: Teachers College Press.
- Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.
- Kilpatrick, J., Swafford, J., and Findell, B., Editors. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.
- Loveless, T. (1997). The second great math rebellion. *Education Week*, October 15, 1997, p. 48, 36.
- Mare, R. D. (1995). Changes in educational attainment and school enrollment. Pp. 155-213 in R. Farley (Ed.), *The state of the union: America in the 1990s*. New York: Russell Sage Foundation.
- Meyer, J. W. (1977). The effects of education as an institution. *American Journal of Sociology*, 83, 55-77.
- Meyer, J. W. (1980). Levels of the educational system and schooling effects. In C. E. Bidwell and D. M. Windham (Eds.), *The analysis of educational productivity: Issues in macroanalysis*. Volume 2. Cambridge, MA: Ballinger.
- Milesi, C., and Gamoran, A. (2003). Effects of class size on student achievement in kindergarten. Madison, WI: Unpublished manuscript, University of Wisconsin, Department of Sociology.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics (1991). *Professional standards for teaching mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- NICHD Early Child Care Research Network.. (2002). Early child care and children's development prior to school entry. *American Educational Research Journal*, 39, 133-164.

- Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Second edition. Thousand Oaks, CA; Sage.
- Ravitch, D. (2001). It is time to stop the war. Pp. 210-228 in T. Loveless (Ed.), *The great curriculum debate: How should we teach reading and math?* Washington, DC: Brookings.
- Rice, J. K. (1999). "The impact of class size on instructional strategies and the use of time in high school mathematics and science courses." *Educational Evaluation and Policy Analysis*, 21, 215-229.
- Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, 40, 411-451.
- Rowan, B., and Miracle, Jr., A. J. (1983). Systems of ability grouping and the stratification of achievement in elementary schools. *Sociology of Education*, 56, 133-144.
- Shavit, Y., and Blossfeld, H.-P. (1993). *Persistent inequality: Changing educational attainment in thirteen countries*. Boulder, CO: Westview.
- Slavin, R. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57, 293-336.
- Snow, C. E., Burns, S., and Griffin, P., Editors. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- U.S. Department of Education. *Digest of Education Statistics, 2001*. Document no. NCES 2002-130. Washington, DC: 2002.
- Viadero, D. (2002). Study: Full-day kindergarten boosts academic achievement. *Education Week*, April 17, p.14.
- West, J., Denton, K., and Germino-Hausken, E. (2000). *America's kindergartners: Findings from the Early Childhood Longitudinal Study, kindergarten class of 1998-1999, fall 1998*. Document no. NCES-2000-070. Washington, DC: National Center for Education Statistics.