Available online at www.sciencedirect.com
ScienceDirect
Economics of Education Review 26 (2007) 67–82
Economics of Education Review
www.elsevier.com/locate/econedurev

Does full-day kindergarten matter? Evidence from the first two years of schooling

Philip DeCicca*

Department of Economics, McMaster University, Hamilton, Ont., Canada L8S 4M4

Received 28 February 2005; accepted 12 April 2005

Abstract

Over the past three decades, enrollment in full-day kindergarten has grown considerably—from roughly one-tenth to just over half of US kindergartners today. Full-day kindergarten reappeared first in the 1960s as an intervention designed to help disadvantaged children "catch up" to their peers through additional schooling. More recently, it has gained popularity among non-poor parents and schools, so that children presently enrolled in full-day programs are, on average, very similar to their half-day counterparts in baseline test scores as well as other child, parent and school characteristics. Using longitudinal data, I estimate the impact of full-day kindergarten on standardized test scores in mathematics and reading, as children progress from kindergarten to first grade. I find that full-day kindergarten has sizeable impacts on academic achievement, but the estimated gains are short-lived, particularly for minority children. Given the additional expense of full-day kindergarten, information regarding the size and duration of gains should be of great interest to policy makers.

JEL classification: I21; I28

Keywords: Human capital; Educational economics

1. Introduction

Over the past three decades, the fraction of US kindergartners enrolled in full-day programs has risen from roughly one-tenth to a slight majority (US Census, 2002). Despite the extra cost of providing full-day kindergarten, it remains popular where it exists and is growing in popularity where it does not. To date, there exists little systematic evidence regarding its possible effects on academic achievement and even less information on their persistence over time. Using longitudinal data, I examine the impact of full-day kindergarten attendance on standardized test scores in mathematics and reading, as children progress from kindergarten to first grade.

I find evidence that full-day kindergarten has positive and practically important effects on early human capital formation. However, the estimated differences are short-lived, as they fall dramatically over the course of an additional year, towards the end of first grade. This pattern is especially striking for black and Hispanic full-day kindergartners who

*Tel.: +905 525 9140x24129; fax: +905 521 8232.
E-mail address: decicca@mcmaster.ca.

0272-7757/$ - see front matter
doi:10.1016/j.econedurev.2005.04.003

see significant short-run advantages disappear more completely than their white peers. Hispanic children, in particular, exemplify this pattern. In what follows, "short-run" refers to end of kindergarten test scores and "longer-run" refers to end of first grade test scores.

In the next section, I provide background on the question of interest, briefly discussing the history of full-day kindergarten and reviewing the relevant literature. I also briefly discuss why full-day kindergarten might affect children's human capital formation and provide some sample comparisons between full and half-day kindergartners in order to understand how much they differ on a detailed set of observed characteristics. As will be seen, full and half-day kindergartners appear strikingly similar today, due principally to the historical evolution of full-day kindergarten. Section 3 discusses the data, emphasizing their particular advantages in producing meaningful estimates of the relationship of interest. The basic empirical strategy, which exploits the fact that kindergartners were given mathematics and reading tests at multiple points in time, is also presented in this section. Most importantly, tests were first administered to children shortly after entering kindergarten, prior to much exposure to the treatment of interest. Throughout the paper, I refer to child performance on these initial tests as *baseline* scores. In conjunction with companion tests given toward the ends of kindergarten and first grade, these tests are designed to allow assessment of child learning in reading and mathematics. Section 4 presents results from models estimated separately by race and also by race and gender in later specifications. Section 5 provides an important extension. In particular, I present evidence on *why* short-run effects disappear by focusing on *when* they disappear. Section 6 concludes the paper.

2. Background and motivation

2.1. The movement (back) to full-day kindergarten

In the United States, the history of kindergarten dates back to the late 1800s, when it began as a full-day program. Kindergarten grew in popularity through the early 20th century and remained full-day until the US entered World War II in the 1940s. During the war, the need for labor in war-time industries drew many women into the labor force—some for the first time and some from other employment sectors (Goldin, 1991). Among the latter category were teachers, who at that time were almost uniformly women, especially in the elementary grades. As part of the larger "war effort", schools across the country began to cut their kindergarten classes back to half-day in order to free up additional labor (Jones, 2002; Oelerich, 1984). After WWII, tremendous growth in the number of young children (i.e., the early Baby Boom cohorts) reinforced the trend towards half-day kindergarten (Jones, 2002; Ulrey, Alexander, Bender & Gillis, 1982).

Kindergarten retained its half-day character until the 1960s and 1970s when full-day kindergarten began to reemerge as a way of improving the academic preparation of children deemed "at-risk".[1] As such, the early re-adopters of full-day kindergarten tended to be poorer schools and school districts, serving predominantly minority children. The prevailing view was that full-day kindergarten, which typically consists of a 5–6 h school day, rather than the typical 2 and one-half to 3 h day in half-day kindergarten, would provide more opportunity for these children to "catch up" academically to their less disadvantaged peers. Over time, many non-poor districts, perhaps in partial response to the rise in dual-earner and single-parent families, began to implement full-day programs. Today, full-day kindergarten is the norm, albeit by a slight margin. As I discuss in great detail later in this section, the variegated reemergence of full-day kindergarten has led to full and half-day kindergartners who are presently very similar on many dimensions.

Much of the very recent push for full-day kindergarten has occurred at the state level. Presently, nine state governments mandate full-day kindergarten and 26 provide financial incentives to encourage school districts to provide it. Much of the enabling legislation occurred in the 1990s. For example, of the nine mandating states, seven initiated full-day kindergarten-only regimes after 1990. Other states may soon follow as evidenced by the comments of Nevada Governor Kenny Guinn in his 2003 State of the State address:

> "…to realize that vision we need to create a generation of young Nevadans with stronger, sharper, and more sophisticated skills. Therefore, I propose that we start at the beginning by

[1] One notable exception is Hawaii, which instituted full-day kindergarten in the 1950s (Gorton, 1968).

providing full-day kindergarten for our children.''

Full-day kindergarten remains popular, despite being much more expensive than the half-day variety. These higher costs are somewhat intuitive. Consider a school that is required to switch from traditional half-day to full-day kindergarten. Holding constant class size, the number of days per week that class meets, the length of the school year and other potentially offsetting (i.e., cost-saving) behavior, this school now requires twice as many teachers for a fixed number of pupils. In addition, full-day kindergarten offers fewer opportunities to share resources such as desks, books and computers relative to a half-day regime. Finally, full-day kindergarten may involve higher quasi-fixed costs. For example, a school might need additional classroom space since it will no longer be able to use the same room two times per day, as with a half-day regime. More concretely, Ohio's Office of Education Oversight (1997) estimates that full-day kindergarten costs over 70 percent more than traditional half-day kindergarten in terms of per pupil expenditure. The report notes that full-day kindergarten is not twice as expensive as half-day kindergarten primarily due to savings in transportation costs.[2]

2.2. *Why might full-day kindergarten matter?*

The rationale for full-day kindergarten is simple: the more time children spend in school the more they will learn. In turn, it is thought that this additional learning will lead to improved academic outcomes as children move into later grades. In the context of a human capital accumulation story, early investments reduce the cost of future ones, so a larger initial or early stock of human capital has the potential to influence later, perhaps even adult, levels of human capital. More generally, if one fails to learn the ''basics'', it may inhibit subsequent learning so that ''catching up'' is prohibitively costly. As noted by Nobel laureate Gary S. Becker in his 1989 Ryerson Lecture:

> ''Large differences among young children grow over time with age and schooling because children learn more easily when they are better prepared. Therefore, even small differences among children…are frequently multiplied over time into large differences when they are teenagers.''

In the present case, children who leave kindergarten with relatively better reading skills are likely to be more successful in learning new material in the first grade and beyond, especially since the material taught in early elementary school tends to be sequential in nature (Siefert, 1993). Mathematics provides perhaps an even better example. Mathematical learning tends to be quite sequential in nature, so if one masters the basic concepts early it is likely that the burden of future learning will be lowered. To the extent that learning at this level is indeed sequential, it is possible that full-day kindergarten has effects that persist over time. In this spirit, recent work links the introduction of kindergartens in the South to increased educational attainment for blacks (Cascio, 2003). While not directly related to the present question, it suggests that additional early childhood education may have persistent impacts since final educational attainment is realized much later, typically in late adolescence or early adulthood. As described in Section 3, the longitudinal nature of my data allows me to test for persistence, albeit over a much shorter amount of time.

2.3. *Related literature*

The preceding logic implicitly assumes that extra ''seat time'' provided by full-day kindergarten is devoted to learning activities. Available evidence lends support to this assumption. For example, Hough and Bryde (1996) and Elicker and Mathur (1997) show that teachers in full-day kindergarten settings spend more time with children individually and in small groups, relative to teachers in half-day programs. As a result, full-day kindergartens are more able to integrate the on-going ''push down'' of academic material traditionally presented in the first grade since they have more time to incorporate these concepts and/or related learning activities. Full-day kindergarten, however, is not without its detractors: some warn that an early emphasis on academic learning, at the expense of the traditional play-based curriculum, may harm children emotionally and, consequently, academically as well (Natale, 2001; Olsen & Zigler, 1989).

[2]In particular, schools must only bus children two times per day under a full-day regime, instead of four times per day when operating morning and afternoon sessions under half-day kindergarten.

While there has been considerable attention in the non-economics education literature, economists have shown little interest in the full vs. half-day kindergarten debate. The existing literature suggests that full-day kindergarten's impact on academic and social outcomes is somewhat mixed, but taken as a whole tends to imply that full-day kindergarten's pros outweigh its cons. Some studies find relatively large gains (e.g., Gullo, 2000), while others do not (Karweit, 1992). It should be noted, however, that existing work has several limitations. For example, several studies focus on the experiences of a particular school or school district, some focus on particular types of students and others ignore concerns related to non-random sorting.

Given the abundance of closely related work, the lack of attention by economists is surprising. In particular, much attention has been given to the pre-kindergarten program, Head Start, and its possible effects on child academic, social and health outcomes. As recently as the middle and late 1990s, economists studying Head Start have linked participation with higher test scores, reduced grade repetition and an increased likelihood of receiving recommended childhood immunizations (Currie & Thomas, 1995, 1999). Even more recently, Garces, Thomas, and Currie (2002) find evidence that Head Start effects persist into early adulthood. Many of the reasons that one might suspect Head Start to improve child outcomes also apply to full-day kindergarten.[3]

2.4. Comparing full and half-day kindergartners

A common concern in a study such as this is that "treated" and "untreated" individuals differ on unobserved traits that are correlated with treatment status and also independently affect the outcome of interest. To the extent that this is true, traditional estimators may yield misleading estimates of program effects. In the present context, the concern is that children who attend full-day kindergarten differ appreciably, and unobservably, from half-day children. And while it is impossible to compare children on the basis of unobserved characteristics, some observable comparisons may be suggestive of the degree of non-random sorting into kindergarten type. At a minimum, one can ascertain if the two groups are so different as to make comparison very difficult or perhaps even meaningless.

To reduce unobserved heterogeneity, I restrict my sample to children enrolled in public schools, who are first-time kindergartners and who did not change schools during kindergarten. I also perform the analyses by race group to further reduce this sort of heterogeneity (e.g., unobserved school quality). Table 1 provides several comparisons by kindergarten type for white, black and Hispanic children, respectively.[4] The first two columns of Table 1 show that among white children, the two types of kindergartners are strikingly similar. Perhaps most importantly, mean baseline mathematics and reading scores, as well as their corresponding standard deviations, are virtually identical and not statistically different from each other. In particular, the mean half-day math score is 1.2 percent higher than the full-day average and the corresponding reading score difference is less than 0.05 percent, with respective absolute t-ratios of 1.26 and 0.05. Since tests are designed to yield normally distributed scores, the sample moments in Table 1 suggest that white full and half-day kindergartners belong to very similar baseline test score distributions. Beyond baseline scores, the fraction of children residing with two parents, the fraction whose families ever participated in AFDC/TANF and the fraction of kindergartners with working mothers are all quite similar. There are more substantial differences with respect to family income and parents' education, though these suggest a degree of "negative" selection into full-day kindergarten, but the differences are, practically speaking, small.[5]

Among black and Hispanic children, the means suggest that the two groups are slightly less similar than full and half-day whites. As seen in the middle

[3]Note, however, that Head Start programs often involve non-academic interventions (e.g., promotion of available social services or parenting skills classes), while full-day kindergarten focuses more narrowly on classroom learning. As a result, I consider only learning-related outcomes, unlike the Head Start literature discussed above. Note also that all models with covariates include controls for prior Head Start participation.

[4]The samples used to produce the figures in Table 1 correspond exactly to the samples upon which model estimates in Tables 2–4, which I present in Section 4, are generated.

[5]Also, while there is little difference in maternal employment between birth and kindergarten, mothers of white children in full-day kindergarten are more likely to work 35 or more hours per week. However, the direction of causality may be from kindergarten type to full vs. part-time work (cf., Lemke, Witte, Querault, & Witt, 2000). Much smaller employment differences are seen for the mothers of minority children. All models estimated include a series of indicators for mother's employment status.

Table 1
Selected sample characteristics, by kindergarten type (means & standard deviations)

	White		Black		Hispanic	
	HDK	FDK	HDK	FDK	HDK	FDK
Baseline mathematics score	21.03 (7.19)	20.79 (7.14)	16.45 (5.50)	16.39 (5.12)	17.10 (6.18)	17.43 (6.03)
Baseline reading score	23.02 (8.26)	23.03 (8.12)	19.14 (6.78)	19.72 (6.45)	19.13 (7.20)	19.90 (7.76)
Family income (in 1000s)	64.1 (55.3)	56.3 (56.9)	30.7 (25.5)	25.4 (27.5)	40.8 (30.0)	37.1 (33.4)
Child age (in months)	68.5 (4.1)	68.8 (4.0)	67.5 (3.9)	68.1 (4.0)	67.2 (3.9)	68.6 (4.0)
Mother's age (in years)	34.0 (5.7)	33.2 (5.9)	31.5 (7.8)	31.4 (7.7)	31.6 (6.3)	31.8 (6.4)
At least one parent HS grad or higher	0.97	0.96	0.89	0.84	0.82	0.83
At least one parent college grad or higher	0.39	0.34	0.15	0.10	0.15	0.14
Mother in labor force	0.70	0.74	0.79	0.84	0.71	0.70
Mother works 35+ hours per week	0.41	0.47	0.57	0.58	0.46	0.49
Mom worked between birth & kindergarten	0.76	0.78	0.82	0.79	0.71	0.68
Family ever received AFDC/TANF	0.12	0.13	0.50	0.50	0.28	0.26
Parent expects kid college grad or higher	0.74	0.73	0.73	0.72	0.77	0.80
Child resides in two parent family	0.86	0.86	0.44	0.39	0.78	0.71
Ever participated in Head Start	0.07	0.10	0.38	0.45	0.21	0.22
Child disabled	0.15	0.15	0.13	0.12	0.10	0.13
Child born 2+ weeks premature	0.17	0.16	0.18	0.19	0.17	0.17
Sample size	2944	2615	323	1122	608	552

Notes: Samples consist of public school children who are first-time kindergartners and did not change schools during kindergarten. Additional restrictions make these samples correspond exactly to those that generate regression model estimates in Tables 2–4. The only exception is household income, which is based on slightly smaller samples (5075 for whites, 1216 for blacks and 1012 for Hispanics) due to missing information. Note that differences in baseline math and reading scores between full and half-day kindergartners are not statistically different from zero at conventional levels for any race group (for whites: $t = -1.26$ and 0.05, for blacks: $t = -0.17$ and 1.42, and for Hispanics: $t = 0.92$ and 1.75, for math and reading, respectively). FDK is full-day kindergarten and HDK is half-day kindergarten. Standard deviations are in parentheses for non-binary variables.

two columns of Table 1, black full-day kindergartners enjoy a 3 percent advantage in the baseline reading test score, but have an average baseline mathematics score which is 0.4 percent lower than their half-day peers. Neither difference is statistically significant at conventional levels ($t = 0.17$ and 1.42, respectively). While the reading-score advantage suggests some degree of positive selection for black full-day kindergartners, the education level and income of the parents of black full-day kindergartners imply a slight degree of negative selection. With respect to baseline test scores, the degree of positive selection into full-day kindergarten is greatest among Hispanic children. The last two columns of Table 1 show that Hispanic full-day kindergarteners have baseline reading scores that are 4 percent higher than their half-day peers. However, corresponding mathematics scores are only 1.9 percent higher. Once again, neither difference in means is statistically different from zero at conventional levels, though the reading score difference could be considered marginally significant ($t = 0.92$ and 1.75 for math and reading, respectively). As with black children, there is also some evidence that full-day kindergartners of Hispanic origin are negatively sorted into this status. So, while minority full and half-day kindergartners are not as similar as white full and half-day children, the differences are relatively minor.

3. Data and empirical strategy

3.1. Key variables

The variables of greatest interest are the math and reading test scores, which are my outcomes of interest, and kindergarten status. As noted earlier, these tests were administered to children at multiple points in time. Of greatest importance to this study is that tests were given early in the first year of schooling, before much exposure to the kindergarten curriculum. Following these baseline tests, students were then re-assessed at two later points in time—towards the end of the kindergarten year and towards the end of first grade. This testing structure was designed explicitly to measure children's longitudinal gains in subject-specific achievement (NCES, 2002b). Since the amount of time between tests may influence achievement gains, all models include separate month of assessment indicators for initial and subsequent tests.

Though the tests differ over time, they contain common and overlapping elements. The mathematics tests assess number recognition, counting, comparing and ordering numbers, solving word problems and interpreting simple graphs. The reading tests include questions to assess the basic reading skills, vocabulary/word comprehension, knowledge of the alphabet, phonetics, listening skill and reading comprehension.[6] While a variety of scores based on these tests are available in the ECLS-K, the analysis presented here uses Item Response Theory (IRT)-adjusted scores rather than, for example, the raw number of correct answers provided. These particular scores adjust for the fact that the tests were not standardized per se, but instead asked different questions to different children, depending on their answers to a set of initial "routing" questions. This sort of adaptive testing is considered by psychometricians to be more efficient compared to pure standardized testing, where all students take the same examination, and also reduces the potential for "ceiling" and "floor" effects which can affect the measurements of gains over time (Lord, 1980). See the *ECLS-K Psychometric Report* and/or Chapter 3 of the *ECLS-K User's Guide*, especially pages 3-2 through 3-5, for more detailed information on these tests including test validity, reliability, differential item response and test-taker motivation (NCES, 2002b, 2002c).

Kindergarten type is the covariate of greatest interest. According to national data, nearly 60 percent of all kindergarteners were enrolled in full-day kindergarten in 2000 (US Census, 2002). In the ECLS-K, which sampled children who entered kindergarten in academic year 1998–1999, the corresponding fraction is approximately 53 percent. This slight difference may arise from the fact that the Census report, based on data from the October 2000 Current Population Survey, includes private schools while my sample includes only public school students. Indeed, the fraction enrolled in full-day kindergarten rises to 57 percent when I include private school children.

At the school-level, 92 percent of public schools in my sample (656 out of 714) offer either full or half-day kindergarten. Of these 714 schools, 348 (49 percent) provide full-day kindergarten and 308 (43 percent) provide half-day kindergarten. The remaining 58 schools (8 percent) provide both types. This preponderance of one type or the other, as opposed to offering both kindergarten types, is consistent with what can be gleaned from the literature on full-day kindergarten. More importantly, it suggests that within-school sorting is not widespread. That is, to the extent that kindergarten type is determined at the school-level and there is not as much scope for choosing full vs. half-day kindergarten within a particular school, there are fewer possibilities for non-random sorting since families would have to change (public) schools to change kindergarten type. Later, as a robustness check, I estimate models that exclude children in schools with both types of kindergarten and find very little difference relative to models that include children in all three school types. Finally, note that the "all or nothing" nature of full-day kindergarten effectively precludes the inclusion of school fixed effects in my models.

3.2. Empirical strategy

The ECLS-K was designed, in part, to assess the value-added of kindergarten (NCES, 2002b). As such, standardized tests in math and reading were administered to nearly all sample members near the beginning and end of kindergarten. The existence of test scores prior to much exposure to schooling provides a baseline of the child's ability in these subject areas. Subsequent test scores, measured near the ends of kindergarten, first, third and fifth grades, provide the opportunity to assess both short and longer-run impacts of full-day kindergarten. I examine end-of-kindergarten and end-of-first grade scores in this paper.

Given the longitudinal nature of this information, the preferred statistical approach is one that accounts for individual heterogeneity. To fix ideas, consider test scores as an outcome of interest. If test score is regressed on kindergarten type in a standard cross-sectional model, a reasonable concern is omitted variable bias resulting from non-random sorting. For example, perhaps full-day children possess greater "readiness to learn" than their half-day counterparts *prior to* entering kindergarten,

[6]The tests were developed especially for the ECLS-K, but are based largely on existing and generally accepted instruments including the Children's Cognitive Battery (CCB), Peabody Individual Achievement Test—Revised (PIAT-R), Peabody Picture Vocabulary Test-3 (PPVT-3), Primary Test of Cognitive Skills (PTCS) and the Woodcock–Johnson Psycho-Educational Battery—Revised (WJ-R).

but this fact is not captured in the model. If so, the estimated impact of full-day kindergarten is likely to be overstated. However, if the impact of this unobserved difference is relatively constant over time, its effects should be embedded in the baseline test scores included in ECLS-K data. Hence, examining test score "growth" over time should eliminate, or at least mitigate, any associated bias.

I take a conceptually similar approach, controlling for the appropriate baseline test score as a right-hand side variable. More precisely, I estimate variants of the following general specification:

$$\mathrm{TS}_{i,t+j} = \Psi \mathrm{TS}_{it} + \gamma \mathrm{FDK}_{it} + \alpha C_{it} + \delta F_{it} + \theta S_{it} + \varepsilon_{i,t+j}, \quad (1)$$

where TS is subject-specific test score, FDK is a full-day kindergarten indicator, C is a set of child-specific variables, F is a set of family specific variables and S is a set of school and classroom-specific variables and "j" determines whether the model estimates short or longer-run effects. That is, $j=1$ refers to end of kindergarten and $j=2$ refers to end of first grade. All models are estimated via ordinary least squares.

Empirically, my goal is to estimate γ in the presence of this baseline score.[7] In essence, this is a more flexible way of estimating test score growth. Indeed, if ψ equals one, the model specified above would be equivalent to differencing the dependent variable, which would correspond to a difference-in-differences approach. Though not reported, I find that estimates from a difference-in-differences model without covariates are nearly identical to estimates derived from my Eq. (1) with the restriction that $\alpha=\delta=\theta=0$. These estimates are available from me upon request.

A couple of related items deserve mention. First, in practice, estimates of ψ are always relatively close to one. However, since ψ's are estimated very precisely, they are often statistically different from one, indicating that differencing the dependent variable may not be appropriate. Hence, I retain the more flexible specification implied by Eq. (1). Second, in all models presented, covariates are measured as of the initial survey wave out of necessity. While I observe test scores for the entire sample at three different points in time, I cannot time difference covariates since only a small subset are measured in all three relevant waves of the ECLS-K. While certainly a methodological weakness, the relatively short period of time between the beginning of kindergarten and the end of first grade (roughly 18–20 months), and the even shorter period between the beginning and end of kindergarten (roughly 9 months), imply that such differencing would result primarily in noisy measures of intertemporal change.

Finally, in addition to kindergarten type, several other potentially relevant school-level characteristics are available in the ECLS-K. I use information on school size, class size, length of the school year, number of days per week school is in session, public school type (i.e., regular, school of choice or magnet school) and whether the child's parents chose their current residence on the basis of the local schools. To the extent that these school-level features are correlated with the type of kindergarten offered, and exert an independent effect on achievement, they represent important controls. Further, schools, especially those required to provide more expensive full-day kindergarten, may offset additional costs by increasing class size, reducing the length of the school year or the number of days class meets per week, etc., if they are able to do so.[8] However, to the extent that such offsetting behavior is directly attributable to full-day kindergarten, ignoring such school responses involves estimating a more general effect. In other words, the two estimates are conceptually different. Therefore, as noted, I estimate models with and without school-level covariates. In all cases, robust standard errors are adjusted for clustering at this level of aggregation.

3.3. Analysis sample

While four waves of the ECLS-K were gathered by the end of first grade, I use only three—fall kindergarten (1998), spring kindergarten (1999) and spring first grade (2000)—since the fourth, fall first grade (1999), deliberately sampled only 30 percent

[7] I also estimate models with higher-order terms in baseline test score to allow for a more flexible functional form than is implied by Eq. (1) which imposes a linear relationship between baseline and subsequent test scores. Inclusion of these higher order terms had no appreciable impact on the relevant estimate of γ.

[8] In my analysis sample, which I describe in great detail in the next sub-section, I find only small differences in class size (20 vs. 21 students per class) and length of school year (177 vs. 178 days per year) for full and half-day kindergartners, respectively. With respect to days of class per week, however, there is some evidence of offsetting behavior. In particular, while only 1 percent of half-day children attend school less than 5 days per week, roughly 6 percent of full-day children do so.

of original respondents. Restricting my sample to those with math and reading scores in all three waves, discernible kindergarten type, and a completed initial parent questionnaire yields a sample of 13,025 children. Further restricting the sample to public school children who are first-time kindergarteners and do not switch schools during the kindergarten year leaves a sample of 9632 children. Since analysis is done separately by race, I limit the sample to those races with enough sample size to support estimation by gender. Doing so produces an analysis sample of 8599 children, of whom 5785 are white, 1583 are black and 1231 are Hispanic.

As previously mentioned, I estimate models with and without school-level covariates. Without the school-level characteristics, complete case analysis yields a sample of 7303, which includes 5075 white, 1216 black and 1012 Hispanic children. With the school-level covariates included, the corresponding sample size drops to 5734, including 4189 white, 861 black and 684 Hispanic children. Most of the reduction in sample size is due to missing information on household income and school characteristics.[9] To understand the potential sensitivity of my results to the impact of missing data, I include separate binary indicators for those individuals with missing information on these covariates.[10] This increases sample size to 8164, which includes 5559 white, 1445 black and 1160 Hispanic children. These numbers represent 96 percent (5559 of 5785), 91 percent (1445 of 1583) and 94 percent (1160 of 1231) of the white, black and Hispanic analysis samples, respectively. Since estimates, and their pattern, are very similar across the different samples, I report results from models that use this latter, and most complete, sample.

4. Results

In what follows, I refer frequently to short and longer-run estimates of the impact of full-day kindergarten. Recall that *short-run* refers to performance on an end-of-kindergarten reading or math test and that *longer-run* refers to performance on a similar test, designed to evaluate academic progress, roughly 1 year later. Since I estimate models separately by race group, I report estimates in a similar fashion. In each race-specific sub-section, I present short-run results followed by corresponding longer-run estimates. Further, since the literature on child cognitive development suggests the possibility of differences in the learning patterns of girls and boys, I also present estimates by gender (cf., Carr & Jessup, 1997; Fennema & Sherman, 1977).

Tables 2–4 present regression estimates from three different specifications—Column 1 presents results from a simple regression of test score on kindergarten type, Column 2 presents a regression that adds an extensive set of covariates, but not the subject-specific baseline test score, and Column 3 presents a model that adds the appropriate baseline test score to the model in Column 2.[11] All specifications are variants of Eq. (1). In all cases, Column 3 contains my preferred specification. These three tables each include two tables (e.g., Table 2A and B), where the former displays short-run estimates and the latter presents longer-run estimates. Both sets of estimates are generated from balanced samples, though corresponding results are very similar when generated from unbalanced samples. Table 5 presents short and longer-run estimates by gender, but only for the preferred specification.

Given the large volume of estimates, I limit discussion to the estimated impacts of full-day kindergarten, sometimes to the exclusion of other potentially interesting results. In all cases, estimated differences in the performance of full-day kindergartners relative to their half-day counterparts are presented as a percentage of the standard deviation of the model-specific dependent variable and are discussed in terms of ‘‘gains’’ or ‘‘advantages’’. Relevant tables also include estimated differences as a percentage of the model-specific mean. Discussion centers on estimates from my most preferred

[9]Of the analysis sample (8599 children), about 11 percent are missing household income information and about 12 percent are missing various school characteristics. No other covariate is missing for more than 3 percent of cases.

[10]As part of this strategy, I converted the variables household income, length of school year, number of days per week school meets and class size each into a series of discrete indicators.

[11]This structure allows the reader to compare the estimated effect of full-day kindergarten via a simple conditional mean difference, a naïve model that does not control for the appropriate baseline test score, and finally a model that also includes the baseline score. While my data are not experimental by design, estimates of the effect of full-day kindergarten are very close across these three specifications, especially for short-run estimates. Each of these tables contains results from a Hausman test for the equality of full-day kindergarten coefficients in Models 2 and 3 relative to Model 1, the simple regression. See pp. 338–342 of Johnston and DiNardo (1997) for details. Finally, note that I fail to reject the equality of full-day kindergarten coefficients for Model 1 vs. Model 3 in 11 of 12 cases.

Table 2
Short- and longer-run regression estimates for white children

	Mathematics			Reading		
Selected covariates	1	2	3	1	2	3
(A) Short-run regression estimates—*Dependent variable: End-of-kindergarten test score*						
Full-day kindergarten (FDK)	0.979***	1.830***	1.437***	1.754***	2.444***	1.843***
	(0.347)	(0.316)	(0.241)	(0.441)	(0.393)	(0.263)
	[11.7%]	[21.8%]	[17.1%]	[17.8%]	[24.8%]	[18.7%]
	{3.3%}	{6.2%}	{4.9%}	{5.3%}	{7.4%}	{5.5%}
Maximum of parent education						
Less than HS	—	−2.273*** (0.651)	−0.822** (0.387)	—	−2.551*** (0.680)	−1.127** (0.451)
Some college	—	1.002*** (0.269)	−0.076 (0.180)	—	1.241*** (0.312)	0.286 (0.206)
College degree	—	2.644*** (0.343)	0.090 (0.220)	—	3.247*** (0.427)	0.473* (0.260)
Master's degree or higher	—	4.177*** (0.423)	0.381 (0.306)	—	5.334*** (0.523)	0.888*** (0.317)
Child female	—	−0.132 (0.207)	−0.187 (0.143)	—	1.877*** (0.232)	0.605*** (0.155)
Baseline test score	—	—	0.893*** (0.014)	—	—	0.929*** (0.012)
R^2	0.005	0.22	0.66	0.01	0.20	0.67
Dependent mean/standard deviation	29.50/8.40	29.50/8.40	29.50/8.40	33.21/9.87	33.21/9.87	33.21/9.87
Hausman test (*p*-values)	—	0.02	0.38	—	0.24	0.63
Sample size	5559	5559	5559	5559	5559	5559

Notes: Sample restricted to white public school students who are first-time kindergartners and do not switch schools during kindergarten. Model 1 is a simple regression. Models 2 and 3 are identical except Model 3 includes the appropriate baseline test score. With the exception of Model 1, all models also include household income (7 categories), child age (in months), child disability status, family type (4 categories), household size, ever in Head Start, mother's age, mother's current work status (4 categories), day care status, child birth weight (in ounces), whether child born premature, parental educational expectations (6 categories), urbanicity (7 categories), region (4 categories), number of students in kindergarten classroom (4 categories), number of days per week school meets (3 categories), number of days in school year (3 categories), whether magnet school, whether school of choice, whether parents chose residence for school-related reasons, school enrollment (6 categories), and indicators corresponding assessment months. Each Hausman test compares the FDK coefficient of Model 2 or 3 to the Model 1 estimate; corresponding test statistic is distributed $\chi^2(1)$. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets, for FDK coefficient only. Robust standard errors, adjusted for clustering at the school level, in parentheses.

	Mathematics			Reading		
Selected covariates	1	2	3	1	2	3
(B) Longer-run regression estimates—*Dependent variable: End-of-1st grade test score*						
Full-day kindergarten (FDK)	0.550*	0.995***	0.656**	0.551	1.325***	0.711
	(0.331)	(0.307)	(0.264)	(0.553)	(0.502)	(0.437)
	[6.5%]	[11.7%]	[7.7%]	[4.3%]	[10.3%]	[5.5%]
	{1.2%}	{2.2%}	{1.4%}	{1.0%}	{2.3%}	{1.2%}
Maximum of parent education						
Less than HS	—	−2.358*** 0.702)	−1.149** (0.574)	—	−4.530*** (0.986)	−3.074*** (0.851)
Some college	—	1.565*** (0.311)	0.677*** (0.251)	—	1.640*** (0.439)	0.692* (0.354)
College degree	—	3.227*** (0.359)	1.123*** (0.284)	—	4.064*** (0.540)	1.311*** (0.423)
Master's degree or higher	—	4.751*** (0.404)	1.624*** (0.345)	—	6.066*** (0.656)	1.650*** (0.540)
Child female	—	−0.375* (0.213)	−0.423** (0.178)	—	2.450*** (0.314)	1.176*** (0.266)
Baseline test score	—	—	0.739*** (0.017)	—	—	0.928*** (0.019)
R^2	0.003	0.19	0.48	0.004	0.18	0.46
Dependent mean/standard deviation	45.42/8.49	45.42/8.49	45.42/8.49	57.82/12.82	57.82/12.82	57.82/12.82
Hausman test (*p*-values)	—	0.05	0.51	—	0.06	0.65
Sample size	5559	5559	5559	5559	5559	5559

Notes: Sample restricted to white public school students who are first-time kindergartners and do not switch schools during kindergarten. Model 1 is a simple regression. Models 2 and 3 are identical except Model 3 includes the appropriate baseline test score. With the exception of Model 1, all models also include the extensive set of covariates listed in Table 2A. Each Hausman test compares the FDK coefficient of Model 2 or 3 to the Model 1 estimate; the corresponding test statistic is distributed $\chi^2(1)$. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets, for FDK coefficient only. Robust standard errors, adjusted for clustering at the school level, are in parentheses. *= $0.10 \geqslant p > 0.05$, **= $0.05 \geqslant p > 0.01$ and ***= $p \leqslant 0.01$.

Table 3
Short- and longer-run regression estimates for black children

	Mathematics			Reading		
Selected covariates	1	2	3	1	2	3
(A) Short-run regression estimates—*Dependent variable: End-of-kindergarten test score*						
Full-day kindergarten (FDK)	0.947*	1.214***	0.842**	1.423**	1.636***	0.996*
	(0.530)	(0.466)	(0.350)	(0.688)	(0.626)	(0.534)
	[13.0%]	[16.7%]	[11.1%]	[15.8%]	[18.2%]	[11.1%]
	{4.0%}	{5.2%}	{3.6%}	{5.0%}	{5.7%}	{3.5%}
Maximum of parent education						
Less than HS	—	−0.743 (0.499)	−0.039 (0.399)	—	−0.945 (0.597)	0.197 (0.436)
Some college	—	2.062*** (0.446)	0.641* (0.328)	—	2.431*** (0.551)	1.033*** (0.354)
College degree	—	1.769** (0.716)	0.248 (0.495)	—	2.964*** (1.001)	0.395 (0.707)
Master's degree or higher	—	2.704* (1.398)	1.164 (0.921)	—	5.427*** (1.601)	2.733*** (1.035)
Child female	—	0.165 (0.358)	−0.026 (0.252)	—	1.043** (0.451)	0.343 (0.310)
Baseline test score	—	—	1.011*** (0.031)	—	—	0.990*** (0.036)
R^2	0.008	0.22	0.61	0.01	0.22	0.62
Dependent mean/ standard deviation	23.40/7.26	23.40/7.26	23.40/7.26	28.52/8.99	28.52/8.99	28.52/8.99
Hausman test (p-values)	—	0.54	0.81	—	0.71	0.45
Sample size	1445	1445	1445	1445	1445	1445
(B) Longer-run regression estimates—*Dependent variable: End-of-1st grade test score*						
Full-day kindergarten (FDK)	−0.177	0.092	−0.312	−0.099	−0.266	−1.044
	(0.701)	(0.625)	(0.544)	(1.050)	(1.111)	(1.003)
	[−2.0%]	[1.1%]	[−3.6%]	[−0.7%]	[−2.0%]	[−7.8%]
	{−0.5%}	{0.2%}	{−0.8%}	{−0.2%}	{−0.5%}	{−2.1%}
Maximum of parent education						
Less than HS	—	−0.701 (0.620)	0.045 (0.524)	—	−2.122** (0.921)	−0.819 (0.774)
Some college	—	2.517*** (0.561)	1.067** (0.425)	—	3.493*** (0.797)	1.781*** (0.635)
College degree	—	2.979*** (0.847)	1.442** (0.614)	—	3.534*** (1.315)	0.460 (1.109)
Master's degree or higher	—	4.134*** (1.425)	2.601** (1.265)	—	8.346*** (2.191)	5.061*** (1.798)
Child female	—	0.075 (0.418)	−0.088 (0.329)	—	2.539*** (0.718)	1.727*** (0.597)
Baseline test score	—	—	1.041*** (0.044)	—	—	1.171*** (0.063)
R^2	0.003	0.21	0.51	0.007	0.19	0.45
Dependent mean/ standard deviation	38.70/8.69	38.70/8.69	38.70/8.69	50.36/13.32	50.36/13.32	50.36/13.32
Hausman test (p-values)	—	0.62	0.80	—	0.84	0.25
Sample size	1445	1445	1445	1445	1445	1445

Notes: Sample restricted to black public school students who are first-time kindergartners and do not switch schools during kindergarten. Model 1 is a simple regression. Models 2 and 3 are identical except Model 3 includes the appropriate baseline test score. With the exception of Model 1, all models also include the extensive set of covariates listed in Table 2A. Each Hausman test compares the FDK coefficient of Model 2 or 3 to the Model 1 estimate; the corresponding test statistic is distributed $\chi^2(1)$. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets, for FDK coefficient only. Robust standard errors, adjusted for clustering at the school level, are in parentheses. *= $0.10 \geqslant p > 0.05$, **= $0.05 \geqslant p > 0.01$ and ***= $p \leqslant 0.01$.

specification. Finally, since there is little difference in the estimated impact of full-day kindergarten with and without them, all models I discuss include school-level covariates.

4.1. White children

Short-run estimates, contained in Table 2A, imply that white full-day kindergarteners outscore

Table 4
Short- and longer-run regression estimates for Hispanic children

Selected covariates	Mathematics 1	Mathematics 2	Mathematics 3	Reading 1	Reading 2	Reading 3
(A) Short-run regression estimates—Dependent variable: End-of-kindergarten test score						
Full-day kindergarten (FDK)	1.381**	1.237**	1.219***	2.328***	3.155***	2.278***
	(0.578)	(0.605)	(0.436)	(0.695)	(0.810)	(0.623)
	[17.7%]	[15.8%]	[15.6%]	[24.4%]	[33.1%]	[23.9%]
	{5.4%}	{4.9%}	{4.8%}	{7.7%}	{10.5%}	{7.5%}
Maximum of parent education						
Less than HS	—	−1.285** (0.571)	0.066 (0.400)	—	−0.969 (0.749)	−0.085 (0.584)
Some college	—	1.848*** (0.491)	0.758** (0.343)	—	3.068*** (0.680)	0.731 (0.452)
College degree	—	3.235*** (0.807)	0.511 (0.565)	—	4.092*** (1.002)	0.787 (0.678)
Master's degree or higher	—	2.176* (1.231)	−0.449 (0.810)	—	3.391** (1.591)	−0.648 (1.019)
Child female	—	0.190 (0.404)	−0.084 (0.272)	—	2.096*** (0.542)	0.726** (0.370)
Baseline test score	—	—	0.957*** (0.025)	—	—	0.937*** (0.029)
R^2	0.02	0.27	0.67	0.03	0.23	0.63
Dependent mean/standard deviation	25.49/7.81	25.49/7.81	25.49/7.81	30.18/9.52	30.18/9.52	30.18/9.52
Hausman test (p-values)	—	0.77	0.74	—	0.15	0.93
Sample size	1160	1160	1160	1160	1160	1160
(B) Longer-run regression estimates—Dependent variable: End-of-1st grade test score						
Full-day kindergarten (FDK)	0.695	−0.633	−0.730	0.961	0.826	−0.141
	(0.578)	(0.646)	(0.554)	(0.695)	(1.139)	(1.064)
	[8.3%]	[−7.6%]	[−8.7%]	[7.4%]	[6.4%]	[−1.1%]
	{1.7%}	{−1.5%}	{−1.8%}	{1.8%}	{1.5%}	{−0.3%}
Maximum of parent education						
Less than HS	—	−1.214* (0.731)	−0.067 (0.638)	—	−2.851*** (1.067)	−1.858** (0.920)
Some college	—	1.557*** (0.577)	0.647 (0.489)	—	3.732*** (0.922)	1.194 (0.747)
College degree	—	3.158*** (0.842)	0.817 (0.674)	—	6.404*** (1.313)	2.776** (1.132)
Master's degree or higher	—	2.367* (1.300)	0.058 (1.115)	—	2.540 (2.090)	−1.886 (1.582)
Child female	—	−0.376 (0.441)	−0.584 (0.366)	—	2.954*** (0.757)	1.520** (0.604)
Baseline test score	—	—	0.802*** (0.039)	—	—	0.997*** (0.050)
R^2	0.007	0.22	0.47	0.005	0.20	0.45
Dependent mean/standard deviation	41.94/8.37	41.94/8.37	41.94/8.37	53.46/12.98	53.46/12.98	53.46/12.98
Hausman test (p-values)	—	0.01	0.02	—	0.87	0.19
Sample size	1,160	1,160	1,160	1,160	1,160	1,160

Notes: Sample restricted to Hispanic public school students who are first-time kindergartners and do not switch schools during kindergarten. Model 1 is a simple regression. Models 2 and 3 are identical except Model 3 includes the appropriate baseline test score. With the exception of Model 1, all models also include the extensive set of covariates listed in Table 2A. Each Hausman test compares the FDK coefficient of Model 2 or 3 to the Model 1 estimate; the corresponding test statistic is distributed $\chi^2(1)$. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets, for FDK coefficient only. Robust standard errors, adjusted for clustering at the school level, are in parentheses. *= $0.10 \geqslant p > 0.05$, **= $0.05 \geqslant p > 0.01$ and ***= $p \leqslant 0.01$.

their half-day counterparts in both mathematics and reading. In mathematics, I estimate that full-day kindergartners have roughly a 17 percent gain relative to their half-day peers, while in reading the corresponding advantage is nearly 19 percent. The estimates upon which these gains are based are both very precisely estimated. Examining these short-run estimates by gender (see the top panel of Table 5) reveals somewhat larger gains for full-day boys relative to full-day girls in both reading and

Table 5
Full-day kindergarten coefficient estimates, by race and gender

	Mathematics				Reading			
	Short-run		Longer-run		Short-run		Longer-run	
	Males	Females	Males	Females	Males	Females	Males	Females
White children								
Full-day kindergarten (FDK)	1.587***	1.212***	0.759**	0.626**	2.151***	1.487***	0.740	0.761
	(0.341)	(0.256)	(0.350)	(0.314)	(0.329)	(0.341)	(0.544)	(0.515)
	[18.0%]	[15.3%]	[8.5%]	[7.8%]	[21.5%]	[15.4%]	[5.6%]	[6.2%]
	{5.4%}	{4.1%}	{1.7%}	{1.4%}	{6.6%}	{4.4%}	{1.3%}	{1.3%}
R^2	0.67	0.65	0.49	0.50	0.68	0.66	0.46	0.48
Sample size	2829	2730	2829	2730	2829	2730	2829	2730
Black children								
Full-day kindergarten (FDK)	1.188**	0.652	0.061	−0.463	0.889	0.903	−0.223	−1.367
	(0.479)	(0.467)	(0.703)	(0.749)	(0.806)	(0.621)	(1.478)	(1.311)
	[15.8%]	[9.3%]	[0.7%]	[−5.6%]	[9.6%]	[10.4%]	[−1.6%]	[−10.7%]
	{5.1%}	{2.8%}	{0.2%}	{−1.2%}	{3.2%}	{3.1%}	{−0.5%}	{−2.6%}
R^2	0.68	0.57	0.53	0.52	0.65	0.62	0.49	0.44
Sample size	709	736	709	736	709	736	709	736
Hispanic children								
Full-day kindergarten (FDK)	1.404***	0.901	−1.276	−0.416	1.869***	2.621***	−0.236	−0.555
	(0.520)	(0.631)	(0.872)	(0.700)	(0.709)	(0.790)	(1.344)	(1.254)
	[18.4%]	[11.3%]	[−14.8%]	[−5.1%]	[20.3%]	[27.0%]	[−1.8%]	[−4.4%]
	{5.6%}	{3.5%}	{−3.0%}	{−1.0%}	{6.5%}	{8.3%}	{−0.5%}	{−1.0%}
R^2	0.69	0.69	0.49	0.53	0.64	0.67	0.45	0.53
Sample size	594	566	594	566	594	566	594	566

Notes: Samples restricted to public school students who are first-time kindergartners and do not switch schools during kindergarten. All models include the appropriate baseline test score and the extensive set of covariates listed in Table 2A, with the exception of child gender. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets. Robust standard errors, adjusted for clustering at the school level, are in parentheses. *= $0.10 \geqslant p > 0.05$, **= $0.05 \geqslant p > 0.01$ and ***= $p \leqslant 0.01$.

math. In reading, the estimates suggest that boys who are full-day kindergartners exhibit a nearly 22 percent advantage over their male half-day kindergarten counterparts, while corresponding girls have a smaller, yet still substantial, advantage of roughly 15 percent. These boys also have larger gains in math, but the difference is much smaller (18.0 and 15.3 percent for boys and girls, respectively).

Longer-run estimates, contained in Table 2B, tell a much different story. For example, the 17 percent advantage in math, reported in Table 2A, falls by more than half, to less than 8 percent. The corresponding 19 percent advantage in reading slides even further to just over 5 percent. While much closer to zero than short-run estimates in Table 2A, the longer-run estimates contained in Table 2B are likewise precisely estimated, though the reading coefficient in my most preferred specification should be considered only marginally significant. Hence, the gains of white full-day kindergartners, relative to their half-day peers, decline substantially over the course of an additional year so that they may now lack practical significance.

Examining longer-run gains by gender (see the top panel of Table 5) shows virtually no difference in the longer-run estimates in either the mathematics score (7.8 percent for girls and 8.5 percent for boys) or the reading score (6.2 percent for girls and 5.6 percent for boys). Overall, the estimated pattern for white children is one of relatively large differences by the end of kindergarten that depreciate substantially 1 year later, by the end of first grade.[12]

[12]Moving from column 2 to column 3 in Tables 2–4, the estimated effects of family and individual covariates are reduced in magnitude substantially. This suggests the baseline test score is highly correlated with these characteristics. Corresponding standard errors also drop considerably, especially in short-run models. In other words, the addition of the baseline score reduces residual variance, as one might expect.

4.2. Black children

Similar to, but to a lesser extent than white children, black full-day kindergartners outperform their half-day peers in both mathematics and reading by the end of kindergarten. As seen in Table 3A, black full-day kindergartners have gains in math and reading scores that are roughly eleven percent higher than their half-day counterparts. Both of these gains are estimated relatively precisely, though the reading estimate should be considered only marginally significant relative to conventional levels ($p-\text{value}=0.06$). Unlike white children, estimates suggest possible gender differences among black full-day kindergartners. For example, I estimate that full-day kindergarten boys exhibit a short-run mathematics score gain that is roughly 16 percent higher than their half-day peers. Note that despite the relatively small sample size, this gain is estimated very precisely. Conversely, I find that black girls enrolled in full-day kindergarten have only a 9 percent gain relative to their half-day peers. Estimated short-run reading score differences are virtually identical for black girls and boys.

Relative to short-run estimates, longer-run results paint a much different picture. Moving from Table 3A and B, it is apparent that estimated short-run gains have largely disappeared. Further, as can be seen in the middle panel of Table 5, the 16 percent short-run advantage enjoyed by black male full-day kindergarteners is estimated, albeit imprecisely, at less than 1 percent in the longer-run. So, if there are short-run gains to black boys who attend full-day kindergarten, they appear to have vanished 1 year later, by the end of first grade. The corresponding results for girls paint an even bleaker picture, as point estimates imply that these full-day kindergartners actually score lower than their half-day counterparts, especially in reading where estimates imply a nearly 11 percent *dis*advantage for black girls enrolled in full-day kindergarten. However, a null relationship cannot be rejected in these longer-run models.

4.3. Hispanic children

To a greater extent than either white or black children, I find that Hispanic full-day kindergartners outperform their half-day counterparts in both mathematics and reading in the short-run. For example, Table 4A shows that Hispanic full-day kindergartners exhibit a short-run mathematics score gain that is nearly 16 percent higher and a corresponding reading score gain that is nearly 24 percent higher than their half-day counterparts. Both estimated differences are statistically different from zero at conventional levels. Given concern over the academic achievement of Hispanic children, who tend to lag behind white and black children, these findings likely would be interpreted as "good news" by educators and policymakers alike.[13] Examining these short-run estimates by gender (see the bottom panel of Table 5) demonstrates that full-day kindergarten tends to benefit Hispanic boys relatively more than girls in math, and vice versa in reading, though differences are small.

Moving to the longer-run results reported in Table 4B, the story changes dramatically. I find no evidence that Hispanic full-day kindergartners retain any of their sizeable short-run advantages. Indeed, I find that these children actually score lower than their half-day peers 1 year later. While not statistically different from zero, preferred specifications imply that Hispanic full-day kindergartners have longer-run mathematics and reading score disadvantages of roughly 9 and 1 percent, respectively. Note, however, that these estimated disadvantages are small, in absolute value, relative to corresponding short-run estimates.

Examining these longer-run gains by gender (see the bottom panel of Table 5) suggests that both Hispanic girls and boys score lower in mathematics relative to their half-day peers, though these gains are not very precisely estimated. In sum, the pattern of large short-run gains, followed by much smaller longer-run differences appears most pronounced for Hispanic children.[14]

Overall, my findings suggest that any short-run differences in math and reading proficiency due to full-day kindergarten have vanished by the end of an additional year of schooling. Given that kindergarten type is not randomly assigned, it is important to note that the pattern of my findings is not, in any straightforward manner, consistent with differential selection into full-day kindergarten. For example, if

[13] As of October 2000, the fraction of Hispanics aged 16–24 classified as "status" dropouts was 27.8 percent, as compared to 6.9 percent of whites and 13.1 percent of blacks (NCES, 2002a).

[14] Though not reported, I perform several sensitivity analyses of estimates presented in the previous three sub-sections. I find that the magnitudes of my estimates and their temporal pattern continue to obtain. These estimates are available upon request.

Table 6
Beginning of 1st grade regression estimates—Dependent variable: Start-of-1st grade test score

	Mathematics			Reading		
Selected covariates	White	Black	Hispanic	White	Black	Hispanic
Full-day kindergarten (FDK)	1.260***	0.471	1.663	2.311***	−1.930*	1.255
	(0.411)	(0.963)	(1.034)	(0.531)	(1.099)	(1.173)
	[13.9%]	[5.6%]	[19.3%]	[18.7%]	[−18.5%]	[10.7%]
	{3.6%}	{1.7%}	{5.4%}	{5.8%}	{−5.7%}	{3.4%}
Maximum of parent education						
Less than HS	−0.836 (0.726)	0.550 (0.886)	−0.322 (1.067)	−0.296 (1.145)	−0.155 (0.872)	−1.508 (1.379)
Some college	0.335 (0.431)	−0.095 (0.581)	1.789** (0.774)	0.499 (0.502)	0.973 (0.876)	1.243 (1.038)
College degree	0.547 (0.446)	−0.454 (1.186)	3.339** (1.424)	0.330 (0.609)	−0.989 (1.468)	1.514 (2.197)
Master's degree or higher	0.950* (0.544)	1.404 (2.486)	1.015 (1.886)	1.777** (0.712)	2.159 (1.906)	2.908 (2.627)
Child female	−0.202 (0.278)	−0.170 (0.569)	0.516 (0.780)	0.956** (0.395)	1.162 (0.786)	2.642*** (0.997)
Baseline test score	0.916*** (0.026)	1.127*** (0.068)	0.905*** (0.068)	1.095*** (0.027)	1.062*** (0.065)	1.002*** (0.066)
R^2	0.65	0.66	0.68	0.64	0.67	0.72
Dependent mean/standard deviation	34.96/9.07	28.27/8.41	30.85/8.61	40.11/12.33	33.96/10.46	36.49/11.72
Sample size	1,620	392	312	1,620	392	312

Notes: Samples restricted to public school students who are first-time kindergartners and do not switch schools during kindergarten. Models also include the extensive set of covariates listed in Table 2A. Percent of dependent standard deviation in square brackets and percent of dependent mean in curly brackets, for FDK coefficient only. Robust standard errors, adjusted for clustering at the school level, are in parentheses. *= $0.10 \geqslant p > 0.05$, **= $0.05 \geqslant p > 0.01$ and ***= $p \leqslant 0.01$.

the gains associated with full-day kindergarten persisted over time, one may worry that full-day kindergartners are positively selected in unobserved ways. Likewise, if unobserved school quality differences are responsible for the short-run differences between full and half-day children, it seems unlikely that these differences would disappear over such a short period of time. Though not reported, this pattern remains intact when I restrict my sample to those children who attend the same school in kindergarten and first grade.

5. Why do full-day effects fade over time?

To this point, I have presented evidence that gains associated with full-day kindergarten fall off sharply just 1 year later, towards the end of first grade. The next natural question is: Why do these short-run advantages fade? While this is a difficult question to answer, an understanding of *when* gains are lost may provide information about *why* they are lost. For example, if gains are lost over summer vacation, the explanation may be linked to the child's home or neighborhood environment, and not directly school-related. Indeed, education researchers have documented that gains associated with many interventions peak for a short time and then decline when the intervention is ended or otherwise disrupted (Cooper, Nye, Charlton, Lindsey, & Greathouse, 1996; Entwisle & Alexander, 1992).[15] Conversely, if gains are sustained over the summer only to be lost during the next school year, this is more consistent with a school-related explanation (e.g., lack of coordination between kindergarten and first grade curricula). This latter explanation, of course, does not preclude a role for the home or neighborhood environment; indeed, the two explanations are not mutually exclusive.

I use individuals from the ECLS-K's randomly chosen 30 percent sub-sample to investigate the timing of observed patterns. In particular, these data allow me to understand if short-run gains have disappeared by the start of first grade, consistent with summer fallback, or if they persist over the summer, only to be lost sometime during first grade. Table 6 presents results from regressions following my most preferred specification, the only exception being that the dependent variable is now measured near the start of first grade, when the 30 percent sub-sample was collected.

[15]This phenomenon is often referred to as "summer fallback". Related findings have led to calls for year-round schooling, in the hope that achievement gains might be sustained over time.

I find evidence of summer fallback for black children, but much less for Hispanic, and especially, white children. More specifically, while black full-day kindergartners show advantages of roughly 11 percent in math and reading by the end of kindergarten, comparable estimates decline to 5.6 percent and −18.5 percent, respectively, by the start of first grade. Results for Hispanic children are more mixed. While Hispanic full-day kindergarteners gain slightly in their mathematics advantage (15.6–19.3 percent), their advantage in reading slips from nearly 24 percent to just under 11 percent, relative to their half-day counterparts. Estimates for white children are very consistent with short-run results and hence offer even less evidence of summer fallback. In particular, I estimate that at the beginning of first grade white full-day kindergarteners have a 13.9 percent advantage over their half-day peers in math and 18.7 percent in reading, compared to 17.1 and 18.7 percent, respectively, at the end of kindergarten—prior to summer vacation.

Taken as a whole, the evidence presented suggests the gains of black children have disappeared by the start of first grade, while the gains of white and Hispanic children apparently diminish during the school year. Since the short-run gains of black children are smaller than whites or Hispanics, this is not too surprising. Nevertheless, the finding is consistent with literature on summer fallback (Cooper et al., 1996). From a policy perspective, additional effort to understand these losses is likely worthwhile.

6. Conclusions

The estimated pattern of results suggests that full-day kindergarten substantially raises the math and reading achievement of children of all races. However, these gains are much smaller in magnitude when measured via similar tests just one year later. In other words, the short-run impact of full-day kindergarten has depreciated considerably by the end of first grade. The observed pattern is even more striking for minority children since some of the specifications imply that these full-day kindergartners actually perform worse than their half-day counterparts by the end of first grade. Hispanic children, in particular, exemplify this pattern.

This pattern runs contrary to the notion that full-day kindergarten augments child human capital in a manner that allows for improved learning as children progress through school. While the estimates show substantial depreciation for all groups of children, declines are shallowest for whites. Given existing socioeconomic differences between the races, it is possible that differences in home environment contribute significantly to the larger losses for Hispanic, and especially black, children. This explanation is consistent with the summer fallback evidence presented. Another possible explanation is that black and Hispanic children are relegated to poorer quality schools that teach at a lower level, on average, than those attended by their white peers. Of course, both types of explanations may contribute simultaneously to the observed patterns.

Given the limited availability of funds that can be devoted to early childhood education and the considerable costs associated with the provision of a full-day program, it is reasonable to want an improved understanding of its academic returns. My findings suggest that, on average, the academic returns associated with full-day kindergarten are quite low or non-existent. However, it should be noted that a complete evaluation of full-day kindergarten is beyond the scope of this paper. In addition to potential academic returns, issues related to child socialization and full-day kindergarten's implicit child care subsidy are also key ingredients in constructing a more complete assessment and hence merit future attention from researchers.

Acknowledgements

I thank Rodney Andrews, Charlie Brown, Kerwin Charles, Mike Chernew, Julie Cullen, Angela Dills, John DiNardo, Rich Hirth, Pat Kline, Sara LaLumia, Justin McCrary and Diane Whitmore for helpful comments. All remaining errors are mine.

References

Carr, M., & Jessup, D. (1997). Gender differences in first grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *98*(2), 318–328.

Cascio, E. (2003). Getting a head start or falling further behind? Black-white relative schooling attainment and the introduction of kindergartens in the South. Working paper.

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

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review*, *85*(3), 341–364.

Currie, J., & Thomas, D. (1999). Does Head Start help Hispanic children? *Journal of Public Economics*, *74*(2), 235–262.

Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly*, *12*(4), 459–480.

Entwisle, D., & Alexander, K. (1992). Summer setback: Race, poverty, school composition and math achievement in the first two years of school. *American Sociological Review*, *53*(Feb.), 72–84.

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and affective factors. *American Educational Research Journal*, *19*(1), 51–71.

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, *92*(4), 999–1012.

Goldin, C. (1991). The role of World War II in the rise of women's work. *American Economic Review*, *81*(4), 741–756.

Gorton, H. (1968). A study of the kindergarten program: Full or half day? ERIC Document 012 327.

Gullo, D. (2000). Long-term educational effects of half-day vs. full-day kindergarten. *Early Child Development and Care*, *160*, 17–24.

Hough, D. & Bryde, S. (1996). The effects of full-day kindergarten on student's academic performance, Paper presented at the 1996 AERA meetings.

Johnston, J., & DiNardo, J. (1997). *Econometric methods* (4th ed.). New York: McGraw-Hill.

Jones, S. (2002). Effect of all-day kindergarten on student cognitive growth: A meta-analysis. Doctoral dissertation, University of Kansas.

Karweit, N. (1992). The kindergarten experience. *Educational Leadership*, *49*(6), 82–86.

Lemke, R., Witte, A., Querault, M., & Witt, R. (2000). Child care and the welfare to work transition. NBER working paper #7583.

Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

Natale, J. (2001). Early learners: Are full-day academic kindergartens too much, too soon? *American School Board Journal*, *188*(3), 22–25.

Oelerich, M. (1984). Should kindergarten children attend school all day every day? *The Journal of College of Education*, *Fall*, 13–16.

Ohio State Legislative Office of Education Oversight. (1997). *An overview of full-day kindergarten*. Columbus, OH.

Olsen, D., & Zigler, E. (1989). An assessment of the all-day kindergarten movement. *Early Childhood Research Quarterly*, *4*(2), 167–186.

Siefert, K. (1993). Cognitive development and early childhood education. In D. Spodeck (Ed.), *Handbook of research on the education of young children*. New York: McMillan.

Ulrey, G., Alexander, K., Bender, B., & Gillis, H. (1982). Effects of length of school day on kindergarten performance and parent satisfaction. *Psychology in the Schools*, *19*, 238–242.

US Census Bureau. (2002). *The population profile of the United States: 2000 (Internet Release)*, www.census.gov/population/www/pop-profile/profile2000.html, Washington, D.C.

US Department of Education, National Center for Education Statistics. (2002a). *Digest of education statistics 2001*, NCES 2002-130, by Thomas Snyder. Production Manager, Charlene M. Hoffman. Washington, DC.

US Department of Education, National Center for Education Statistics. (2002b). *User's manual for the ECLS-K first grade public-use data files and electronic codebook*, NCES 2002-135, Washington, DC.

US Department of Education, National Center for Education Statistics. (2002c). Early Childhood Longitudinal Study—Kindergarten Class of 1998-99 (ECLS-K), In D.A. Rock, J.M. Pollack (Eds.), Psychometric report for kindergarten through first grade, NCES 2002-05, Educational Testing Service, Elvira Germino Hausken, project officer. Washington, DC.