First in the Class?
Age and the Education Production Function

Elizabeth Cascio
Associate Professor of Economics
Dartmouth College

Diane Whitmore Schanzenbach
Associate Professor of Human Development and Social Policy
Faculty Fellow, Institute for Policy Research
Northwestern University

Version: January 2014

DRAFT
Please do not quote or distribute without permission.

Abstract

Cascio and Schanzenbach estimate the effects of having older peers using data from an experiment where children of the same age were randomly assigned to different kindergarten classrooms. Exploiting this experimental variation in conjunction with variation in expected kindergarten entry age to account for negative selection of some of the older school entrants, they find that exposure to older kindergarten classmates increases test scores up to eight years after kindergarten and may reduce the incidence of grade retention and increase the probability of taking a college-entry exam. These findings are consistent with broader peer effects literature documenting positive spillovers from having higher-scoring peers and suggest that, holding constant own age, being old relative to one's peers is not beneficial.

1.   **Introduction**

In 1968, only four percent of U.S. six year olds were not in first grade. By 2010 this figure had risen to almost 20 percent.[1] Only one-third of the recent trend is estimated to be explained by increases in the minimum kindergarten entry ages established by U.S. states; the rest is thought to be due to teachers deeming more kindergartners unready to progress to first grade, and to more parents choosing to delay their child's entrance into school (Deming and Dynarski, 2008). Estimates over the past fifteen years put the rate of delay as high as 9 to 10 percent (e.g., West, Meek, and Hurst, 2000).

Anecdotally, the parental decision to delay is motivated by concerns that a child will be permanently disadvantaged if not among the biggest and brightest when he starts school. Being old *relative to* one's classmates can matter for achievement over the long term if children are tracked early on the basis of test scores, which are strongly correlated with age when children are young. Placement in the top academic track can be self-reinforcing, since it may tackle more advanced material and move more quickly through a given curriculum. At the same time, relatively old school entrants might become relatively more motivated for school or self-confident because of their relative standing in the class. Importantly, the implied result is zero-sum: when relatively old students gain, relatively young students lose, becoming less engaged with school, being placed on lower academic tracks, etc. This can create an unsustainable "race to the top" (see Gootman, 2006, and Weil, 2007).

Is there any evidence to support the existence of relative age effects? Older children do perform better on tests than younger children in the same grade, even as late as middle school.[2] However, this performance differential alone is not conclusive evidence that *relative*

[1] These are the authors' tabulations from the 1968 and 2010 October Current Population Survey (CPS) School Enrollment Supplements. For the purposes of this calculation, we record the small number of individuals in second grade or higher as being in first grade.
[2] Bedard and Dhuey (2006) show this for a number of countries using data on fourth and eighth graders from the 1995 and 1999 Trends in International Mathematics and Science Study (TIMSS). Similar findings have been documented in country-specific studies of the United States (Datar, 2006; Elder and Lubotsky, 2009), Sweden

age matters: Older kindergartners have greater preparation for school in an absolute sense, not just in relation to their peers; age also signals cognitive development as well as the flow of investments made in a child over his lifetime, so that older entrants' edge on tests, even in middle school, may not even have anything to do with school *per se*.[3] Direct estimates of relative age effects call for information on the ages of a child's peers, which is generally lacking in the survey data frequently used in this literature. As a result, there is limited evidence to date on whether children are better positioned to succeed if they are old relative to their peers when they start school.

We address this identification problem using data from one of the largest educational experiments ever undertaken in the United States – Tennessee's Project STAR (Student-Teacher Achievement Ratio). Though designed to study the effects of reduced class size (Schanzenbach, 2007), Project STAR is well suited for estimating the effects of relative age.[4] In particular, the design of Project STAR allows us to observe children who entered school at the same age, but were randomly assigned to kindergarten classmates with different ages on average. We can also follow these children over time. While such an exercise would be feasible in other data, such as matched student-teacher administrative data, Project STAR helps to ensure that variation in relative age at the classroom level is random. In a non-experimental setting, by contrast, concerned parents may pressure school administrators to place their children in kindergarten classrooms where they would be relatively old.

(Fredriksson and Öckert, 2006), Chile (McEwan and Shapiro, 2008), and Germany (Puhani and Weber, 2005), among others.

[3] In fact, the impossibility of separately estimating the effects of age at school entry and age at test or observation on elementary school test scores is cited by Angrist and Pischke (2009, p. 7) as an example of a "fundamentally unidentified question." Black et al. (2011) separate the effects of age at school entry and age at test or observation by estimating these for army entrance exams given to young adults.

[4] We are not the first to exploit the random assignment of children and teachers in Project STAR to classes, not just to class sizes, to gain insights into the education production function (Dee, 2004; Dee and Keys, 2004; Whitmore, 2005; Schanzenbach, 2006; Chetty et al., 2011). Others have also exploited the random assignment of students and teachers to classes of different sizes to estimate peer effects (Boozer and Cacciola, 2001; Graham, 2008).

Our empirical strategy also accounts for the possible negative selection of delayed school entrants by exploiting variation across students in "expected" school entry age given exact date of birth and Tennessee's birthday cutoff for kindergarten entry. Our estimates therefore isolate the effects of older peers being older separately from the effects of their being on average lower ability. This aspect of our identification strategy is similar to that taken by Elder and Lubotsky (2009), who estimate the effects of having older peers in one's school grade cohort using nationally-representative survey data.

We find little evidence to support the hypothesis that there are benefits to being relatively old in one's kindergarten classroom, holding constant own age. Own age taken as given, children who are older relative to their kindergarten classmates – that is, have younger peers – appear to lose out on several dimensions. For example, they score worse on achievement tests, both at the end of kindergarten and in middle school. They are also more likely to have been retained in grade by middle school and are less likely to take a college-entrance exam, though these effects are estimated less precisely.

Thus, our findings suggest, on net, *positive* – and potentially persistent – spillovers from having older classmates in kindergarten. Age could have positive spillover effects through several channels. As noted above and also shown below, older children score higher on tests, particularly at the start of school. In addition, older children should be more ready for a given curriculum and are therefore potentially less disruptive. While surprising from the perspective of academic and popular discussions of school entry age effects that suggest an advantage to being among the oldest, our findings are thus broadly consistent with the peer effects literature that finds students benefit from being exposed to higher-quality peers.[5]

---

[5] A number of studies have found that exposure to higher-achieving or less disruptive peers has benefits for a child's own achievement and behavior. On the effects of exposure to higher-achieving peers, see Hanushek et al. (2003), Ding and Lehrer (2007), Duflo et al. (2011), and Kugler et al. (2012). Figlio (2007) and Carrell and Hoekstra (2010) provide evidence that exposure to more disruptive peers can be harmful. See Sacerdote (2011) for a complete review of this literature.

2. Background

Our analysis will focus on test performance and several other academic outcomes of one school entry cohort. The model of interest is:

$$(1) \qquad y_{it} = \beta_{0t} + \beta_{1t} a_i + \beta_{2t} f(a_i, \mathrm{A}_{-i}) + \varepsilon_{it}$$

where y_{it} represents an outcome for individual i in year t, and a_i is his observed age (in years) in the fall of kindergarten. In this application, as in most in this literature, a_i is perfectly correlated with the age at which an outcome is measured, a_{it}. We therefore refer to their combined effect, β_{1t}, as the "absolute age effect." $f(\cdot,\cdot)$ represents relative age, taking as arguments own age a_i and the vector of ages of child i's peers, A_{-i}; we refer to its coefficient, β_{2t}, as the "relative age effect." Thus, relative age is some function of a child's own age and the ages of his peers. ε_{it} represents unobserved determinants of the outcome of interest.

Most studies to date have lacked data on A_{-i}, and hence have omitted relative age when estimating model 1. As a result, most existing estimates of the coefficient on a_i are reduced-form, capturing the effects of both absolute age and relative age. This has generated uncertainty about correct interpretation of this reduced-form coefficient in the literature; its interpretation nevertheless has important implications. If the coefficient on age is truly a relative age effect, postponing a child's school entry – either by choice or through policy – has negative consequences for other children. Likewise, a principal's manipulation of class age distributions may benefit some children at the expense of others. On the other hand, if the reduced-form coefficient on age represents an absolute age effect, such spillovers from changing class age distributions do not exist.

We tackle the question of interpretation by using a data set where we observe the ages of each respondent's peers at the start of school.[6] Thus, we are able to measure relative

---

[6] An alternative approach would be to impose functional form restrictions to support one interpretation over another. For example, in one of the earliest papers in this literature, Bedard and Dhuey (2006) argue that their

4

age and control for it when estimating model 1. Yet, two questions arise in specifying this model.

First, what is the correct way to measure a child's relative age? For simplicity and consistency with Elder and Lubotsky (2009), the first paper to address this issue, we assume that $f(a_i, \mathbf{A}_{-i}) = a_i - \bar{a}_{-i}$, where $\bar{a}_{-i}$ represents the average age of a i's peers. Thus, we assume that relative age is linear, continuous, increasing in own age, and decreasing in the average age of peers. This captures the zero-sum implications of manipulating relative age. For example, in order for a given student to become relatively old, another student must become relatively young. As defined, relative age thus has implications for the distribution of achievement, but not its level.

Second, which peers are relevant for relative age effects? A key advantage of our data is that we can observe relative age at the classroom level. To make the most of the experimental variation in Project STAR, we focus on relative age in the kindergarten classroom. A potential drawback of this approach is that we are only able to detect whether within-classroom interactions in kindergarten drive relative age effects, not whether relative age effects exist at all. For example, if age in relation to classmates determines placement into reading groups, but children do not learn to read until first grade or later, we may not detect relative age effects when they truly exist, or our estimates of relative age effects could be biased downward. Fortunately, we are able to explore the sensitivity of our findings by estimating models where relative age is defined in relation to a child's first-grade classmates, and the results are little changed.[7]

As mentioned above, Elder and Lubotsky (2009) were the first to tackle these questions, and we measure relative age in the same way that they do. However, with regard

---

positive estimates of the coefficient on age for eighth-grade test scores capture the effect of relative age – not age at school entry – on the basis of a model where the effect of age at school entry is assumed non-linear.
[7] This is not surprising, as relative age in the kindergarten and first grade classroom has a correlation of 0.96.

to the question of which peers determine relative age effects, they use nationally representative data where it is only possible to measure relative age at the school (grade) cohort level.[8,9] Such estimates capture the possibility that children who are relatively young in their cohorts are likely to be among the youngest in their classrooms at some point during their elementary-school careers, and potentially repeatedly. They also capture the possibility that school administrators, not just teachers, target the youngest students in a cohort for lower academic tracks. In an attempt to reconcile our estimates with those of Elder and Lubotsky, as well as to provide insight into causal mechanisms, we also consider models where relative age is defined in relation to other children in the entire entry cohort of a child's elementary school, and not just his own classroom. These estimates do not rely on the experimental variation in classroom assignment in our data.

3.   Empirical Strategy

Our primary estimates focus on a child's relative age in his kindergarten classroom and, for ease of estimation and to parallel the peer effects literature, are based on the following model:

$$(2) \qquad y_{itk} = \alpha_{0t} + \alpha_{1t} a_i + \alpha_{2t} \bar{a}_{-i,k} + x'_{ik} \delta_t + \varepsilon_{itk}.$$

y_{itk} now represents the outcome of individual i in year t who was assigned to classroom k in kindergarten, and a_i remains his observed age at the start of kindergarten. Given that Project STAR included only one academic cohort and all students were tested (or observed) at the same time, a_i is perfectly correlated with the age at which outcomes are measured. We are therefore only able to identify their combined effect, as noted above. $\bar{a}_{-i,k}$ is the average age of a child's kindergarten classmates; x_{ik} is a vector including fixed (predetermined)

---

[8] Fredriksson and Öckert (2006) use administrative data from Sweden to estimate a version of model 1 where relative age is measured as child's rank in the age distribution of his ninth-grade school cohort.
[9] In principle, there are other situations where relative age varies while absolute age remains constant. For example, changes in the minimum age at school entry change the expected relative age of some children without changing the year in which they should enter school (see Bedard and Dhuey, 2012).

6

characteristics of the individual, other (predetermined) characteristics of his kindergarten classmates, characteristics of his teacher, class size, and school indicators;[10] and ε_{itk} denotes unobserved determinants of outcomes.

As shown in Elder and Lubotsky (2009), the effect of relative age, $a_i - \bar{a}_{i,k}$, can be obtained simply by re-expressing model 2:

$$(2') \qquad y_{itk} = \beta_{0t} + \beta_{1t}a_i + \beta_{2t}\left(a_i - \bar{a}_{-i,k}\right) + x'_{ik}\delta_t + \varepsilon_{itk}$$

where $\beta_{0t} = \alpha_{0t}$, $\beta_{1t} = \alpha_{1t} + \alpha_{2t}$ and $\beta_{2t} = -\alpha_{2t}$. Thus, the relative age effect is $\beta_{2t} = -\alpha_{2t}$, where α_{2t} is the effect of peer average age on outcomes at time t from model 2. The expectation is that $\beta_{2t} > 0$. In words, holding constant own (entry) age, being relatively old improves academic outcomes; alternatively, holding constant own (entry) age, having older peers lowers outcomes on average ($\alpha_{2t} < 0$). The absolute age effect, $\beta_{1t} = \alpha_{1t} + \alpha_{2t}$, is the sum of the coefficients on own and peer average age in model 2. Thus, if the relative age effect is positive, we would expect the coefficient on a_i to fall relative to its value in the more commonly estimated specification where relative age is omitted from the model.

A comparison of models 2 and 2′ makes clear that a relative age effect is indistinguishable from the effect of having younger peers. Having younger peers may improve a child's position in the classroom "pecking order" – the positive spillover that we conceive of as a "relative age effect." But it may also have negative spillovers, e.g., by exposing a child to lower-performing or more disruptive children in the classroom. Thus, even though we can re-express the parameters in model 2 to derive a relative age effect – as in model 2′ – the coefficient α_{2t} is inherently a reduced-form one, capturing the net effect of these competing spillovers. It is important to keep this reduced-form interpretation in mind as we proceed.

---

[10] We include school fixed effects in the model because classrooms were randomly assigned within schools.

Two identification problems arise when estimating model 2, both of which we can address. First, because kindergarten retention and delay are at the discretion of parents and teachers, older kindergartners may differ in unobservable ways from younger kindergartners, leading ordinary least squares (OLS) estimates of α_{1t} to be biased. In particular, children delayed or retained are likely negatively selected, imparting a downward bias on OLS estimates. We take the same empirical approach as previous researchers to this identification problem, constructing an instrument for a_i using information on a child's birthday and the date by which new school entrants are to reach a specified age.[11] In particular, children in the Project STAR cohort were to have turned age five on or before September 30. As a result, a child who turned five on September 30 should have enrolled in kindergarten a month before, while her counterpart who turned age five on October 1 is expected to have entered kindergarten about one month before her sixth birthday.

The solid line in Figure 1 illustrates the relationship between birthday and this "expected" age at the start of kindergarten, ea_i, under the assumption that the school year begins September 1. The figure shows clearly that actual age on September 1, 1985 (daily averages of which are represented by the hollow circles) is strongly but not perfectly related to ea_i. Two-stage least squares (TSLS) estimates of α_{1t} will be consistent if ea_i is uncorrelated with unobserved determinants of achievement, or if $E\left(ea_i \varepsilon_{itk} \mid x'_{ik} \delta_t \right) = 0$.[12] While we cannot test this assumption directly, we show below that expected age at the start of kindergarten generally does not predict a child's observed characteristics.

---

[11] See, for example, Bedard and Dhuey (2006) and Elder and Lubotsky (2009), among others. A similar approach can be used to construct an instrument for years of completed schooling in models of youth test scores (e.g., Gormley and Gayer, 2005; Cascio and Lewis, 2006) and of adult outcomes (e.g., Dobkin and Ferreira, 2010; McCrary and Royer, 2011).

[12] A sufficient condition for this assumption to be satisfied is that birthday is randomly assigned, but it is possible to identify α_{1t} under weaker assumptions, e.g., under the assumption that expected school-entry age is randomly assigned conditional on some flexible function of birthday that is smooth through the cutoff date. We have estimated a regression-discontinuity model like this as a robustness check and found similar results to those reported below (results available on request).

Second, there is likely to be sorting of children across kindergarten classrooms, possibly on unobservable characteristics related to the ages of other children. Our use of experimental data implies that no such sorting should have taken place, in which case $E\!\left(\bar{a}_{-i,k}\varepsilon_{itk} \mid x_{ik}'\delta_t\right)=0$. However, even with random assignment, older peers may be on average relatively low ability because some of these children will have been held back by their parents or teachers on the basis of their perceived school readiness. Thus, OLS estimates of α_{2t} pick up the effects of having lower-ability peers in addition to the effects of having older peers.

To ensure that the peer effect we identify derives from age rather than ability, we combine the identification strategy described above with the random variation in classroom assignment from Project STAR. Specifically, we instrument for $\bar{a}_{-i,k}$ with the average expected age of a child's peers, $\overline{ea}_{-i,k}$. Given the strong correlation between ea_i and a_i shown in Figure 1, it is not surprising that we find a strong correlation between these variables. This instrument should also be valid because variation in the birthday composition of a child's classmates should be uncorrelated with his latent academic potential due to the experimental design. As above, we can informally test this assumption by examining the relationship between $\overline{ea}_{-i,k}$ and observables.

Thus, while the OLS and TSLS estimates of α_{2t} are in principle both unbiased, they are estimates of different parameters. The explanation is intuitive: given that older classmates may have been delayed or previously retained in kindergarten, OLS estimates of α_{2t} pick up not just the effect of having older classmates in kindergarten, but also the effect of having "overage" classmates in kindergarten, who are likely of below-average ability. By contrast, the TSLS estimates identify how having older peers of the *same* innate ability at the start of school affects the average child's academic performance.

As described above, Elder and Lubotsky (2009) use a similar identification strategy to estimate the effects of relative age. However, in using nationally-representative survey data, they have a source of variation in peer average age at their disposal that we do not. In particular, in their case, in addition to exploiting variation across schools in the birthday compositions of grade cohorts, there is variation across states in peer average expected age – and hence in relative age – holding the birthday composition of cohorts fixed, due to variation across states in school entry age regulations. By contrast, our identification relies only on the transitory (and random) fluctuations in the age distributions of kindergarten classrooms within schools, and hence holds school entry age regulations constant. In other words, Elder and Lubotsky exploit variation in age groupings at the cohort level that can be induced by changes in state regulations on school entry, while our paper uses variation in age groupings that can be manipulated by elementary school principals within cohorts and a given school entry age regime.[13] We discuss below how the source of variation might affect our estimates.

4. Data

4.1. Sample and Summary Statistics

Project STAR was an experiment designed to study the effects of class size on student achievement. Kindergarten students and teachers in 79 Tennessee schools were randomly assigned to three different class types – small (with target enrollment of 13-17 students), regular (with target enrollment of 22-25 students), and regular with a full-time teacher's aide – in the fall of 1985.[14] This cohort was to have maintained its class type

---

[13] Our calculations from the Common Core of Data suggest that over 95 percent of kindergarten students attend a school with more than 30 students in their grade, implying that most schools have multiple kindergarten classrooms and therefore principals face the choice of assigning students across classrooms.
[14] Children entering the experiment in grades one through three, either by moving into the school or having been retained in grade the previous year, were also added to existing classes through random assignment. Each of the 79 schools had enrollment sufficient to accommodate at least one class of each type and were thus slightly larger than the state average. To ensure sufficiently large samples of poor and minority children, Project

through third grade, after which all participants were returned to regular-sized classes. Random assignment of children to class types took place within schools.

Our analysis exploits the fact that most, if not all, Project STAR participants would have been randomly assigned to classrooms, not just class types, as a result of the experimental design. We focus on peers in the kindergarten classroom because non-random transitions across class types (and classrooms) were less problematic in kindergarten than they would later become.[15] However, we do estimate model 2 at the first grade classroom level in a robustness check. As described above, we also estimate model 2 at the school-cohort level for comparison to Elder and Lubotsky (2009), but these estimates do not rely in any way on the experiment.

We restrict our estimation sample to participants for whom each of the available background characteristics (birthday, race, gender, free lunch status in kindergarten) and kindergarten teacher characteristics (experience, education, and race) is observed.[16] However, to construct average age and other characteristics of kindergarten classmates, we use all available data on the Project STAR kindergarten cohort, not just those individuals with non-missing outcomes data.

Table 1 gives summary statistics on these students and their teachers in our estimation sample. As has been found in nationally-representative data, children in our sample tend to be older at the start of kindergarten than expected (5.43 years old versus 5.38 years old), as shown in Panel A. However, the sample is not nationally representative. As shown in Panel B, nearly half of Project STAR participants received free or reduced-price

STAR schools were also disproportionately drawn from inner cities. A comparison of Project STAR schools to other Tennessee schools is provided in Schanzenbach (2007).

[15] Using administrative data for 18 Project STAR schools, Krueger (1999) found that only five of 1581 participants did not attend their initially assigned class type in kindergarten. As the experiment continued, however, ten percent of students made transitions across class types. Anecdotally, these class switches were largely the result of student misbehavior, which might plausibly be related to age at school entry or the average age of kindergarten classmates.

[16] This results in us dropping only 75 observations. The observations dropped are not significantly predicted, individually or jointly, by the instrumental variables for age and average age of peers.

lunch in kindergarten, and 33 percent were black (Panel B). By comparison, only 15.4 percent of five year olds in the U.S. were black in fall 1985.[17] About 17 percent of kindergarten teachers were black, 35 percent had master's degrees, and 10 percent had less than 2 years of experience (Panel D). Students were roughly equally divided across class size types (Panel E).

Our main outcomes come from tests administered to STAR participants through the end of high school. In the spring of kindergarten, STAR participants were administered the Stanford Achievement Test.[18] For participants still attending public school in Tennessee, we have scores on the Comprehensive Test of Basic Skills (CTBS) in grades 5 through 8, regardless of year attended.[19] Both tests are multiple-choice standardized tests with reading and math components. We average the reading and math scale scores on each test, then standardize this average to have a mean of zero and a standard deviation of one using data on all STAR participants with non-missing test scores in a given year. Thus, coefficient estimates in the test-score models are in standard deviation (σ) units. The last panel of Table 1 shows that the kindergartners in our sample are slightly positively selected from the pool of all Project STAR participants, scoring on average 0.12σ above the mean in spring 1994. This positive selection may arise because it was not mandatory for the Project STAR cohort to attend kindergarten, and kindergarten attendees may have been positively selected.

Our analysis focuses on test scores at the end of kindergarten (in spring 1986) and in spring 1994, when STAR participants progressing through school normally would have been completing eighth grade. We choose spring 1994 because many existing studies have

---

[17] There are authors' calculations from the 1985 October CPS School Enrollment Supplement.

[18] Scores on the Stanford Achievement Test are also available for STAR participants in grades 1 through 3 who did not leave a STAR school or repeat or skip a grade during the experiment. Unfortunately, our instrumental variables are related to attrition during the experiment, so we are unable to use these data. This is most likely driven by the relationship of these variables to grade repetition (shown below), but we are not able to observe the reason that children attrite from the sample, so we cannot confirm this.

[19] This is true as long as a child attended grades 5 through 8 at some point between 1990-91 and 1996-97. Test scores were also collected in 1989-90, but are not available for a large, non-random subset of children who attended school in Memphis because the tests were not universally given there in that year.

considered the relationship between age and test scores in eighth grade. We present

estimates for the year that the cohort was *expected* to be in eighth grade instead of eighth-

grade test scores because our sample includes individuals in the same school-entry cohort,

and we wish to have our estimates be consistently interpretable across tests taken at different

points in time. Because either one of the treatment variables may have affected grade

progression, we also estimate separate models for whether a child was enrolled below eighth

grade when tested in spring 1994. One quarter of students tested are deemed to have been

retained by this measure (Panel F).

Our final outcome measure is an indicator for whether a respondent took the ACT

or SAT college-entrance exam. College-entry test information on STAR participants was

collected from graduating classes through 1999 (i.e. for students who graduated early, on-

time, or no later than one year behind "normal" grade progression) from all high schools in

the U.S.[20] Perhaps not surprisingly, individuals in our sample are less likely to have taken the

ACT or SAT (47 percent) than individuals in the U.S. overall (lower panel of Table 1).[21]

4.2.   Is Variation in the Instruments Exogenous?

Before turning to the estimates, it is useful to demonstrate that expected age at

school entry appears to be randomly assigned and to establish that the experiment generated

random variation in the average expected age of a child's kindergarten classmates. To this

end, Table 2a gives the coefficients on ea_i and $\overline{ea}_{-i,k}$, along with their joint significance, in

models where observed characteristics are the dependent variables. In Table 2b, we present

the results from a similar exercise where the dependent variables are instead indicators for

---

[20] See Krueger and Whitmore (2001) for more information about how the CTBS and college entrance exam data were collected. Observation of other outcomes, like high school grades, is selected on our instrumental variables.

[21] For example, using the National Longitudinal Educational Study, Bedard and Dhuey (2006) report an ACT/SAT test-taking rate of 60 percent.

whether several key outcomes are observed.[22] The underlying regressions also include school fixed effects, because random assignment of children to class types took place within schools. Standard errors are consistent for heteroskedasticity and correlation of error terms among children in the same kindergarten classroom.

The estimates are consistent with random assignment. As shown in Table 2a, in only one case does expected age predict an observed correlate of test scores (column 1, respondent receipt of free or reduced price lunch), and in no instance does the average expected age a respondent's peers significantly predict his own background or the average background of his peers, or his kindergarten teacher or class characteristics (column 2).[23] The instruments are jointly significant for only three of thirteen variables, and in one of these cases (teacher race), only marginally so (column 3). As shown below, including these variables as controls has little impact on the TSLS point estimates. Table 2b shows that the instruments also do not predict observation of our dependent variables. This suggests that observations with missing values of dependent variables are random with respect to the identifying variation in a_i and $\bar{a}_{-i,k}$, and that the estimates presented below are not biased by sample selection.

5.   Results

5.1.   Conventional Estimates

Table 3 presents OLS and TSLS estimates of the coefficients on own age and peer average age from model 2. The dependent variable is the standardized average of math and reading scores at the end of kindergarten. Unless otherwise noted, all specifications hereafter

---

[22] ACT or SAT test-taking is observed for all Project STAR participants.
[23] The specification checks in Chetty et al. (2011), which include a wider range of background characteristics that are not available to us in this study, also support random assignment to kindergarten classrooms in Project STAR. We have also re-estimated our models on students attending the only kindergarten class of its size within their schools (where assignment should have been random). The estimates are noisier, but broadly consistent with those reported below for the full sample (results available on request).

include fixed effects for school attended in kindergarten, and estimation accounts for heteroskedasticity and arbitrary correlation of error terms within kindergarten classrooms.

We first present estimates of the coefficient on age at kindergarten entry from a version of model 2 that excludes $\bar{a}_{-i,k}$. This model is the same as that widely estimated in the school entry age literature, and so provides a useful benchmark. The OLS estimates, shown in Panel A, imply that STAR participants who were one year older at the start of kindergarten scored on average 0.24σ higher on a standardized test at the end of kindergarten (column 1). However, these estimates will be biased if children previously retained or delayed in entering kindergarten are selected on unobservables. As described above, we confront this possibility by comparing children who *should* have entered kindergarten with a one-year difference in age, given their birthdays; if birthday is randomly assigned, these children will be on average identical in all other ways. As shown in Panel B of the same column, TSLS estimates using expected age at kindergarten entry as an instrument imply that the test-score differential between two otherwise identical children who enter kindergarten with a one-year difference in age is a significantly higher 0.71σ.[24]

These estimates are comparable to those previously found in nationally-representative data for the U.S. For example, applying a similar identification strategy to data on a more recent kindergarten cohort, Elder and Lubotsky (2009) find that an additional year of age at school entry is associated with a 0.87σ difference in math test performance and a 0.61σ difference in reading test performance in the spring of a child's kindergarten year.[25] Like us, Elder and Lubotsky (2009) and Bedard and Dhuey (2006) also find that OLS estimates of the coefficient on age are significantly lower than their TSLS counterparts, suggesting that students who are older by actual age (but not predicted age) are negatively

---

[24] First-stage estimates for the specifications in Table 3 are presented in Table A1.
[25] When we estimate separate models by subject on our data, we arrive at TSLS estimates (with additional controls) of 0.69   and 0.524   for math and reading, respectively, at the end of kindergarten.

selected. When we estimate similar models for test scores at the end of eighth grade, we also obtain findings that are quite similar to those previously documented for the U.S.[26] Unlike Bedard and Dhuey (2006), however, we do not find that older school entrants are more likely to take a college entrance exam.[27] In general, however, the consistency of our estimates with the existing literature suggests the possible broader applicability of inferences made from our data.

5.2.   Full Model Estimates

As discussed above, the estimates presented in column 1 confound the effects of absolute age with the effects of relative age. In column 2, we add the average age of a child's kindergarten classmates to the column 1 specification; in Panel B, we instrument for this average using the average expected age of a child's kindergarten classmates, $\overline{ea}_{-i,k}$. First, notice that regardless of the method of estimation, the addition of $\overline{a}_{-i,k}$ to the model changes the coefficient on age at the start of kindergarten very little. The stability of the own age coefficient is consistent with random assignment of a child's kindergarten classmates.

Second, the TSLS estimate of the effect of having older peers is statistically significant (Panel B), while the OLS estimate is not (Panel A). The TSLS estimate isolates the effect of having older peers; the OLS estimate conflates the effect of having older peers with

---

[26] In particular, our TSLS estimates imply that a one-year increase in age at school entry is associated with a 0.215   boost in test scores nine years later – a substantially smaller difference than that observed at the end of kindergarten – and a 18.9 percentage point reduction in the likelihood of being below grade when tested. Elder and Lubotsky (2009) and Bedard and Dhuey (2006) find math test score differences of approximately 4 percentile points, or roughly 0.13 , between eighth graders who entered school in the late 1970s with a one-year difference in age. When we re-estimate the model for actual eighth grade test scores (regardless of year attended), we find that a one-year difference in entry age is associated with a 0.12   difference in math test performance. Elder and Lubotsky (2009) also find that a one-year increase in age at school entry lowers the likelihood of having been retained by eighth grade by 15.1 percentage points.

[27] Bedard and Dhuey (2006) find that being one year older (in eighth grade) raises the probability of taking the ACT or SAT by 11.1 percentage points. By contrast, we find that being one year older (at the start of kindergarten) raises the probability of taking the ACT or SAT by only 0.2 percentage points (standard error=0.024). However, it is important to note that their estimates are for individuals who were in eighth grade in the same year, not individuals who started kindergarten at the same time, so our estimates may not be strictly comparable. Our sample also has a higher minority share and is poorer than the national average (Table 1), and the reduced-form relationship between age and longer-term outcomes may be stronger in a nationally-representative sample.

the effect of having peers who are on average of lower innate ability. The estimates thus imply that having peers who are older but on average the same ability as younger peers has a greater positive impact on end-of-kindergarten test scores than having peers who are older but on average lower scoring; said differently, peers who are "overage" for grade appear to have negative spillovers.[28] Consistent with this interpretation, the estimates shown in column 3, which add controls to the model from column 2, suggest that the OLS estimate picks up the effects of other peer attributes that might be correlated with delay or retention. In particular, the OLS coefficient on $\bar{a}_{-i,k}$ falls by 37 percent (from 0.178 to 0.112), largely as a result of controlling for the fraction of kindergarten classmates who are black, eligible for free or reduced-price lunch, or female. By contrast, the TSLS coefficient falls only 14 percent (from 0.847 to 0.724) with the inclusion of controls.

Finally, as shown in column 2 of Panel B, TSLS estimates of α_{2k} are statistically significant and positive (0.847σ). From model 2′, recall that the implied effect of relative age on test scores at the end of kindergarten ($t=k$) is $-\alpha_{2k}$ (or -0.847σ), while the effect of absolute age on test performance is the sum of the coefficients on own and peer average age, $\alpha_{1k} + \alpha_{2k}$ (0.716σ + 0.847σ = 1.563σ). The estimates in column 2 thus imply that the effect of relative age on test performance at the end of kindergarten is *negative*. On the flip side, these estimates imply that omission of relative age from model 2 leads to an *understatement* of the absolute age effect. Indeed, our TSLS estimates reject the hypothesis that the relative age effect at the end of kindergarten is positive and also strongly reject the null hypothesis that the true absolute age effect at the end of kindergarten is zero, or that $\alpha_{1k} = -\alpha_{2k}$. Our

---

[28] Consistent with this idea, Lavy et al. (2012) use the fraction of an individual's peers who are repeaters or delayers in grade cohorts of Israeli high schools as a measure of his exposure to lower-ability peers. They find that higher exposure to low-ability peers, so defined, is associated with lower test scores. When we estimate similar models at the kindergarten classroom level in Project STAR, we arrive at the same conclusion (results available on request).

findings are thus inconsistent with the interpretation of existing estimates of reduced-form age effects, such as those presented in column 1, as relative age effects.

While this might be surprising from the perspective of the literature on (and popular discussion of) school entry age, our findings of (net) positive spillovers from older classmates are consistent with studies showing that students benefit from exposure to higher-scoring peers. Indeed, when we use the average expected age of a child's kindergarten classmates as an instrument for their average test performance at the end of kindergarten – instead of their true average age at the start of the year – the TSLS point estimates are in the middle of the (wide) range of prior estimates of the achievement effects of having higher-scoring peers (see Sacerdote 2011, Table 4.2) (results available on request).

How large is the positive spillover from having older kindergarten classmates? To interpret magnitudes, note that the typical child is unlikely to have the option of enrolling in a kindergarten classroom where his classmates were a full year older. Indeed, the standard deviation of $\bar{a}_{-i,k}$ is only 0.09 years (Table 1). Thus, for the average kindergartner in our sample, assignment to peers one standard deviation older on average is associated with a 0.076σ improvement in test performance at the end of kindergarten ($0.09*0.874\sigma$). With the inclusion of additional controls (column 3), the effect size falls only slightly to 0.065σ ($0.09*0.724\sigma$).

Panel A of Table 4 presents TSLS estimates of the coefficients on age at the start of kindergarten and the average age of a child's kindergarten classmates for later outcomes – test scores and grade retention eight years after the end of kindergarten, and whether the respondent took the ACT or SAT. For later test scores (column 1), the TSLS coefficient on classmate average age is smaller than it was in kindergarten, but is still a marginally significant

0.43σ (implying an effect size of approximately 0.039σ).[29] Older kindergarten classmates are also associated a higher likelihood of taking a college-entrance exam and a lower probability of being below grade, but these estimates are not significant at conventional levels. However, the estimates remain precise enough to rule out even small positive effects of being relatively old.[30] We also reject the hypothesis that absolute age has no impact for most post-kindergarten outcomes, and in the one case where we do not (took ACT or SAT, in column 3), the p-value is 0.103.

5.3. Interpretation

These findings are somewhat different than those previously documented for the U.S. by Elder and Lubotsky (2009), who as described above estimate the effects of having older schoolmates in kindergarten. Their estimated impacts of having older schoolmates on kindergarten and later test scores are smaller than ours. They also find that having older peers at the school level *increases* the likelihood that a child will be retained in grade. Though the standard errors on our estimates make it difficult for us to rule out their estimated effects, it is useful to explore whether changing our peer reference group or the source of variation in peer average age might bring our point estimates more in line.

To examine whether estimation of peer effects at the classroom instead of school level generates the differences in findings across studies, the remaining columns of Table 3 and Panel B of Table 4 present estimates where average age of peers is calculated at the school level, once again in kindergarten. We omit school fixed effects from these models and

---

[29] The effects of having older kindergarten classmates on test scores thus appear to "fade out" as a child ages. However, Cascio and Staiger (2012) point out that if a standard deviation in test scores maps into a larger difference in knowledge in later grades – i.e., if knowledge accumulates faster than it depreciates – some of this observed "fadeout" on test z-scores might be a "statistical artifact." Their estimates (column 4, Table III) imply that the effect size for spring 1994 z-scores should be scaled up by a factor of roughly 1.22 to be comparable to that for spring of kindergarten. Thus, some of the observed fadeout in our context is not a statistical artifact.

[30] For example, the lower bound on the 95 percent confidence interval on the TSLS coefficient on $\bar{a}_{-i,k}$ in column 3 is -0.032. Applying the same metric for calculating effect sizes as above, we conclude that relative age is highly unlikely to have truly generated more than a 0.0029 percentage point increase in the probability of taking a college-entrance exam (0.032*0.09).

now cluster standard errors at the school, instead of classroom, level. We continue to instrument for the average age of a child's kindergarten schoolmates with their average expected age; we also include a vector of average kindergarten schoolmate characteristics (fraction female, fraction black, and fraction free lunch) among the additional controls (column 5 of Table 3, all columns in Table 4). The TSLS coefficients on peer average age from these models follow the same pattern as those on average kindergarten classmate age. Though the estimated effect of having older schoolmates is systematically larger than that of having older kindergarten classmates, it is impossible for us to rule out that the effects of older peers in kindergarten classrooms drive the effects of older peers at the school level.[31] Thus, a different definition of the relevant peer group in kindergarten does not appear to be the factor underlying the differences between our point estimates.

On the other hand, by focusing on peers at the school grade cohort level, Elder and Lubotsky's estimates allow for the possibility that relative age effects emerge later than kindergarten. Put differently, our models might overstate the learning advantages of having older peers – or understate relative age effects – by focusing on kindergartners, for whom the mechanisms for perpetuating relative age effects (like ability tracking) may not yet be in place; indeed, the goal of kindergarten has historically been socialization, not the acquisition of academic skills. To examine this possibility, we re-estimated our models using variation in classroom age distributions in first grade, when children have historically been first exposed to reading (and ability grouping).[32] The estimates (available on request) are broadly consistent with those reported in Table 3 and Table 4; for example, they imply that having older peers (or being relatively young) raises end of first grade and 1994 test scores and

---

[31] We lack sufficient data to estimate the peer effects at the kindergarten school cohort and kindergarten classroom level simultaneously. Also note that these estimates are not identified using experimental variation, so should be viewed with caution. Indeed, when we perform the tests in Table 2a substituting the average expected age of kindergarten schoolmates for the average expected age of kindergarten classmates (and excluding school fixed effects), average expected age of kindergarten schoolmates is a highly significant predictor of six observables.

[32] All specification checks that we employed in Tables 2a and 2b pass for the first-grade classroom analysis.

SAT/ACT taking rates and reduces the probability of subsequent grade retention. However, the positive spillovers of older first grade classmates are generally smaller than those from having older kindergarten classmates.

Yet another possible explanation, mentioned above, may lie in our use of different sources of variation in peer average age. We exploit idiosyncratic variation in birthday distributions across classrooms within Tennessee elementary schools, while Elder and Lubotsky combine variation in birthday distributions across schools with differences across states in school-entry cutoff dates. Thus, Elder and Lubotsky's estimates in part represent the effects of peers who are older because of school entry age regulations, while ours represent the effects of older classmates holding the school entry age regime fixed. Our findings demonstrate that – all else equal – higher quality peers (i.e. older, which proxies for higher scoring and likely better behavior) positively impact their classmates' achievement. While we are unable to explore the hypothesis with our data from only one state, when school-entry regulations, not variation in birthdays, drive an increase in classmates' ages, as in Elder and Lubotsky's case, curricula may be more advanced and expectations of students may be heightened. In this case, any achievement advantages that come from having older peers may be counterbalanced if not overwhelmed by the potential for more advanced curricula to draw attention to a relatively young child's inability to keep up.[33]

6. Heterogeneity

As earlier noted and as the above discussion makes evident, α_{2t} is in practice a reduced-form parameter. It captures the net result of the competing achievement effects of being relatively old and of having older peers. Our estimates suggest that, on average, the

---

[33] Alternatively, Elder and Lubotsky's estimates may be subject to more attenuation bias from measurement error than those presented here. They observe only some of a given student's schoolmates, while we observe all kindergartners in a given classroom or school. However, this hypothesis cannot explain the sign difference on our estimates for grade retention.

positive spillovers from older peers win out. Thus, while an effect of being relatively old might exist, it is on average overwhelmed by the benefits of having older classmates.

A brief analysis of heterogeneity in our estimates further supports this interpretation. The likelihood of delayed school entry is more common for some subgroups in the population. For example, boys are much more likely to be held back than girls. Children from higher-income families are also more likely to start kindergarten at an age older than expected (West, Meek, and Hurst, 2000). Relative age effects may be more important for these groups. If this is the case, the positive spillovers of having older kindergarten classmates found in the full sample should be less evident – or not evident at all – for these groups.

To investigate this possibility, we estimated separate models by gender and free or reduced-price lunch status in kindergarten. We present the TSLS estimates by gender in Table 5 and by free lunch status in Table 6. The estimates suggest that being relatively old may have a greater benefit for boys than for girls and for higher-income children than for lower-income children. In particular, the positive effects of older peers tend to be larger and more often statistically significant for girls – particularly for test scores and the ACT/SAT taking rate (Table 5, Panel B) – and for children receiving free or reduced-price lunch (Table 6, Panel B). The difference in TSLS coefficients on α_{2t} between lower-income and higher-income children is starker than that between boys and girls, and in the case of taking the ACT or SAT, it is in fact statistically significant.

While the estimates are somewhat too imprecise to draw strong conclusions, this heterogeneity in our estimates is consistent with the existence of a relative age effect. Yet, even for these boys and higher-income children, the net effect of exposure to older peers is positive.

7.   Conclusion

In this paper, we have estimated the net effects of having older peers – and being young relative to one's peers – using data from an experiment where children of the same age were randomly assigned to different classrooms in kindergarten. We find that, holding own age constant, children who were young relative to their kindergarten classmates generally performed no worse on achievement tests, were no more likely to be retained, and were no less likely to take the ACT or SAT. In fact, though some of our estimates are imprecise, having older classmates appears to have made the average child better off. We arrive at similar conclusions when we use a child's kindergarten schoolmates as a peer group, but the estimates are less precise. And while our analysis of heterogeneity suggests that relative age effects might be more present for boys and higher-income children – consistent with the higher incidence of delay among these groups – on net, boys and higher-income children still appear to benefit from having older kindergarten classmates.

These findings suggest the reduced-form age effects that have frequently been estimated in the literature on school entry age are better interpreted as absolute age effects than relative age effects. In other words, while our data do not allow us to determine whether the own age effects reflect differences in developmental trajectories, age at observation, or true effects of entering school later, there is something about being older in an absolute sense that drives the observed positive impacts. The driving factor does not appear to be being one of the oldest in the classroom.

Our findings are also consistent with broader peer effects literature documenting positive spillovers from having higher-quality peers. More specifically, they add another dimension to the evidence on how kindergarten classrooms matter over the long term (Chetty et al. 2011): The ages of a child's kindergarten classmates – a feature of the

23

kindergarten experience than can be easily manipulated by school principals – appear to have a lasting impact on his or her own school outcomes.
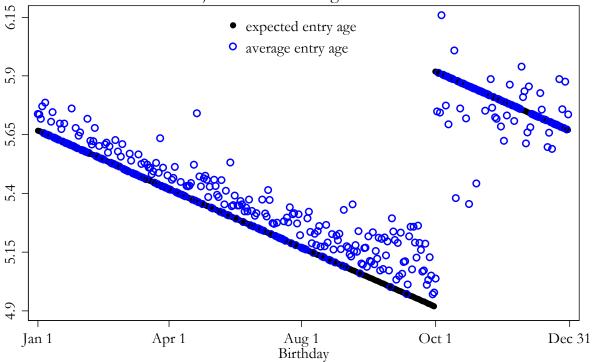
Acknowledgments

References

Angrist, J., Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, Princeton University Press.

Bedard, K., Dhuey E., 2006. The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. The Quarterly Journal of Economics 121 (4), 1437-1472.

Bedard, K., Dhuey E., 2012. School Entry Policies and Skill Accumulation Across Directly and Indirectly Affected Men. Journal of Human Resources 47 (3), 643-683.

Black, S., Devereux, P., Salvanes K., 2011. Too Young to Leave the Nest? The Effects of School Starting Age. Review of Economics and Statistics 93 (2) 455-467.

Boozer, M.A., Cacciola, S.E., 2001. Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data. Yale University Economic Growth Center Working Paper 832.

Carrell, S., Hoekstra, M., 2010. Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. American Economic Journal: Applied Economics 2 (1), 211-228.

Cascio, E.U., Lewis, E.G., 2006. Schooling and the Armed Forces Qualifying Test: Evidence from School Entry Laws. The Journal of Human Resources 41 (2), 294-318.

Cascio, E.U., Staiger, D., 2012. Knowledge, Tests, and Fadeout in Educational Interventions. NBER Working Paper 18038.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D., Yagan, D., 2011. How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. Quarterly Journal of Economics 126 (4), 1593-1660.

Datar, A., 2006. Does Delaying Kindergarten Entrance Give Children a Head Start? Economics of Education Review 25 (1), 43-62.

Dee, T.S., 2004. Teachers, Race, and Student Achievement in a Randomized Experiment. The Review of Economics and Statistics 86 (1), 195-210.

Dee, T.S., Keys, B.J., 2004. Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment. Journal of Policy Analysis and Management 23 (3), 471-488.

Ding, W., Lehrer, S.F., 2007. Do Peers Affect Student Achievement in China's Secondary Schools? Review of Economics and Statistics 89 (2), 300-312.

Dobkin, C., Ferreira, F., 2010. Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes? Economics of Education Review 29 (1), 40-54.

Duflo, E., Dupas, P., Kremer, M., 2011. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. American Economic Review 101 (5), 1739-1774.

Elder, T.E., Lubotsky, D.H., 2009. Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers. The Journal of Human Resources 44 (3), 641-683.

Figlio, D., 2007. Boys Named Sue: Disruptive Students and Their Peers. Education Finance and Policy 2 (4), 376-394.

Fredrikksson, P., Öckert, B., 2006. Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance. IFAU Working Paper 2006:12.

Gootman, E., 2006. Preschoolers Grow Older as Parents Seek an Edge. The New York Times (October 19, 2006).

Gormley, W.T., Gayer T., 2005. Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program. The Journal of Human Resources 40 (3), 533-558.

Graham, B.S., 2008. Identifying Social Interactions through Conditional Variance Restrictions. Econometrica 76 (3), 643-660.

Hanushek, E.A., Kain, J.F., Markman, J.M. Rivkin, S.G., 2003. Does Peer Ability Affect Student Achievement? Journal of Applied Econometrics 18 (5), 522-544.

Krueger, A.B., 1999. Experimental Estimates of Education Production Functions. The Quarterly Journal of Economics 114 (2), 497-532.

Krueger, A.B., Whitmore, D.M., 2001. The Effect of Attending a Small Class in the Early Grades on College Test-Taking and Middle School Test Results: Evidence from Project STAR. The Economic Journal 111, 1-28.

Kugler, A., Imberman S., Sacerdote, B., 2012. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. American Economic Review 102 (5), 2048-2082.

Lavy, V., Paserman M.D., Schlosser, A., 2012. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. The Economic Journal 122, 208-237.

McCrary, J., Royer, H., 2011. The Effect of Maternal Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. American Economic Review 101 (1), 158-195.

McEwan, P., Shapiro, J., 2008. The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates using Exact Birth Dates. The Journal of Human Resources 43 (1), 1-29.

Puhani, P.A., Weber, A.M., 2005. Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age at School Entry in Germany. IZA Discussion Paper 2005.

Sacerdote, B., 2011. Peer Effects in Education: How Might They Work, How Big are They, and How Much Do We Know Thus Far? In: Hanushek, E.A., Machin, S., Ludger Woessman, L. (Eds.), Handbook of the Economics of Education, ed. 1, vol. 3, no. 3, Elsevier, pp. 249-277.

Schanzenbach, D.W., 2006. Classroom Gender Composition and Student Achievement: Evidence from a Randomized Experiment. Mimeo, University of Chicago.

Schanzenbach, D.W., 2007. What Have Researchers Learned From Project STAR? *Brookings* Papers on Education Policy 2006/07, 205-228.

Weil, E., 2007. When Should a Kid Start Kindergarten? The New York Times Magazine (June 3, 2007).

West, J., Meek, A., Hurst, D., 2000. Children Who Enter Kindergarten Late or Repeat Kindergarten: Their Characteristics and Later School Performance. NCES 2000-039. Washington, D.C., U.S. Department of Education.

Whitmore, D.M., 2005. Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. American Economic Review 95 (2), 199-203.

Figure 1. Expected and Average Kindergarten Entry Age by Birthday
Project STAR Kindergarten Cohort

Notes: Figure plots the age at which Project STAR participants born on each day of the calendar year would have been expected to enter kindergarten, given Tennessee's regulation that entering kindergartners must be aged five by September 30 (darkened circles). Figure also plots the average age of Project STAR participants on September 1, 1985 (hollow circles).

Table 1
Descriptive statistics for Project STAR kindergarten cohort

	Mean	Standard deviation	N
	(1)	(2)	(3)
A. Age variables			
Expected age at the start of kindergarten	5.38	0.28	6248
Age in at the start of kindergarten	5.43	0.35	6248
Average expected age of kindergarten classmates	5.38	0.07	6248
Average age of kindergarten classmates	5.43	0.09	6248
B. Demographic/SES variables			
Female	0.49	-	6248
Black	0.33	-	6248
Free/reduced-price lunch (in kindergarten)	0.49	-	6248
C. Other characteristics of kindergarten classmates			
Fraction female	0.49	0.12	6248
Fraction black	0.33	0.41	6248
Fraction free/reduced-price lunch (in K)	0.49	0.28	6248
D. Kindergarten teacher characteristics			
Black	0.17	-	6248
Has MA	0.35	-	6248
Has <2 years experience	0.10	-	6248
E. Kindergarten class characteristics			
Small	0.30	-	6248
Regular with side	0.35	-	6248
F. Outcome variables			
Standardardized test score, end of kindergarten	0.00	1.00	5719
Standardardized test score, 8 years later	0.12	0.99	4436
Ever retained, 8 years later	0.25	-	4508
Took ACT/SAT	0.47	-	6248

Notes: Sample includes individuals with non-missing demographic/SES variables, kindergarten classmate characteristics, kindergarten teacher characteristics, and kindergarten class characteristics. See text for more details.

Table 2a
Predictive power of instrumental variables for background characteristics

| | Coefficient (standard error) on: | | |
| | Expected age at start of K | Expected average age of K classmates | *P* on joint significance |
Dependent variable:	(1)	(2)	(3)
Demographic/SES variables			
Female	0.013	-0.013	0.826
	(0.022)	(0.095)	
Black	0.007	0.019	0.732
	(0.010)	(0.041)	
Free/reduced-price lunch (in kindergarten)	-0.054 ***	-0.077	0.018
	(0.019)	(0.092)	
Other characteristics of kindergarten classmates			
Fraction female	-0.002	-0.024	0.878
	(0.005)	(0.091)	
Fraction black	0.001	0.02	0.885
	(0.002)	(0.040)	
Fraction free/reduced-price lunch (in kindergarten)	-0.004	-0.127	0.028
	(0.005)	(0.091)	
Kindergarten teacher characteristics			
Black	-0.018	-0.315	0.096
	(0.012)	(0.220)	
Has MA	-0.016	-0.303	0.616
	(0.019)	(0.352)	
Has <2 years experience	-0.003	-0.047	0.781
	(0.015)	(0.274)	
Kindergarten class characteristics			
Small	0.032	0.562	0.226
	(0.024)	(0.440)	
Regular with aide	0.009	0.127	0.296
	(0.024)	(0.435)	

Notes: Each row represents a different regression. All regressions are based on 6248 observations and include school fixed effects. Standard errors are clustered on kindergarten classroom. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 2b
Predictive power of instrumental variables for observation of dependent variables

Dependent variable is dummy=1 if observed:	Coefficient (standard error) on:		
	Expected age at start of K	Expected average age of K classmates	*P* on joint significance
	(1)	(2)	(3)
Standardardized test score, end of kindergarten	0.009	0.073	0.376
	(0.012)	(0.061)	
Standardardized test score, 8 years later	0.031	-0.033	0.293
	(0.020)	(0.075)	
Ever retained, 8 years later	0.031	-0.019	0.297
	(0.020)	(0.075)	

Notes: Each row represents a different regression. All regressions are based on 6248 observations and include school fixed effects. Standard errors are clustered on kindergarten classroom. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3

Estimated effects of entry age and peer average entry age on end-of-kindergarten test scores

	Dependent variable: Average test score, end of K (standardized)				
	(1)	(2)	(3)	(4)	(5)

A. Ordinary least squares

	(1)	(2)	(3)	(4)	(5)
Age at start of kindergarten	0.242***	0.244***	0.283***	0.250***	0.277***
	(0.038)	(0.039)	(0.037)	(0.045)	(0.038)
Average age of peers:					
in kindergarten classroom		0.178	0.112		
		(0.267)	(0.251)		
in kindergarten school cohort				0.907	-0.258
				(0.871)	(0.767)
R-squared	0.23	0.23	0.30	0.01	0.12
Root MSE	0.885	0.885	0.844	0.997	0.942
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$		0.129	0.133	0.194	0.981

B. Two-stage least squares[a]

	(1)	(2)	(3)	(4)	(5)
Age at start of kindergarten	0.707***	0.716***	0.683***	0.732***	0.699***
	(0.057)	(0.058)	(0.055)	(0.061)	(0.057)
Average age of peers:					
in kindergarten classroom		0.847**	0.724*		
		(0.407)	(0.374)		
in kindergarten school cohort				2.538*	2.314
				(1.351)	(1.531)
Root MSE	0.900	0.901	0.856	1.017	0.963
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$		0.0003	0.0004	0.021	0.057
Observations	5,719	5,719	5,719	5,719	5,719
Controls[b]?	N	N	Y	N	Y
School fixed effects?	Y	Y	Y	N	N

Notes: Each column and panel of the table presents estimates from a different regression. The dependent variable in all regressions is the standardized average of reading and math Stanford Achievement Test scores at the end of kindergarten. Standard errors (in parentheses) are clustered on kindergarten classroom in columns 1 to 3 and on school attended in kindergarten in columns 4 and 5. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.
[a] The instrument for a child's age at the start of kindergarten is his expected age given his birthday and the September 30th kindergarten entry cutoff birthdate in Tennessee. The instrument for the average age of child's peers is their average expected age.
[b] Dummies for whether child is female, black, or received free/reduced-price lunch in K; fractions of K classmates with these characteristics; whether the kindergarten teacher is black, has an MA, or has 0 to 1 years of experience; and dummies for whether kindergarten class is small or regular sized with teacher's aide. The specifications in column 5 also include the fractions of a child's schoolmates in kindergarten who were female, black, or received free or reduced-price lunch.

Table 4

TSLS estimates of the effects of entry age and peer average entry age on later outcomes

	Dependent variable: Eight years later:		
	Average score	Retained in any grade	Took ACT or SAT
	(1)	(2)	(3)
A. Entry age + classroom level peer effect			
Age at start of kindergarten	0.221 ***	-0.191 ***	0.004
	(0.065)	(0.026)	(0.024)
Average age of peers in K classroom	0.430 *	-0.136	0.172
	(0.248)	(0.115)	(0.106)
Root MSE	0.890	0.405	0.465
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$	0.010	0.006	0.103
B. Entry age + school level peer effect			
Age at start of kindergarten	0.225***	-0.192***	0.004
	(0.069)	(0.025)	(0.027)
Average age of peers in K school cohort	0.789	-0.459	0.205
	(0.630)	(0.316)	(0.279)
Root MSE	0.902	0.413	0.469
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$	0.116	0.0431	0.456
Observations	4,436	4,508	6,248

Notes: Each column and panel of the table presents estimates from a different regression. All models are estimated using two-stage least squares. The specification in Panel A is the same as that in Table 3 column 3. The specification in Panel B is the same as that in Table 3 column 5. Standard errors (in parentheses) are clustered on kindergarten classroom in Panel A and on school attended in kindergarten in Panel B. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5

TSLS estimates of the effects of entry age and peer average entry age, by gender

	Average Score End of K (1)	Average score (2)	Retained in any grade (3)	Took ACT or SAT (4)
		Dependent variable: Eight years later:		
		A. Male		
Age at start of kindergarten	0.645***	0.343***	-0.279***	0.059
	(0.073)	(0.101)	(0.042)	(0.038)
Average age of peers in K classroom	0.627	0.190	-0.154	-0.003
	(0.459)	(0.381)	(0.162)	(0.142)
Root MSE	0.843	0.978	0.445	0.463
P: $\alpha_{age} = - \alpha_{average\ age\ of\ peers}$	0.009	0.174	0.009	0.693
Observations	2,936	2,191	2,234	3,213
		B. Female		
Age at start of kindergarten	0.708***	0.122	-0.115***	-0.048
	(0.077)	(0.075)	(0.034)	(0.033)
Average age of peers in K classroom	0.824**	0.613**	-0.062	0.331**
	(0.407)	(0.308)	(0.152)	(0.160)
Root MSE	0.868	0.799	0.362	0.468
P: $\alpha_{age} = - \alpha_{average\ age\ of\ peers}$	0.000	0.019	0.250	0.079
Observations	2,783	2,245	2,274	3,035
P for age: $\alpha_{male} = \alpha_{female}$	0.000	0.007	0.000	0.046
P for average age: $\alpha_{male} = \alpha_{female}$	0.586	0.771	0.482	0.290

Notes: Each column and panel of the table presents estimates from a different regression. All models are estimated using two-stage least squares and are based on the specification presented in Table 3, column 3. Standard errors (in parentheses) are clustered on kindergarten classroom. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6

TSLS estimates of the effects of entry age and peer average entry age, by free/reduced-price lunch status

	Average Score End of K (1)	Average score (2)	Retained in any grade (3)	Took ACT or SAT (4)
	Dependent variable: Eight years later:			
	A. Not free/reduced-price lunch in K			
Age at start of kindergarten	0.773***	0.231***	-0.188***	0.029
	(0.081)	(0.086)	(0.034)	(0.036)
Average age of peers in K classroom	0.555	0.267	0.003	-0.003
	(0.413)	(0.297)	(0.124)	(0.154)
Root MSE	0.892	0.850	0.347	0.473
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$	0.002	0.097	0.151	0.870
Observations	2,951	2,398	2,420	3,217
	B. Free/reduced-price lunch in K			
Age at start of kindergarten	0.601***	0.186*	-0.193***	-0.000
	(0.075)	(0.095)	(0.043)	(0.038)
Average age of peers in K classroom	0.900*	0.429	-0.318	0.345**
	(0.494)	(0.459)	(0.205)	(0.150)
Root MSE	0.805	0.925	0.461	0.454
P: $\alpha_{age} = -\alpha_{average\ age\ of\ peers}$	0.005	0.193	0.013	0.026
Observations	2,768	2,038	2,088	3,031
P for age: $\alpha_{lunch} = \alpha_{no\ lunch}$	0.012	0.487	0.034	0.730
P for average age: $\alpha_{lunch} = \alpha_{no\ lunch}$	0.183	0.549	0.127	0.043

Notes: Each column and panel of the table presents estimates from a different regression. All models are estimated using two-stage least squares and are based on the specification presented in Table 3, column 3. Standard errors (in parentheses) are clustered on kindergarten classroom. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A1
First-stage estimates for Table 3 TSLS models

	(1)	(2)	(3)	(4)	(5)
	A. Dependent variable: Age at start of K				
Expected age at start of K	0.812***	0.813***	0.816***	0.816***	0.818***
	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)
Average expected age peers in K classroom		0.067	0.062		
		(0.052)	(0.051)		
Average expected age of peers in K school cohort				0.225	0.099
				(0.195)	(0.186)
Partial F-stat on instrument(s)	2916	1506	1523	1334	1298
Observations	5,719	5,719	5,719	5,719	5,719
	B. Dependent variable: Average age of peers				
Expected age at start of K		0.001	0.001	0.005*	0.003
		(0.003)	(0.003)	(0.003)	(0.002)
Average expected age peers in K classroom		0.855***	0.855***		
		(0.050)	(0.050)		
Average expected age of peers in K school cohort				1.104***	0.953***
				(0.183)	(0.174)
Partial F-stat on instrument(s)		1075	1087	107.7	154.2
Observations		5,719	5,719	5,719	5,719
Controls[a]?	N	N	Y	N	Y
School fixed effects?	Y	Y	Y	N	N

Notes: Standard errors (in parentheses) are clustered on kindergarten classroom in columns 1 to 3 and on school attended in kindergarten in columns 4 and 5. ***, **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

[a] Dummies for whether child is female, black, or received free/reduced-price lunch in K; fractions of K classmates with these characteristics; whether the kindergarten teacher is black, has an MA, or has 0 to 1 years of experience; and dummies for whether kindergarten class is small or regular sized with teacher's aide. The specifications in column 5 also include the fractions of a child's schoolmates in kindergarten who were female, black, or received free or reduced-price lunch.