NBER WORKING PAPER SERIES

FIRST IN THE CLASS? AGE AND THE EDUCATION PRODUCTION FUNCTION

Elizabeth Cascio
Diane Whitmore Schanzenbach

Working Paper 13663
http://www.nber.org/papers/w13663

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2007

We are grateful to Sandra Black, Kristin Butcher, David Card, Damon Clark, Ethan Lewis, Jens Ludwig, Heather Royer, Douglas Staiger, and seminar participants at Case Western Reserve University, Dartmouth College, the Federal Reserve Bank of San Francisco, the University of Florida, Wellesley College, and Association for Public Policy Analysis and Management Fall Conference for comments that helped to improve the paper. All errors are our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

First in the Class? Age and the Education Production Function
Elizabeth Cascio and Diane Whitmore Schanzenbach
NBER Working Paper No. 13663
December 2007
JEL No. I20,J18,J24

ABSTRACT

Older children outperform younger children in a school-entry cohort well into their school careers. The existing literature has provided little insight into the causes of this phenomenon, leaving open the possibility that school-entry age is zero-sum game, where relatively young students lose what relatively old students gain. In this paper, we estimate the effects of relative age using data from an experiment where children of the same biological age were randomly assigned to different classrooms at the start of school. We find no evidence that relative age impacts achievement in the population at large. However, disadvantaged children assigned to a classroom where they are among the youngest students are less likely to take a college-entrance exam than others of the same biological age. Controlling for relative age also reveals no long-term effect of biological age at school entry in the aggregate, but striking differences by socioeconomic status: Disadvantaged children who are older at the start of kindergarten are less likely to take the SAT or ACT, while the opposite may be true for children from more advantaged families. These findings suggest that, far from being zero-sum, school-entry age has far-reaching consequences for the level of achievement and achievement gaps between the rich and poor.

Elizabeth Cascio
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
elizabeth.u.cascio@dartmouth.edu

Diane Whitmore Schanzenbach
Harris School
University of Chicago
1155 E. 60th Street
Chicago, IL 60637
schanzenbach@uchicago.edu

I. Introduction

Older children outperform younger children in the same entry cohort well into their school careers.[1] This phenomenon has come to be thought of as a "relative-age effect" (Bedard and Dhuey, 2006), whereby a child's standing in his class by virtue of his relative maturity permanently boosts his achievement – perhaps by placement into the first reading group, or through the self-confidence and attention that come from being at the top of the class.[2] This idea has encouraged many parents to "red-shirt" their children, that is, to provide them with an extra year to grow bigger and smarter before starting kindergarten, just like athletes might delay in joining the college team to get the most out of their years of eligibility.[3] In some communities, this practice has become prevalent enough to be considered "an epidemic," creating a mass of kindergartners "pushing [age] seven" (Gootman, 2006); nationwide, it is surprisingly common.[4]

However, this behavior – as well as the academic assertion that relative age is important – contradicts years of research establishing the importance of starting school young. If children enter school at younger ages, they are subject to more compulsory education, which has been shown to raise earnings (Angrist and Krueger, 1991) and improve a number of other life outcomes (e.g., Lleras-Muney, 2005; Lochner and Moretti, 2004; Oreopoulos, Page, and Stevens, 2006). The marginal education dollar is also thought to be more effective if spent at younger ages, when the brain is still rapidly developing, particularly if there are complementarities between

---

[1] Bedard and Dhuey (2006) use data from a number of countries. Similar findings have been documented in country-specific studies of the United States (Datar, 2006; Elder and Lubotsky, 2007), Sweden (Fredriksson and Öckert, 2006), Chile (McEwan and Shapiro, 2006), and Germany (Puhani and Weber, 2005), among others.
[2] In a sample of school-entry law compliers from the Early Childhood Longitudinal Study of Kindergartners (ECLS-K), students in the oldest twelfth of the cohort had a math score 0.46 standard deviations higher than those in the youngest twelfth at the start of kindergarten. The comparable female-male difference was 0.10 and white-black difference was 0.58 (authors' calculations).
[3] The term "red-shirting" supposedly originates from the red jerseys worn in scrimmages by college athletes delaying their eligibility. See Gootman (2006) and Weil (2007) for popular press coverage of this phenomenon.
[4] In the mid-1990s, 17 percent of U.S. male second graders born between August and October – children with a high probability of being relatively young –entered school after they were first eligible. Among white non-Hispanic boys, this figure was 23 percent (authors' calculations from the 1995 National Household Education Survey).

early skills and later learning (Cunha et al., 2006). If relative age were all that mattered for long-term outcomes, there would be limited scope for early education to raise well-being: It would not matter how old a child was when he began school, as long as he was the eldest in the class.

But does relative age in fact matter at all? To date, most studies have simply estimated the relationship between outcomes and age at the start of kindergarten. However, it is not just the case that older kindergartners tend to be bigger and (temporarily) smarter *in relation to* their peers; they are also *absolutely*, or biologically, older, not just at the beginning of school but throughout their entire school careers. Maturity at the start of school may be an important determinant of a child's academic trajectory regardless of whether he is older or younger than his classmates, and age is correlated with a host of other factors that potentially correlate with achievement. For example, older school entrants may have better outcomes simply because they have accumulated more enriching experiences outside of school – more years of preschool, more visits to museums and books read with parents – not because they have benefited from their relative maturity.

To disentangle relative- from absolute-age effects, it would be helpful for two data requirements to be met. The first is a data set that gives *both* a child's own age at school entry *and* the age composition of his classmates. With few exceptions (Elder and Lubotsky, 2007; Fredriksson and Öckert, 2006), existing studies have relied on data where the latter cannot be observed, making it difficult to compare individuals who are old relative to their classmates to individuals who are relatively young while holding absolute age constant. Second, if such data were available, it would be preferable for a child's classmates to be randomly assigned. This rarely happens in practice, since school administrators may match incoming kindergartners to classes based on maturity, or parents may lobby to have their children placed in classrooms where they are among the eldest students. Without this condition met, however, one cannot easily argue that a child's relative age is unrelated to unobserved determinants of his own success.

We confront these challenges using data from one of the largest educational experiments ever undertaken in the United States – Project STAR (Student-Teacher Achievement Ratio), initiated in Tennessee in the fall of 1985. Though designed to study the effects of reduced class size (Schanzenbach, 2007), Project STAR is ideal for our purposes. In each of 79 participating schools, program administrators collected background information on all kindergartners, including children's ages and the classrooms to which they were assigned. Most kindergartners were also randomly assigned to classrooms within schools, making their classmates' ages unpredictable on the basis of their own characteristics and thus generating exogenous variation in relative age. Furthermore, for most participants, data are available on achievement tests in kindergarten through eighth grade, and whether the ACT or SAT was taken while the student was in high school – a strong predictor of eventual educational attainment. We are able to replicate the findings of earlier reduced-form studies for these outcomes, suggesting that our inferences about relative-age effects are not driven by the fact that these data are not nationally representative.

Our identification strategy combines the empirical approach used in previous reduced-form studies with variation in classroom age distributions induced by the experiment. In particular, to identify the absolute-age effect, we compare children who should have entered school at different ages given their birthdates and the rules governing school entry.[5] Given a child's "expected" age at the start of kindergarten, we then estimate the effect of relative age by comparing children who should have been young among their randomly-assigned classmates to those who should have been relatively old. Because the correct parameterization of the relative-

[5] For example, a Project STAR kindergartner whose birthday was on September 30 should have started school just before turning age five, while his counterpart born on October 1 should have been one full year older. If expected age is unrelated to unobservable determinants of outcomes, it can be used as an instrument for actual age. See Section II.

age effect is unknown, we estimate models where relative age is defined in a number of alternative ways and find that our estimates are quite robust. We also obtain similar inferences when we define relative age using different peer groups, including a child's first-grade classmates and all other children entering kindergarten in the same school.

We find no evidence that relative age matters for test scores or the likelihood of taking a college-entrance exam in the population at large. However, a disadvantaged child assigned to a kindergarten classroom where she is among the youngest students is less likely to take the ACT or SAT, holding absolute age constant. That this is not also true for more advantaged children suggests that subsequent investments may compensate for any negative effects of being relatively young at the start of school. Given that there is no relative-age effect on the earlier test scores of disadvantaged students, the causal link between relative age and long-term outcomes is likely non-cognitive, related more to motivation or self-confidence than to ability grouping.

Taken at face value, these estimates suggest that it might be desirable for disadvantaged children to postpone entrance into kindergarten. However, the opposite is likely to be true. In particular, we also find that disadvantaged school entrants who are biologically older are less likely to take the ACT or SAT than their biologically younger counterparts of the same relative age. For other children, it is possible to rule out a negative effect of the same magnitude, and there is suggestive evidence that school-entry age is positively related to the likelihood of taking one of these tests. The most promising explanation for this pattern of findings is that children from disadvantaged backgrounds are likely to receive an extra year of relatively low-quality care in lieu of kindergarten, while for more advantaged children, another year at home or in preschool may be at least as productive as time spent in school.

Regardless of the explanation, our finding that biological age at school entry predicts long-run well-being suggests that policies which manipulate this age, such as changes to the minimum age at school entry or the expansion of publicly-funded preschools, can have lasting effects on the level of achievement. Nevertheless, we hesitate to make any strong claims about the likely effects of such interventions, which could have effects on curricula and other aspects of the education production function that are not varying systematically with age at school entry in this analysis. Further, the relative-age effects estimated here do not necessarily tell us anything about the achievement gains associated with red-shirting, as our identifying variation derives from children who entered school when they were first eligible by law. That said, the present paper establishes the first – and for the foreseeable future, only – estimates of the long-term effects of being relatively mature at the start of school using variation from an experiment.

The paper proceeds as follows. The next section discusses the identification problem, previous literature, and our empirical approach. Section III describes Project STAR, key variables, and sample characteristics, and provides some preliminary evidence that the identification strategy outlined above is likely to uncover causal effects. In Section IV, we present the baseline results, and in Section V, we present the findings from several specification checks. Section VI presents findings for different subpopulations, and Section VII concludes.

II. Identification

A. The Identification Problem

A number of recent studies have estimated the reduced-form effect on test scores and educational attainment of being older at school entry. The basic model estimated is:[6]

(1) $$y_{it} = \alpha_{0t} + \alpha_{1t} a_i + u_{it}$$

where y_{it} represents the outcome of individual i in year t, a_i is his observed age (in years) in the fall of kindergarten, and u_{it} are unobserved determinants of the outcome of interest. Two challenges

[6] For ease of exposition, we suppress covariates and subscripts for cohort and place of residence.

arise in estimating this model. First, because kindergarten retention and delay are not randomly assigned, older kindergartners may differ in unobservable ways from younger kindergartners, leading ordinary least squares (OLS) estimates of α_{1t} to be biased. Second, holding entry cohort constant, a_i is perfectly correlated with the age at which a person is observed in cross-section data and will be strongly correlated with his age relative to his classmates' at the start of school. α_{1t} is therefore a reduced-form parameter, summarizing the relationship of outcomes to three variables – school-entry age, age at observation, and relative age – at a given point in time.[7]

To address the first identification problem, researchers have generally relied on the same empirical methodology, 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. For example, a child who turns age five on September 1 is expected to enter kindergarten if she resides in state or school district with a September 1 cutoff birthday for school entry, while her counterpart who turns age five on September 2 is expected to enter kindergarten at roughly age six. A sufficient condition for this approach to uncover a consistent estimate of α_{1t} is that birthday is randomly assigned; a necessary condition is that birthday is randomly assigned in the neighborhood of the cutoff date.[8]

To date, researchers have been less successful in addressing the second identification problem. For example, Bedard and Dhuey (2006) posit that test-score differentials associated with an additional year of age at the test date – while potentially quite large when children are young – asymptote as children progress through school. While estimates of α_{1t} do fall dramatically as t increases (or as children age), whether differences in age cease to matter for test

[7] See Appendix for a more formal discussion of this point.

[8] How weak the identifying assumptions are differs across studies. Research using large data sets with exact day of birth, which has generally been for one geographic area, has sometimes required only that expected school-entry age be randomly assigned conditional on some flexible function of birthday that is smooth through the cutoff birthday threshold (e.g., Dobkin and Ferreira, 2007). We estimate a model like this below as a robustness check and find very similar results as in our baseline models. An alternative approach is to use survey data for areas with different cutoffs and condition estimation on fixed effects for quarter of birth (e.g., Bedard and Dhuey, 2006; Elder and Lubotsky, 2007).

scores during the period of compulsory school enrollment is an empirical question that is not yet resolved.[9] Nonetheless, the assumption is convenient, as it implies that α_{1t} only remains positive if there is a true causal effect of either relative age or school-entry age. Bedard and Dhuey then argue that their positive estimates of α_{1t} 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 to be non-linear. Age at school entry and relative age nonetheless remain highly correlated – and their effects are thus inseparable – in this model.[10]

In principle, there are situations where relative age varies while absolute age remains constant. However, in practice, the quasi-experiments that generate such variation may be difficult to exploit or may generate other differences between relatively old and relatively young students. For example, changes in the minimum age at school entry change the relative age of some children without changing the year in which they enter school. Bedard and Dhuey (2007) estimate the reduced-form effects of these regime changes for all children, but lack the data to separate infra-marginal individuals from individuals whose own entry ages are likely to be directly affected and hence cannot separate the effects of absolute age from relative age. Kawaguchi (2006) estimates the relative-age effect by exploiting a large drop in cohort size in Japan in 1966. While this sharply changed the distribution of births across months within two adjacent school-entry cohorts – making some children relatively young who would not otherwise have been so –

[9] The brain continues to develop until around age 30 (Sowell et al., 2003). There is also evidence that age may continue to matter for the test scores of relatively advantaged children well into adolescence, possibly due to relatively enriching experiences outside of school. Elder and Lubotsky (2007) find that test score differentials between older and younger school entrants tend to be larger among children from more advantaged backgrounds from the beginning of kindergarten and through eighth grade. Cascio and Lewis (2006) present evidence that age-based test score differences are likely to be small among minority adolescents, but may be substantial among whites.
[10] Using data pooled across countries with different cutoff dates, Bedard and Dhuey (2006) instrument for a child's observed age in fourth and eighth grade with his expected relative age in his cohort in months, which ranges between 0 (youngest) and 11 (oldest), and interpret the coefficient as the relative age effect. They include dummies for expected age at school entry in a specification check, and find that these coefficient estimates are little changed. However, as country fixed effects are also included in this model, it is identified solely off of the assumption of non-linearity in the effect of expected age at school entry. They could have also used a linear function in expected age at school entry as an instrument for age in the same specification and arrived at exactly the same estimates.

the decline in fertility was generated by superstitions that females born in the year of the Fire-Horse would be "hot-tempered and unmarriageable," raising questions about selection and stigmatization of children born that year.

In this paper, we cut the Gordian Knot of identification using data from Project STAR. Observation of complete kindergarten classrooms in Project STAR allows us estimate the effect a child's age relative to his kindergarten classmates' holding constant the combined linear effect of age at school entry and age at observation. Thus, we are able to identify the effect of relative age without imposing implausible functional form assumptions. Elder and Lubotsky (2007) and Fredriksson and Öckert (2006) take a similar approach, but measure relative age at the school level. The models estimated in these papers are subject to several biases not present here, as described below. Project STAR participants were also randomly assigned peer groups, alleviating concerns over the endogeneity of relative age that exist in these earlier studies.

B. *Empirical Model*

Our primary estimates focus on a child's relative age in his kindergarten classroom and are based on the following model:

$$y_{itk} = \beta_{0t} + \beta_{1t} a_i + \beta_{2t} f(a_i, \mathrm{A}_{-i,k}) + \beta_{3t} a_{-i,k} + \varepsilon_{itk} \tag{2}$$

where 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, a_i will be perfectly correlated with age at observation, as noted above. We refer to their combined effect, β_{1t}, as the absolute-age effect. $f(\cdot,\cdot)$ represents relative age, taking as arguments a_i and the vector of ages of other children in his kindergarten classroom, $\mathrm{A}_{-i,k}$; below, we refer to its coefficient, β_{2t}, as the relative-age effect. Thus, relative age is some (unknown) interaction between a child's own age and the ages of his peers. $a_{-i,k}$ is the average age of a child's classmates in kindergarten, and ε_{itk} denotes unobserved determinants of outcomes.

Two key behavioral assumptions underlie this model. The first is that the correct functional form for relative age is not a linear combination of a_i and $a_{-i,k}$, such as the deviation between a child's own age and the average age of his kindergarten classmates. If this were the case, it would not be possible to separate the relative-age effect from a linear peer effect in age.[11] Since we do not know the correct form of $f(\cdot,\cdot)$, we consider a number of alternatives. Because children who are relatively young by any metric are likely to have older peers on average, holding constant their own age – and older peers may exert a positive direct effect on other students' performance by being better behaved and more knowledgeable – we control directly for $a_{-i,k}$. Said differently, without controlling for $a_{-i,k}$, estimates of the relative-age effect are likely to be biased downward. Thus, one additional contribution of this paper is that we estimate the effect of having more mature peers on average at the start of school.

Second, our primary estimates measure relative age in the kindergarten classroom because of data constraints described below. Estimates of Model 2 therefore allow a test of 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. To explore possible sensitivity of our relative-age estimates to the choice of peer group, we estimate models where relative age is defined in relation to a child's first-grade classmates and to other children in the entry cohort of a child's elementary school. These estimates rely less on the experiment (in the latter case, no experimental variation

[11] Let $f(a_i, \mathrm{A}_{-i,k}) \equiv a_i - a_{-i,k}$. Then Model 2 can be rewritten as $y_{itk} = \beta_{0t} + \gamma_{1t} a_i + \gamma_2 (a_i - a_{-i,k}) + \varepsilon_{itk}$, where $\gamma_{1t} \equiv \beta_{1t} + \beta_{3t}$ and $\gamma_2 \equiv \beta_{2t} - \beta_{3t}$.

is used at all); however, they do provide insights into the causal mechanisms that underlie relative-age effects, and establish another important contribution of this paper.

As noted, Elder and Lubotsky (2007) Fredriksson and Öckert (2006) take a similar approach to separating relative- and absolute-age effects, but these studies have limitations that our data and empirical strategy allow us to overcome. First, the specifications estimated in these studies potentially confound the effects of relative age with those of having more mature peers on average. In both cases, the estimated relative-age effect will be biased downward if having more mature peers raises achievement.[12] Second, relative age is more likely to matter when differences in development or maturity play the greatest role in determining an individual's academic path. We measure relative age at the start of school, while these studies measure relative age in middle school or high school –well after most individuals have established an academic course – in some if not all specifications. Both studies also lack classroom-level data, instead using peers within the same grade and school to define relative age. When we exploit the experimental variation in Project STAR, the identifying assumptions of the model are likely to hold, but when we look across schools – and rely on non-experimental variation – they may not.

C. Identification Strategy

As with Model 1, OLS estimates of the key parameters in Model 2 are likely to be biased. In this case, however, not only is it the case that a_i is endogenous; in most data, there is also likely to be sorting of children across kindergarten classrooms, possibly on unobservable characteristics related to the maturity of other children. As a result, relative age is likely to be correlated with

[12] Using U.S. survey data, Elder and Lubotsky (2007) estimate a version of Model 2 that imposes the restriction that relative age is the deviation between a child's age and the average age of his schoolmates. (However, they do not strictly interpret this estimate as a relative age effect.) Fredriksson and Öckert (2006) use administrative data from Sweden to estimate a version of Model 2 where relative age is measured as child's rank in the age distribution of his ninth-grade school cohort. The average age of his schoolmates is omitted from their model.

unobserved determinants of test scores for two reasons, and the average age of a child's classmates is not likely to be randomly assigned.

To uncover consistent estimates, we combine the identification strategy described above with the random variation in class assignment in Project STAR. Specifically, we instrument for a child's age at the start of kindergarten with his expected age at school entry, ea_i, given his birthday and the school-entry cutoff birthday relevant in our data (September 30). We then use the experiment to construct an instrument for relative age – a child's expected relative age given the expected ages of his classmates, $f(ea_i, \mathrm{EA}_{-i,k})$ – the analog of observed relative age, $f(a_i, \mathrm{A}_{-i,k})$.[13] We also instrument for $a_{-i,k}$ with the average expected age of a child's peers, which we denote $ea_{-i,k}$. Given the strong correlation between ea_i and a_i (see Figure 1), it is not surprising that we find a strong correlation between these variables and their observed counterparts. Provided that the exclusion restriction on ea_i holds, these instruments should also be valid because variation in the birthday composition of a child's classmates should be random. However, if the true underlying model has heterogeneous treatment effects, the resulting two stage least squares (TSLS) estimates identify the average effect of relative age for children who comply with school-entry regulations (Imbens and Angrist, 1994). This is an important caveat, as we are not identifying the gain to being relatively old for a child who is selected into delay by his parents.

In principle, random assignment of children to classrooms implies that observed ages of a child's peers, not just expected ages, are uncorrelated with his own academic potential. If this were the case, an alternative instrument for relative age would be a child's expected place in the *actual* age distribution of his peers (i.e., $f(ea_i, \mathrm{A}_{-i,k})$), and $a_{-i,k}$ could be treated as exogenous.

[13] Functional-form assumptions on $f(\cdot,\cdot)$ are discussed below.

However, this assumption is not likely to hold: As shown below, children who are older than expected in kindergarten (repeaters and delayers) are negatively selected. Estimates of β_{2t} and β_{3t} would therefore also pick up the effect of having lower-ability classmates in kindergarten.[14] In contrast, the instruments described above allow us to identify how having peers of the *same* innate *ability* but *different* levels of *maturity* at the start of school affects the average child's academic performance.

III. Data

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.[15] This cohort was to have maintained its class type through third grade, after which all participants were returned to regular-sized classes. Random assignment of children to class types took place within schools, and provides an experiment that can be used to investigate other aspects of the education production function, not just the effects of class size.[16]

---

[14] Consistent with this idea, Lavy, Paserman, and Schlosser (2007) 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. The estimates are available from the authors on request.

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

[16] We are not the first to exploit the random assignment of children and teachers in Project STAR to classes, not just to class types, to gain insights into the education production function (Dee, 2004; Dee and Keys, 2004; Whitmore, 2005; Schanzenbach, 2006). 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, 2007).

Our analysis exploits the fact that most, if not all, Project STAR participants would have been randomly assigned to classrooms as a result of the experimental design.[17] We focus on measurement of relative age at the kindergarten classroom level because deviations from the experimental protocol were minimized at the experiment's beginning. As noted, we show some specifications where relative age is defined in relation to first-grade classmates to examine the sensitivity of our estimates to the choice of peer group. Because students were re-randomized between regular and aide classrooms at that point, non-random transitions across class types (and classrooms) were less problematic in first grade than they would later become.[18] We also estimate relative-age effects at the school-cohort level for the reasons outlined above, but these estimates do not rely in any way on the experiment.

A. Age and Other Explanatory Variables

With exact date of birth for STAR participants, we are able to define actual age (a_i) at the start of kindergarten and expected age at kindergarten entry (ea_i) to the day. Figure 1 plots ea_i and the average value of a_i against birthday for the STAR kindergarten cohort under the assumption that the school year began on September 1. In fall 1985, children in Tennessee were eligible to begin kindergarten if they turned age five by September 30; as a result, there is a one-year difference in ea_i between children born on September 30 and children born on October 1. Expected kindergarten-entrance age is also strongly related to age at the start of kindergarten in practice.

---

[17] Below, we examine robustness of our estimates to possible sorting across classrooms and find that our results are robust.

[18] 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, relative age in kindergarten classmates, or the average age of kindergarten classmates.

We consider several measures of relative age. The first is a child's rank in his kindergarten classroom age distribution, $\text{rank}_k(a_i)$, normalized to lie between zero (for the youngest child) and one (for the oldest child).[19] Because the correlation between $\text{rank}_k(a_i)$ and a_i is strong enough to render some of our estimates based on this measure uninformative, our analysis also considers two alternatives: (1) an indicator for whether a child is ranked in the bottom quartile of his class age distribution ($1[\text{rank}_k(a_i)<0.25]$); and (2) an indicator for whether a child's own age is more than three months below the average of his classmates ($1[a_i< a_{-i,k} -0.25]$). While these definitions may seem arbitrary, we show below that our estimates are insensitive to the choice of threshold, and that the coefficients on these variables are likely to provide a reasonable approximation to the true relative-age effect. As described above, instrumental variables are the same functions of expected instead of actual age, e.g., the instrument for $1[\text{rank}_k(a_i)<0.25]$ is $1[\text{rank}_k(ea_i)<0.25]$. Given the strength of the relationship shown in Figure 1, it is not surprising that this instrument is a strong predictor of observed relative age.

To construct average age and other characteristics of kindergarten classmates, as well as relative age, we use all available data on the Project STAR kindergarten cohort. However, we restrict the estimation sample to participants who have all background characteristics (birthday, race, gender, free lunch status in kindergarten), kindergarten classmate characteristics (average age, fraction black, fraction female, fraction receiving free lunch), and kindergarten teacher characteristics (experience, education, and race) observed.[20] Table 1 shows that the average child in our sample has a high probability of being poor or a minority. For example, 33 percent of our sample is black; by comparison, in fall 1985, only 15.4 percent of five year olds in the U.S. were

[19] We calculate a child's age rank in his kindergarten classroom, $rank_{ik}$, by lining up children from youngest ($rank_{ik}=1$) to oldest ($rank_{ik}=n_k$). Two students of the same age in a given classroom are assigned the same rank. The normalized rank measure is then $rank_k(a_i)=(rank_{ik}-1)/(n_k-1)$, where n_k is the number of students in kindergarten classroom k with non-missing age variables. Normalized age rank within first grade classrooms and schools is calculated in a similar way.

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

black.[21] However, as has been found in nationally-representative samples, children in our sample tend to be older at the start of kindergarten than expected (5.43 years old versus 5.38 years old).

B. *Outcome Variables*

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. Scores on this 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.[22] We do, however, have scores on the Comprehensive Test of Basic Skills for children in grades 5 through 8, regardless of year attended, for all participants still attending public school in Tennessee.[23] The instrumental variables do not predict observation of these later test scores.[24]

Both tests are multiple-choice standardized tests with reading and math components. To make scores on the different tests comparable, 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 suggests that the kindergartners in our sample are slightly positively selected from the

[21] There are authors' calculations from the 1985 October Current Population Survey. In the ECLS-K, 14.1 percent of kindergartners are black and 29.9 percent qualify for free or reduced-price lunch (authors' calculations).
[22] P-values on the test of the null hypothesis that the instruments do not jointly predict attrition are approximately zero, regardless of the measure of relative age employed. 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.
[23] 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.
[24] P-values on the null hypothesis that the instruments do not jointly predict non-missing test scores in either year are well above 0.1, regardless of the measure employed.

pool of all Project STAR participants, scoring on average 0.12σ above the mean in spring 1994; kindergarten was not mandatory to attend at the time.

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 considered the relationship between age and test scores in approximately 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 any 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.

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.[25] Perhaps not surprisingly, individuals in our sample are less likely to take the ACT or SAT (47 percent) than individuals in the U.S. overall.[26]

C. Is Variation in the Instruments Exogenous?

Before turning to the estimates, it is useful to establish that the experiment generated random variation in the expected age composition of a child's kindergarten classmates and to

[25] See Krueger and Whitmore (2001) for more information about how the CTBS and ACT data were collected. In addition, although we have great interest in measuring non-cognitive skills, the direct measures of these in Project STAR are poor and unusable. There are self-concept and motivation scores, but these are unreliable. Observation of other outcomes, like high school grades, is selected on our instrumental variables.

[26] On the basis of the National Longitudinal Educational Study, Bedard and Dhuey (2006) report an ACT/SAT test-taking rate of 60 percent.

demonstrate that expected age at school entry appears to be randomly assigned. To this end, Table 2 shows p-values on the joint significance of observed characteristics (listed in Table 1) in predicting each of the instrumental variables. The underlying regressions also include school fixed effects, because random assignment of children to class types took place within schools, and standard errors are consistent for heteroskedasticity and correlation of error terms among children in the same kindergarten classroom.

Column 1 shows that the observables (not including the school fixed effects) are marginally significant in predicting a child's own expected entrance age, ea_i (p-value=0.07). While this suggests that this instrument may be correlated with unobservables as well – violating the assumptions needed to identify the effect of absolute age – the covariates do not jointly predict ea_i in sub-samples with non-missing test-score data (not shown, p-values=0.14, 0.12 for non-missing test-score data in 1986 and 1994, respectively). When we control for a smooth function of birthday and identify the absolute-age effect using only the discontinuity in expected entry age at the cutoff-date threshold – and thus no longer rely on birthday to be randomly assigned over its entire support –we also obtain very similar estimates to those from the baseline model (see Table 7, Section V.C).

Estimates presented in the remainder of the table support our earlier assertion that the experiment generated random variation in a child's expected relative age and the average expected age of his peers. In particular, observables do not jointly predict $ea_{-i,k}$ (Column 2). They also do not jointly predicted $rank_k(ea_i)$, set to zero for children expected to be the youngest in their kindergarten classrooms and one for those expected to be the oldest (Column 3), or the expected relative age indicators described above (Columns 4 and 5). When we control for observables below, it is therefore not surprising that it improves the precision of our estimates without really affecting their magnitudes. Overall, these findings are consistent with the identifying assumptions of the model being satisfied.

Two final things to note are the coefficients on ea_i and $ea_{-i,k}$ in the relative-age models. First, children who are biologically older tend to be old relative to their classmates, demonstrating the problem that arises in trying to make inferences about relative-age and absolute-age effects based on estimates of Model 1. Second, holding constant a child's own age, children with younger classmates are relatively old on average. This suggests the importance of controlling for peer average age when attempting to estimate the relative-age effect.

IV. Results

A. Benchmark Estimates

Table 3 presents estimates for the standardized average of math and reading scores at the end of kindergarten.[27] To provide a benchmark to the existing literature, we begin by discussing estimates of the coefficient on age at kindergarten entry from Model 1, shown in Panel A. OLS estimates of this coefficient 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, OLS estimates will be biased downward if children previously retained or delayed in entering kindergarten are negatively selected.

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. Column 2 shows the reduced-form relationship between test scores and expected age at kindergarten entry. The predicted difference in test scores between children born on September 30 and October 1 is

[27] Unless otherwise noted, all specifications hereafter include fixed effects for school attended in kindergarten, and estimation accounts for heteroskedasticity and arbitrary correlation of error terms within kindergarten classrooms.

0.57σ at the end of kindergarten. 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 0.71σ (Column 3).[28] Adding the full vector of additional controls (Column 4) changes the TSLS point estimate little but improves its precision, as anticipated if expected age were in fact randomly assigned.

Importantly, these estimates are comparable to those previously arrived at in nationally-representative data for the U.S.[29] The first two columns of Table 4, Panel A show TSLS estimates of Model 1 (with controls) for test scores in spring 1994 (when those progressing through school normally would have reached eighth grade) and the likelihood of being below grade at this time. A one-year increase in age at school entry is associated with a 0.215σ 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. These findings are also quite similar to those previously documented for the U.S.[30]

These findings suggest the possible broader applicability of inferences made from our data. However, unlike Bedard and Dhuey (2006), we find no evidence that children who enter school at a later age are more likely to take a college-entrance exam (Column 3). In fact, we can rule out positive effects of the magnitude that they document for a nationally representative

---

[28] First-stage estimates for the specification in Column 4 of Table 3 are presented in Appendix Table A1.

[29] For example, applying a similar identification strategy to data on a recent kindergarten cohort, Elder and Lubotsky (2007) 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. When we estimate separate models by subject on our data, we arrive at TSLS estimates (with additional controls) of 0.69σ and 0.524σ for math and reading, respectively, at the end of kindergarten. Like us, Elder and Lubotsky (2007) and Bedard and Dhuey (2006) also find that OLS estimates of the coefficient on age in Model 1 are significantly lower than their TSLS counterparts, suggesting that students who are older by actual age (but not predicted age) are negatively selected.

[30] On the basis of similar models, Elder and Lubotsky (2007) and Bedard and Dhuey (2006) find math test score differences of approximately 4 percentile points, or roughly 0.13σ, between eighth graders who entered school in the late 1970s with a one-year difference in age. When we estimate the model in Column 4 of Table 4, Panel A for eighth grade test scores, we find that a one-year difference in entry age is associated with a 0.14σ difference in math test performance. Elder and Lubotsky (2007) 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.

sample of U.S. students.[31] As noted, however, our sample has a higher minority share and is poorer than the national average. The reduced-form relationship between age and school outcomes may be stronger for more advantaged children. We explore this possibility below, and find that it does have some merit, though this also suggests that existing estimates of the reduced-form relationship between kindergarten entrance age and long-term outcomes may be sensitive to sample composition. Nonetheless, our findings are consistent with other work that finds at best a weak reduced-form relationship between age at school entry and educational attainment for more recent cohorts in the U.S. (Dobkin and Ferreira, 2007).

B. The Effects of Absolute and Relative Age

An important limitation of Model 1 is that the coefficient of age at school entry captures the relationships between test scores and three distinct age concepts. Under the assumptions laid forth above, Model 2 allows us to separate the effects of a child's age in relation to her class from those of being biologically older at school entry or at any point in time thereafter. Panel B of Tables 3 and 4 presents estimates from a specification where relative age is measured with $\text{rank}_k(a_i)$. As discussed, we control directly for the average age of a child's classmates because without doing so, the coefficient on $\text{rank}_k(a_i)$ is likely to be biased downward; we return to a discussion of this peer effect below. We now instrument not only for a_i with a child's expected entry age, but also for $\text{rank}_k(a_i)$ with a child's expected rank in the expected age distribution of his kindergarten class ($\text{rank}_k(ea_i)$) and for $a_{-i,k}$ with the average expected age of a child's kindergarten classmates ($ea_{-i,k}$).

---

[31] Bedard and Dhuey's TSLS estimates imply that being one year older in eighth grade raises the probability of taking the ACT or SAT by 11.1 percentage points. The upper bound 95 percent confidence interval on our TSLS estimate is 6.2 percentage points. 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.

Consider first the estimates of this model for test scores at the end of kindergarten. The TSLS coefficient on $\text{rank}_k(a_i)$ implies that, holding constant a child's age at school entry and the school-entry age of his typical classmate, a child who is the oldest in his kindergarten classroom scores on average 0.138σ *lower* than a child who is the youngest. By contrast, a one-year increase in age at school entry is associated with 0.819σ higher test scores, holding constant relative age (Table 3, Panel B, Column 4). Though the coefficient on $\text{rank}_k(a_i)$ is not statistically significant, it is statistically distinguishable from that on a_i, suggesting that moving a child from a class where she is the youngest to one where she is the oldest is highly unlikely to have the same effect on test scores as simply being one year older at school entry (p-value=0.01).

One drawback of this specification is that estimates are imprecise due to strong collinearity between a_i and $\text{rank}_k(a_i)$. In particular, TSLS estimates of the coefficient on a_i in Panel B are roughly four times as large as they were in Panel A. As a result, in spring 1994, TSLS coefficient estimates on a_i and $\text{rank}_k(a_i)$ (Table 4, Panel B, Column 1) are too imprecisely estimated to reject the null that they are identical (p-value=0.89). We do, however, reject the hypothesis that neither absolute age nor relative age belongs in the model (p-value=0.00). This leads to the somewhat unsatisfying inference that a child's relative age in his kindergarten classroom may contribute to, but cannot fully explain, persistence in the test-score differential between older and younger school entrants.

Anecdotally, however, it is the youngest and oldest children in a classroom or school – not children in the middle of the relative-age distribution – who generate relative-age effects. To capture this possibility in a parsimonious way, Panel C of Tables 3 and 4 presents estimates from a version of Model 2 where relative age is measured with a dummy for whether a child is in the bottom 25 percent of his kindergarten classroom age distribution. The TSLS estimates of the coefficient on this indicator are negative, but small in magnitude and not statistically significant.[32] By comparison, the estimated effects of absolute age are quite large, at 0.68σ and 0.19σ in kindergarten and spring 1994, respectively. Absolute age also contributes significantly to the explanatory power of the models, while relative age does not.

One might argue that this specification is also too restrictive. For example, the specification presented in Panel C forces there to be no test-score differences among children ranked in the top three quartiles of the classroom age distribution – and so no differential effects among those who are the eldest in the class – and no test-score differences among children ranked in the bottom quartile. To examine the plausibility of these restrictions, Figure 2 plots coefficients and 95 percent confidence intervals on the coefficients on a series of indicators for different deciles of the expected rank distribution in models of actual rank and test scores. The coefficient on the indicator for expected rank decile *d* represents the predicted difference in an outcome between children in decile *d* and children expected to be the oldest in their kindergarten class ($\text{rank}_k(ea_i) \geq 0.9$), the dummy for whom is omitted to identify the model. The underlying model includes school fixed effects, ea_i, $ea_{-i,k}$, and the additional controls, and so is analogous to the reduced-form model that underlies the estimates presented in Column 4 of Table 3 and in Table 4.

Figure 2 shows that while expected rank is strongly related to a child's actual rank in his kindergarten classroom age distribution (Panel A), there is no such relationship for test scores (Panels B and C). Further, coefficients on the indicators show no clear pattern, bouncing around zero, and we cannot reject that children in any other decile of the expected age distribution perform differently than those expected to be the eldest in the class. However, relative age does appear to be partially responsible for the rate at which a child progresses through school: Panel

[32] The instrument for this indicator is an indicator for whether a child is expected to be in the bottom quartile of the classroom age distribution, given his birthday and those of his classmates.

D shows that though coefficient estimates on the deciles of the expected rank distribution are not statistically significant, they are consistently above zero for the youngest children in the class, and decrease almost monotonically in moving from the youngest to the oldest in the expected rank distribution. Estimates for grade retention from the more parsimonious specification are presented in second column of Table 4, Panel C. Holding constant entry age, children in the bottom quartile of their kindergarten classroom age distributions are on average 8 percentage points more likely to be below grade. Relative age thus explains almost half of the reduced-form relationship between entrance age and this outcome.

These estimates imply that relative age – at least when peers are limited to kindergarten classmates – significantly affects school progression, but does not affect a child's level of knowledge or skill at any point in time after school entry. Relative age might nonetheless affect educational attainment through non-cognitive channels, such as motivation or self-confidence. All specifications discussed so far imply that being relatively young slightly reduces the probability of taking the ACT or SAT, though the coefficient on relative age is small and never statistically significant (Table 4, Column 3, Panels B and C). The final panel of Figure 2 suggests that this finding is not just an artifact of how $rank_k(a_i)$ enters the models: The probability that individuals in any of the lowest nine deciles of the expected rank distribution take the ACT or SAT are not statistically distinguishable from that of individuals anticipated to be the eldest in their kindergarten classrooms. Holding constant relative age, absolute age also appears to bear no relationship to the likelihood of taking a college-entrance exam.

A final concern about the specifications presented thus far is that children who are relatively young by the rank metric may not be significantly behind relative old students in terms of development at the start of school. For example, children who are ranked low in their class age distributions will not be much younger than their classmates if classroom age distributions are highly compressed. In Panel D of Tables 3 and 4, we therefore measure relative age with an indicator for whether a child is three months younger than the average age of his classmates.[33] The TSLS estimates of the relative-age effect under this specification are slightly less precise, but quite similar in magnitude to those presented in Panel C.[34] Estimates of the relationship between absolute age and outcomes are essentially unchanged.

C. The Effects of Having Older Peers

While the evidence for the full sample is not consistent with strong relative-age effects, Tables 3 and 4 provide reasonably consistent evidence that assignment to a kindergarten classroom with more mature children has a positive impact on a student's own outcomes. In kindergarten (Table 3), being randomly assigned to peers that are one year older increases one's own score by 0.73σ in the specifications shown in Panels C and D of Table 3. To interpret the magnitude of this effect, note that the typical child is unlikely to have the option of enrolling in a kindergarten classroom where his classmates are a full year older. Indeed, the standard deviation of $a_{-i,k}$ 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 0.066σ improvement in test performance at the end of kindergarten. By 1994 (Table 4), the coefficient on classmate age is slightly attenuated to 0.44 to 0.46 (implying an effect size of approximately 0.04σ). Table 4 shows that older kindergarten peers are also associated a higher likelihood of taking a college-entrance

[33] Of course, by the measure defined in Panel C – the bottom 25 percent of the age distribution in each class – there are some children in each classroom that qualify as relatively young. This is not so for the measure in Panel D, which only classifies a child as relatively young if he is more than three months younger than his average classmate. Nonetheless, all kindergarten classes have at least one child defined as relatively young by this measure, and 27 percent of children overall have this label.

[34] We instrument with an indicator for whether an individual's expected age is more than 3 months below the expected average age of his classmates. We have also estimated reduced-form models including indicators for the deviation between expected entry age and the average expected entry age of peers. Given that the direct effects of expected entry age and average expected entry age are also included in the model, the results were even less precisely estimated than the coefficients presented in Figure 2.

exam and a lower probability of being below grade, but these results are not always precisely estimated.

We can only speculate on the mechanism that might link classmates' average maturity to one's own outcomes. For example, more mature classmates may be better behaved, making time spent in school more productive for any given child (Lazear, 2001), or more mature classmates might themselves be better teachers. The fact that the spillover persists long after children are assigned to other classrooms suggests, nonetheless, that having older kindergarten classmates may help to establish a better foundation for subsequent learning. Importantly, these findings are different than those previously documented for the U.S. by Elder and Lubotsky (2007), who estimate the effects of having older schoolmates in kindergarten using variation in birthday distributions across schools and differences across states in school-entry cutoff dates. They find that having more mature peers increases the likelihood that a child will be retained in grade and has effects on kindergarten and later test scores that are weakly positive at best. When peers are older because of school-entry regulations, not because of variation in birthdays, curricula and expectations of students may also be systematically different.[35]

V. Robustness

A. Reference Groups

So far, we have measured a child's relative age using children in his kindergarten classroom. However, if placement into subject-specific (e.g., reading) groups on the basis of ability is a key causal mechanism linking relative age to long-term outcomes, it may be

[35] Elder and Lubotsky also estimate peer effects at the level of the school-entry cohort, not at the classroom level. Below, we show that this is unlikely the reason for the differences in our peer-effect estimates, as our findings are similar at the school level. However, Elder and Lubotsky also observe only some of a given student's schoolmates, while we observe all kindergartners in a given classroom or school. Their estimates may therefore be subject to more attenuation bias from measurement error than those presented here.

unreasonable to expect them to emerge as a function of peers in kindergarten classrooms. Indeed, the goal of kindergarten has historically been socialization, not the acquisition of academic skills. Tennessee schools were also relatively late to introduce kindergartens (Cascio, 2007) and may have therefore been later in adopting the more rigorous academic standards that are common in kindergartens today (West, Denton, and Germino-Hausken, 2000).

While we lack the data to look directly at the effects of relative age on tracking, we are able to re-estimate all of our models using variation in classroom age distributions in first grade, when children have historically been first exposed to reading. The upper panels of Table 5 present our findings. Because transitions across class types became more common as Project STAR progressed, we limit our sample for this analysis to a subset of schools where these transitions were relatively uncommon.[36] The estimated effects of relative age in first-grade classrooms are substantively similar to those for relative age in kindergarten classrooms. For example, there is a significant effect of relative age only for the likelihood of being behind grade by spring 1994.[37] There is also no evidence of positive peer effects in first grade, suggesting that there may be an important interaction between having more mature classmates and curricula.

With information on all kindergartners in the 79 Project STAR schools, we are also able to estimate a model comparable to that in Model 2, but where a child's peers group is defined at the school, not classroom, level. These models capture the possibility that children who are relatively young in their cohorts are likely to be among the youngest in their classrooms at some

[36] Specifically, we limit the sample to schools in which fewer than ten percent of students in a small class in first grade were in a non-small Project STAR class in kindergarten. Other research has documented moves across class types over time at about ten percent of participants (Krueger, 1999). We look at non-random movements only relative to small classes because all regular and regular-aide classes were randomly re-assigned in first grade. Results are similar if we restrict the sample to schools that had no transitions into small classes.

[37] We have estimated similar models on data pooled across kindergarten and first grade classrooms. In these models, relative age is defined as relative age in one's classroom and average age is of peers in one's classroom, regardless of whether kindergarten or first grade. Because some children may be observed twice, yet there may be common shocks within classrooms, we calculate standard errors using the multi-way clustering technique described in Cameron, Gelbach, and Miller (2006). While the estimates from these models are slightly more precise, we still fail to reject the null that relative age does not matter for all outcomes except for the likelihood of being below grade.

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.

These estimates are presented in the lower panels of Table 5. The models exclude school fixed effects, but do include background characteristics of schoolmates among the additional controls.[38] Unsurprisingly, estimates of reduced-form entrance age effects based on this specification (Panel C) are quite comparable to those presented in Panel A of Tables 3 and 4. However, relative-age effects tend to be more precisely estimated and slightly larger in magnitude than those presented above (Panel D). In fact, children in the youngest quartiles of their kindergarten cohort distributions are a marginally significant 4.7 percentage points less likely to have taken the ACT or SAT (Column 3). However, this coefficient is not statistically significant in other specifications, such as those presented in Panels B and D in Table 4 (not shown). We therefore believe that this finding should be viewed as suggestive at best. We also find positive though less precisely estimated effects of peer average age at the school-cohort level, suggesting that the difference between our findings and those of Elder and Lubotsky (2007) are not driven by the choice of reference group. These estimates are not identified off of experimental variation, however, so they should be viewed with caution.[39]

B. *Possible Sorting Across Classrooms*

Having established that our inferences about relative-age effects are not sensitive to how relative age is defined, we now examine the robustness of our estimates to possible sorting of children across classrooms. While we primarily exploit variation in classroom assignment – not

[38] We also cluster standard errors on school, not kindergarten classroom.

[39] We strongly reject the null hypothesis that observables do not jointly predict the average expected age of a child's schoolmates. All specification checks that we employed in Table 2 pass for the first-grade classroom analysis.

class type or size assignment – Project STAR program documents indicate no clear direction about how students were allocated when there were multiple classrooms of a given size within a school. As a result, there may have been sorting across classes for a sub-sample of students, even though the specification checks shown in Table 2 suggest that the allocation was random.[40]

Our first test for sorting is to re-estimate our models for the sub-sample of respondents assigned to the only kindergarten class of its size category within their respective schools.[41] Classroom assignment *should* have been random for these children, since class type was randomly assigned and would have uniquely determined a child's teacher and classmates. The upper part of Table 6 presents, by outcome, TSLS estimates of the specifications presented in Panels A and C of Tables 3 and 4 for this sub-sample. Coefficients on the indicator for being in the bottom quartile of the kindergarten classroom age distribution (Panel B) are similar to those in the full sample. The most notable exception is that the coefficient on relative age is smaller in magnitude and no longer significant in predicting whether a child is below grade by spring 1994 (Column 3). This is primarily due to the fact that, by chance, the youngest children in the kindergarten classrooms of this sub-sample have characteristics that correlate positively with achievement.[42] Unfortunately, this also suggests that the results of this exercise must be viewed with caution.

Non-random transitions across class types – generally into small classes from regular-sized classes – are thought to have been rare in kindergarten, but did occur (Krueger, 1999).

---

[40] In particular, if a school had exactly three kindergarten classes, then randomly assigning students to a class type (small, regular, and regular with aide) was the same as randomly assigning them to a classroom (and classmates). However, in schools with more than one kindergarten class for at least one type, class assignment may not have been random. For example, if a school had four classrooms – typically two small classes, and one of each regular and regular with aide – then a class-type assignment of regular or aide would unambiguously determine a student's teacher, but principals may have had leeway to assign students across the two small classes in a non-random manner.
[41] A similar robustness test is performed in Whitmore (2005), who estimates the effect of class gender composition on test scores. Twenty-six percent of the sample attended a school with exactly three experimental classrooms. Fifty-five percent of the sample attended schools with 4 or 5 classrooms in total, so that at least one classroom was randomly assigned.
[42] For example, younger children in this sample have a higher fraction of female classmates and are less likely to have an inexperienced teacher.

While we have no reason to believe those who switched class types would have been systematically drawn from any part of the classroom age distribution, we have re-estimated Model 2 on the sub-sample to individuals not observed in small classes as an additional robustness check. The results are shown in Panels C and D of Table 6. Here, the estimated effects of relative age are quite similar to those observed for the full sample.[43] These findings also suggest that there is not an interaction between the effects of relative age or absolute age and class size. However, the coefficients on peer average age in these specifications suggest that the true long-term gains from having more mature kindergarten classmates may be non-cognitive: In smaller schools and in regular-sized classes, having older classmates in kindergarten raises the likelihood of taking a college-entrance exam, but does not affect test scores after kindergarten or the likelihood of being below grade.

C. Possible Endogeneity of Birthday

All of our models have imposed the restriction that birthday does not directly affect outcomes. However, there remains the possibility that season of birth matters for reasons beyond its determination of when a child will enter school (Bound, Jaeger, and Baker, 1995; Bound and Jaegar, 2000; Dobkin and Ferreira, 2007). Given that relative age is strongly correlated with a child's age at school entry, misspecification of the relationship between age and outcomes may also bias our estimated effects of relative age.

To examine this possibility, we re-estimated our models using a regression discontinuity approach. In particular, we aligned children by birthday in two intervals on either side of Tennessee's cutoff date (April 1 through September 30 and October 1 through March 30). We then estimated the following model:

[43] Our conclusions from these robustness checks are similar when we consider the specifications shown in Panels B and D of Tables 3 and 4. These estimates are available from the authors on request.

(3) $$y_{itk} = \gamma_{0t} + \gamma_{1t} a_i + \gamma_{2t} f\left(a_i, \mathrm{A}_{-i,k}\right) + \gamma_{3t} a_{-i,k} + \gamma_{4t} normbirth_i + \gamma_{5t} normbirth_i^2 + \gamma_{6t} after_i * normbirth_i + \gamma_{7t} after_i * normbirth_i^2 + \varepsilon_{itk}$$

where $normbirth_i$ represents the number of days between individual *i*'s birthday and September 30 ("normalized birthday"); $after_i$ is an indicator set to one if *i*'s normalized birthday is equal to one (for October 1) or higher; and all other variables are as previously defined. In estimating Model 3, we use $after_i$ as an instrument for age at the start of kindergarten instead of expected age at school entry. Thus, the model identifies the absolute-age effect solely off of the discontinuity in expected age at the cutoff date, allowing birthday to have a smooth direct effect on outcomes.

Table 7 shows TSLS estimates of this model. In general, the absolute-age effects are slightly smaller in magnitude and less precisely estimated than those presented above. However, Panel B shows that we continue to find significant relative-age effects only for the likelihood of being below grade by spring 1994. For no outcome or specification are the birthday controls jointly significant. In the other specifications considered above, we arrive at similar findings.

VI. Heterogeneity

Much like the young in an entry cohort, certain subpopulations begin school at an academic disadvantage. For boys, this disadvantage is likely to arise from natural differences in developmental trajectories by gender, while children from poorer families are likely to have received less investment in their human capital by the time they begin school.[44] Relative age may play a role in reinforcing these developmental lags. Variation across subgroups in the effects of relative age may also obscure differences in the relationship between absolute age and outcomes in the broader population.

[44] See Neal (2006) for an overview of racial gaps in human capital investments prior to school.

We present estimates of our baseline model by gender in Table 8. As one might expect, we find that the reduced-form age gradient in grade retention (Panel A) is significantly steeper for boys than girls (Columns 5 and 6). As shown in Panel B, relative age does contribute to this phenomenon for boys, though the relative-age effects for boys and girls are not statistically distinguishable. Consistent the estimates presented above, we continue to find no evidence that relative age predicts test scores or the likelihood of taking the ACT or SAT. These findings suggest that whatever disadvantages young boys have at the start of school are compensated for as they age.

Table 9 presents estimates of Models 1 and 2 separately for children who did and did not receive free or reduced-price lunch in kindergarten. The reduced-form relationship between age and test performance is stronger for the more advantaged group when children are younger (Columns 1 and 2), perhaps reflecting greater family investments outside of school. However, there is no longer a significant difference across groups in the relationship between absolute age and test scores eight years later (Columns 3 and 4). Panel B shows that there is also no evidence that relative age matters for test performance in either subpopulation.

Unlike our findings for boys, however, we find persistent, long-term relative-age effects for disadvantaged children. For example, free-lunch recipients who ranked among the youngest 25 percent in their kindergarten classrooms are 8.4 percentage points less likely to take the ACT or SAT. This suggests that early socioeconomic disadvantage is an important mediator of the relative-age effect, possibly because disadvantaged parents are less able to correct their child's academic trajectory by compensatory investments outside of school or direct intervention with school administrators. On the other hand, the administrators and teachers in schools attended by more advantaged children may be more aware that relative age can affect a child's academic path and take their own steps to correct it. In any event, our finding of a significant relative-age effect for the likelihood of taking a college-entrance exam, but not for earlier test scores, suggests that these interventions are more likely to take the form of boosting non-cognitive skills – such as motivation or self-confidence – than changing a child's academic track.[45]

Just like these relative-age effects were not revealed in our estimates for the full sample, our finding above that absolute age bears no relationship to long-term outcomes masks substantial heterogeneity in absolute-age effects by socioeconomic status. For example, regardless of the kindergarten classroom to which a free-lunch student is assigned, she is 9.7 percentage points less likely to take a college-entrance exam if she enters kindergarten one year later (Panel B, Column 7). On the other hand, among students not receiving free lunch, those who are one year older at school entry are 6.6 percentage points *more* likely to take the ACT or SAT once the effects of relative age and classmate average age are taken into account (Panel B, Column 8). The negative estimate for free-lunch students is statistically distinguishable from zero. The positive estimate for more advantaged students is imprecisely estimated, though it is possible to rule out a negative effect as large as that observed in the free-lunch population.[46]

What might explain this difference in absolute-age effects by socioeconomic status? On one hand, the suggestion that being older at school entry may lower educational attainment is consistent with previous findings that older school entrants obtain less compulsory education, reaching the minimum age of school exit having spent less time in school (Angrist and Krueger, 1991). However, persons on the margin of high school dropout are also not likely to be simultaneously on the margin of entering a four-year college, and we unfortunately lack the data to look directly at high school dropout.[47] Further, this mechanism cannot explain why school-

---

[45] Consistent with this, we find no evidence that a disadvantaged child's relative age in his first grade classrooms affects his first grade or spring 1994 test scores.

[46] Using the alternative specifications for relative age considered above yields similar findings.

[47] The minimum age of school exit for Project STAR participants (in Tennessee) was 17. The public-use Project STAR data does have information on high school graduation, but it is only available for a subset of students selected

entry age may have a positive impact on college-going for advantaged students. It also seems unreasonable that absolute maturity at the start of kindergarten could potentially serve advantaged students well – leading, for example, to greater confidence and academic placement – but reinforce disadvantage in other populations.

A more promising explanation for these findings is that how children spend the year prior to kindergarten differs dramatically by socioeconomic status. Though we lack any direct evidence of the care that Project STAR participants received prior to entering school, other data suggest that disadvantaged children are either less likely to receive center-based care or education prior to entering school or attend lower-quality programs (e.g., Magnuson and Waldfogel, 2005). The estimated absolute-age effects by socioeconomic status provide some suggestive evidence in support of this hypothesis. For example, the estimates shown in Panel B, Columns 1 and 2 imply that the test-score gap between poor and non-poor students at the end of kindergarten is substantially higher among students who are biologically older at the start of the year (0.366σ higher for every one-year increase in age). Thus, disadvantaged students appear to lose ground when they are not enrolled in school.[48]

The relevant question then becomes how productive this alternative form of care or education is in comparison to kindergarten, particularly the full-day kindergarten programs attended by STAR participants. There are few estimates of the effect of kindergarten attendance on test scores, and the credible estimates that do exist may rest on a very different counterfactual than that which would have been faced by the students in our data.[49] Estimates of the effect of a year of completed schooling on test scores for disadvantaged students in very different settings

on our instrumental variables. We do not use this variable because there is no good way to address this selection problem.

[48] Elder and Lubotsky (2007) also document a similarly-sized expansion of test-score gaps with entrance age using data from the ECLS-K.

[49] For example, DeCicca (2007) estimates the effect of full-day kindergarten on test scores in the first several years of school, but the counterfactual for all children is likely to be a half-day kindergarten program.

(Gormley and Gayer, 2005; Cascio and Lewis, 2006) are nevertheless in the neighborhood of 0.4σ, suggesting that relative to the alternative, a year spent in school for disadvantaged children would prevent the test-score gap from widening with entrance age. If there are strong complementarities between earlier and later learning, as Cunha et al. (2006) argue, the lost year of schooling could have negative long-run consequences for disadvantaged children, while the additional human capital investments received by other older school entrants could prevent them from falling behind, if not better position them to succeed. If this were true, it would also not be surprising that the effect of having older peers on average is significantly positive for disadvantaged children, but negligible otherwise (Panel B of Table 9).

VII. Conclusion

In this paper, we have estimated the effects of relative and absolute age on achievement using data from an experiment where children of the same biological age were randomly assigned to different classrooms at the start of school. We find that children who are young relative to their kindergarten or first-grade classmates generally perform no worse on achievement tests and are no less likely to take the ACT or SAT. We arrive at a similar conclusion when we use a child's kindergarten schoolmates as a peer group, and when we measure a child's place in the classroom or school age distribution in a number of different ways. However, being relatively young may have a lasting impact for disadvantaged children. The reasons for this finding are unclear, though it seems reasonable to believe that disadvantaged children do not receive later investments that compensate for the negative shock of being relatively young.

On the basis of the existing reduced-form literature – the findings of which we have replicated in this paper – one might conclude that biological school entrance age no longer has a meaningful impact on educational attainment. Our findings suggest that this conclusion is misguided. While there is no significant long-run effect of school-entry age for the average child, a disadvantaged child who enters school at a later age stands a high probability of not taking a college-entrance exam, particularly if he does not have the good fortune of being assigned to a kindergarten class where he is relatively old. For other children, the probability of taking the ACT or SAT does not fall by as much and may in fact rise with age at the start of kindergarten. These findings suggest that efforts to provide disadvantaged children with higher-quality care and education prior to kindergarten, as well as changes to state and local rules governing the age of the youngest kindergartner, could substantially affect socioeconomic gaps in educational attainment.

Nevertheless, this study has not directly analyzed any policy intervention. Rather, we have estimated the effects of absolute age and relative age in a cross-section where kindergarten entry regulations and educational opportunities for disadvantaged pre-kindergartners have been held constant. More generally, the age profile of achievement is likely to be strongly related to investments outside of school, which may differ across samples. The strength of our estimated relative-age effects may also rest on features of Tennessee public schools that are not more widely present or on the fact that we are only able to measure relative age when children are young. Attempting to replicate our findings in other settings should be a priority for future research.

Appendix

Consider the education production function

(A1) $y_{it} = \tilde{\alpha}_{0t} + \tilde{\alpha}_{1t} a_i + \tilde{\alpha}_{2t} a_{it} + \tilde{\alpha}_{3t} ra_i + u_{it}$

where a_{it} represents individual i's age at time t, and ra_i represents his relative age. Suppose that $ra_i = a_i + c$, where c represents some normalization (e.g., c is the average age of children in i's entry cohort). Given that $a_{it} = a_i + t$, Model (1) is a simplified version of Model (A1), where $\alpha_{1t} \equiv \tilde{\alpha}_{1t} + \tilde{\alpha}_{2t} + \tilde{\alpha}_{3t}$ and $\alpha_{0t} \equiv \tilde{\alpha}_{0t} + \tilde{\alpha}_{2t} t - \tilde{\alpha}_{3t} c$. The true effect of relative age might not be a linear function of a_i, as discussed above. However, because it is monotonic in a_i, ra_i will still be highly correlated with entry age.

The same identification problem arises when working with data selected on grade, not academic or birth cohort, as is the case in several studies (see for example Bedard and Dhuey (2006) and Elder and Lubotsky (2007)). For example, consider the model

(A2) $y_{ig} = \tilde{\delta}_{0g} + \tilde{\delta}_{1g} a_i + \tilde{\delta}_{2g} a_{ig} + \tilde{\delta}_{3g} ra_i + \tilde{u}_{ig}$,

where g represents grade and all other variables are as previously defined. In this case, $a_{ig} = a_i + g + r_{ig}$, where r_{ig} represents the number of years child i has been retained by grade g. Assuming that $ra_i = a_i + c$, this model can be re-written as $y_{ig} = \delta_{0g} + \delta_{2g} a_{ig} + u_{ig}$, where $\delta_{2g} \equiv \tilde{\delta}_{1g} + \tilde{\delta}_{2g} + \tilde{\delta}_{3g}$, $\delta_{0g} \equiv \tilde{\delta}_{0g} - \left(\tilde{\delta}_{1g} + \tilde{\delta}_{3g}\right) g - \tilde{\delta}_{3g} c$, and $u_{ig} \equiv \tilde{u}_{ig} + \left(\tilde{\delta}_{1g} + \tilde{\delta}_{3g}\right) r_{ig}$.

References

Angrist, Joshua D. and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Education and Earnings?" *The Quarterly Journal of Economics* 106(4): 979-1014.

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

Bedard, Kelly and Elizabeth Dhuey. 2007. "Is September Better Than January? The Effect of Minimum Age at School Entry Laws on Adult Earnings." Mimeo, University of California Santa Barbara.

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

Bound, John, and David A. Jaeger. 2000. "Do Compulsory Attendance Laws Alone Explain the Association Between Earnings and Quarter of Birth?" In *Research in Labor Economics: Worker Well-Being,* ed. Solomon W. Polacheck, 83-108. New York: JAI.

Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90(430):443-450.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2006. "Robust Inference with Multi-Way Clustering." NBER Technical Working Paper 327. Cambridge, MA: National Bureau of Economic Research.

Cascio, Elizabeth U. 2007. "Do Large Investments in Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools." Mimeo, Dartmouth College.

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

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation," in Eric A. Hanushek and Finish Welch, eds., *Handbook of the Economics of Education, Volume 1*, pp. 698-812.

Datar, Ashlesha. 2006. "Does Delaying Kindergarten Entrance Give Children a Head Start?" *Economics of Education Review* 25: 43-62.

Decicca, Philip. 2007. "Does Full-Day Kindergarten Matter? Evidence from the First Two Years of Schooling." *Economics of Education Review* 26(1): 67-82.

Dee, Thomas S. 2004. "Teachers, Race, and Student Achievement in a Randomized Experiment." *The Review of Economics and Statistics* 86, 195-210.

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

Dobkin, Carlos and Fernando Ferreira. 2007. "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" Mimeo.

Elder, Todd E. and Darren H. Lubotsky. 2007. "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers." Mimeo.

Fredricksson, Peter and Björn Öckert. 2006. "Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance." IFAU Working Paper 2006:12. Uppsala: Institute for Labour Market Policy Evaluation.

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

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

Graham, Bryan S. 2007. "Identifying Social Interactions through Conditional Variance Restrictions," Mimeo, University of California Berkeley.

Imbens, Guido and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2): 467-475.

Kawaguchi, Daiji. 2006. "The Effect of Age at School Entry on Education and Income." ESRI Discussion Paper Series No. 162. Tokyo: Economic and Social Research Institute.

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

Krueger, Alan B. and Diane M. Whitmore. 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.

Krueger, Alan B. and Diane M. Whitmore. 2002. "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, eds. J. Chubb and T. Loveless, 11-46. Washington, D.C.: Brookings Institution Press.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2007. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in High and Low Achievers in the Classroom." Mimeo.

Lazear, Edward. 2001. "Educational Production" *The Quarterly Journal of Economics* 116(3): 777-803.

Lleras-Muney, Adriana. 2005. "The Relationship Between Education and Adult Mortality in the United States." *The Review of Economic Studies* 72(1): 189-221.

Lochner, Lance and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self Reports." *The American Economic Review* 94(1): 155-189.

Magnuson, Katherine A. and Jane Waldfogel. 2005. "Early Childhood Care and Education: Effects on Ethnic and Racial Gaps in School Readiness." *The Future of Children* 15(1): 169-196.

McEwan, Patrick and Joseph Shapiro. 2006. "The Consequences of Delayed Primary School Enrollment in a Developing Country." Forthcoming, *The Journal of Human Resources*.

Neal, Derek A. 2006. "Why Has Black-White Skill Convergence Stopped?" *Handbook of Economics of Education,* edited by Eric Hanushek and Finis Welch, Elsiver.

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens. 2006. "Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics* 24(4): 729-760.

Oreopoulos, Philip. 2007. "Would More Compulsory Schooling Help Disadvantaged Youth? Evidence from Recent Changes to School-Leaving Laws." Forthcoming in *An Economic Framework for Understanding and Assisting Disadvantaged Youth*, ed., Jonathan Gruber.

Puhani, Patrick A. and Andrea M. Weber. 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. Bonn: Institute for the Study of Labor.

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

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

Sowell, Elizabeth R., Bradley S. Peterson, Paul M. Thompson, Suzanne E. Welcome, Amy L. Henkenius, and Arthur W. Toga. 2003. "Mapping Cortical Change Across the Human Life Span." *Nature Neuroscience* 6: 309-315.

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

West, Jerry, Kristen Denton, and Elvira Germino-Hausken. 2000. *America's Kindergartners*. Washington, D.C.: U.S. Department of Education.

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

Table 1 - Descriptive Statistics for Project STAR Kindergarten Cohort

	Mean (1)	SD (2)	N (3)
Age Variables:			
Expected Age in K	5.38	0.28	6248
Age in K	5.43	0.35	6248
Average Expected Age of K Classmates	5.38	0.07	6248
Average Age of K Classmates	5.43	0.09	6248
Expected Age Rank in K Classroom	0.50	0.30	6248
Age Rank in K Classroom	0.50	0.30	6248
Demographic/SES Variables:			
Female	0.49	-	6248
Black	0.33	-	6248
Free/Reduced-Price Lunch (in K)	0.49	-	6248
K Classmate Characteristics:			
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
K Teacher Characteristics:			
Black	0.17	-	6248
Has MA	0.35	-	6248
Has <2 Years Experience	0.10	-	6248
K Class Characteristics:			
Small	0.30	-	6248
Regular with Aide	0.35	-	6248
Outcome Variables:			
Math/Reading Average z-score, Spring 1986 (K)	0.00	1.00	5719
Math/Reading Average z-score, Spring 1994	0.12	0.99	4436
Below grade, Spring 1994	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 2 - Exogeneity of the Instrumental Variables

	Dependent Variable:				
	ea_i	$ea_{-i,k}$	$rank_k(ea_i)$	$1(rank_k(ea_i)<0.25)$	$1(ea_i<ea_{-i,k}-0.25)$
	(1)	(2)	(3)	(4)	(5)
Expected Age in K (ea_i)			1.045	-1.112	-1.097
			(0.006)	(0.009)	(0.012)
Average ea_i of K Classmates ($ea_{-i,k}$)			-1.044	1.094	1.121
			(0.007)	(0.018)	(0.056)
Covariates:					
Demographic/SES Variables	X	X	X	X	X
K Classmate Characteristics	X	X	X	X	X
K Teacher Characteristics	X	X	X	X	X
K Class Characteristics	X	X	X	X	X
P-value on covariates (not inc. instr.)	0.07	0.16	0.88	0.21	0.12

Notes: Each column represents a different regression. Covariates are listed in Table 1. Each regression includes school fixed effects and is based on 6248 observations. Standard errors (in parentheses) are clustered on K classroom.

Table 3 - Estimates of the Effect of Absolute Age, Relative Age in Kindergarten Classroom, and Age of Kindergarten Classmates on Test Scores at the End of Kindergarten

	OLS	RF†	TSLS	TSLS
	(1)	(2)	(3)	(4)
		A. Age Only		
a_i	0.242	0.574	0.707	0.674
	(0.038)	(0.044)	(0.057)	(0.054)
		B. Relative Age=Rank		
a_i	-0.538	0.649	0.841	0.819
	(0.097)	(0.183)	(0.204)	(0.194)
$rank_k(a_i)$	0.973	-0.063	-0.126	-0.138
	(0.111)	(0.172)	(0.197)	(0.188)
$a_{-i,k}$	0.958	0.706	0.72	0.585
	(0.283)	(0.389)	(0.469)	(0.428)
P-value: a_i=$rank_k(a_i)$	0.00	0.04	0.02	0.01
P-value: Relative Age=0	0.00	0.71	0.52	0.46
P-value: a_i=Relative Age=0	0.00	0.00	0.00	0.00
		C. Relative Age=Dummy for Bottom 25%		
a_i	0.050	0.568	0.698	0.681
	(0.049)	(0.060)	(0.081)	(0.078)
$1(rank_k(a_i)<0.25)$	-0.236	-0.013	-0.016	-0.002
	(0.034)	(0.036)	(0.048)	(0.047)
$a_{-i,k}$	0.361	0.787	0.866	0.726
	(0.267)	(0.330)	(0.410)	(0.376)
P-value: Relative Age=0	0.00	0.72	0.74	0.97
P-value: a_i=Relative Age=0	0.00	0.00	0.00	0.00
		D. Relative Age=Dummy for 3 Months Below $a_{-i,k}$		
a_i	0.030	0.571	0.683	0.680
	(0.046)	(0.062)	(0.097)	(0.092)
$1(a_i<a_{-i,k}-0.25)$	-0.246	-0.011	-0.026	-0.002
	(0.034)	(0.038)	(0.061)	(0.059)
$a_{-i,k}$	0.419	0.784	0.874	0.726
	(0.269)	(0.332)	(0.412)	(0.378)
P-value: Relative Age=0	0.00	0.77	0.67	0.97
P-value: a_i=Relative Age=0	0.00	0.00	0.00	0.00
Additional Controls?	N	N	N	Y

Notes: Each column and panel of the table presents coefficients and standard errors 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. All regressions include school fixed effects and are estimated based on 5719 observations. Additional controls include 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. Standard errors (in parentheses) are clustered on K class. †Coefficients shown in this column are on the instrumental variable analogs of the variables shown.

Table 4 - TSLS Estimates of the Effect of Absolute Age, Relative Age in Kindergarten Classroom, and Age of Kindergarten Classmates on Other Outcomes

	Dependent Variable:		
	Math and Reading Average, 1994	Below Grade, 1994	Took ACT/SAT
	(1)	(2)	(3)
		A. Age Only	
a_i	0.215	-0.189	0.002
	(0.065)	(0.026)	(0.024)
		B. Relative Age=Rank	
a_i	0.142	-0.104	-0.019
	(0.228)	(0.105)	(0.095)
$rank_k(a_i)$	0.079	-0.087	0.023
	(0.224)	(0.105)	(0.095)
$a_{-i,k}$	0.508	-0.223	0.194
	(0.352)	(0.159)	(0.142)
P-value: a_i=$rank_k(a_i)$	0.89	0.93	0.82
P-value: Relative Age=0	0.72	0.41	0.81
P-value: a_i=Relative Age=0	0.00	0.00	0.96
		C. Relative Age=Dummy for Bottom 25%	
a_i	0.188	-0.095	-0.022
	(0.087)	(0.036)	(0.039)
$1(rank_k(a_i)<0.25)$	-0.028	0.080	-0.022
	(0.056)	(0.026)	(0.027)
$a_{-i,k}$	0.461	-0.226	0.197
	(0.255)	(0.117)	(0.111)
P-value: Relative Age=0	0.62	0.00	0.42
P-value: a_i=Relative Age=0	0.00	0.00	0.71
		D. Relative Age=Dummy for 3 Months Below $a_{-i,k}$	
a_i	0.209	-0.086	0.002
	(0.100)	(0.045)	(0.044)
$1(a_i<a_{-i,k}-0.25)$	-0.009	0.083	-0.002
	(0.063)	(0.032)	(0.030)
$a_{-i,k}$	0.439	-0.214	0.173
	(0.252)	(0.115)	(0.110)
P-value: Relative Age=0	0.89	0.01	0.95
P-value: a_i=Relative Age=0	0.00	0.00	0.99
N	4436	4508	6248

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. All models are estimated using TSLS and include school fixed effects and the full vector of controls listed in the notes to Table 3. Standard errors (in parentheses) are clustered on K class.

Table 5 - TSLS Estimates for Relative Age and Average Age of Peers Within First Grade Classroom or School

	Dependent Variable:			
	Math and Reading Average †	Math and Reading Average 1994	Below Grade, 1994	Took ACT/SAT
	(1)	(2)	(3)	(4)
	Relative Age Defined in 1st Grade Classroom			
	A. Age Only			
a_i	0.485	0.092	-0.148	-0.012
	(0.062)	(0.068)	(0.033)	(0.035)
	B. Relative Age = Dummy for Bottom 25%			
a_i	0.460	0.058	0.054	-0.087
	(0.113)	(0.140)	(0.063)	(0.062)
$1(rank_1(a_i)<0.25)$	-0.021	-0.027	0.159	-0.06
	(0.083)	(0.105)	(0.048)	(0.045)
$a_{-i,1}$	0.174	0.363	-0.17	-0.026
	(0.417)	(0.295)	(0.122)	(0.143)
p: Relative Age=0	0.80	0.80	0.00	0.18
p: a_i=Relative Age=0	0.00	0.41	0.00	0.36
N	4782	3728	3797	5089
	Relative Age Defined at School Level (within Kindergarten Cohort)			
	C. Age Only			
a_i	0.705	0.223	-0.190	0.005
	(0.057)	(0.069)	(0.026)	(0.026)
	D. Relative Age = Dummy for Bottom 25%			
a_i	0.646	0.131	-0.069	-0.048
	(0.087)	(0.090)	(0.038)	(0.038)
$1(rank_s(a_i)<0.25)$	-0.044	-0.075	0.098	-0.043
	(0.050)	(0.055)	(0.027)	(0.025)
$a_{-i,s}$	2.368	0.653	-0.406	0.262
	(1.008)	(0.547)	(0.257)	(0.224)
p: Relative Age=0	0.38	0.18	0.00	0.09
p: a_i=Relative Age=0	0.00	0.00	0.00	0.21
N	5719	4436	4508	6248

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. Models in Panels A and B inc lude school fixed effects and the full vector of additional controls lists in the notes to Table 3, but evaluated for 1st grade classrooms. Models in Panels C and D exclude school fixed effects, but include the full vector of controls listed in the notes to Table 3, as well as the fraction of a child's kindergarten schoolmates who are female, black, or received free or reduced-price lunch. Standard errors (in parentheses) are clustered on first grade classroom in Panels and B and school in Panels C and D. † Test scores in this column are for 1st grade in Panels A and B and kindergarten in Panels C and D.

Table 6 - Robustness of TSLS Estimates to Non-Random Transitions Across Kindergarten Classrooms

	Dependent Variable:			
	Math and Reading Average K	Math and Reading Average 1994	Below Grade, 1994	Took ACT/SAT
	(1)	(2)	(3)	(4)
	Subsample in Unique Kindergarten Class of Type within School			
	A. Age Only			
a_i	0.749	0.221	-0.201	0.015
	(0.072)	(0.088)	(0.036)	(0.033)
	B. Relative Age = Dummy for Bottom 25%			
a_i	0.770	0.259	-0.166	-0.021
	(0.106)	(0.118)	(0.049)	(0.053)
$1(rank_k(a_i)<0.25)$	-0.007	0.03	0.032	-0.034
	(0.065)	(0.082)	(0.036)	(0.036)
$a_{-i,k}$	1.293	0.06	-0.194	0.242
	(0.471)	(0.325)	(0.169)	(0.154)
p: Relative Age=0	0.91	0.71	0.38	0.35
p: a_i=Relative Age=0	0.00	0.03	0.00	0.55
N	2835	2204	2235	3088
	Subsample Not in Small Kindergarten Classes			
	C. Age Only			
a_i	0.652	0.201	-0.161	-0.002
	(0.059)	(0.078)	(0.031)	(0.030)
	D. Relative Age = Dummy for Bottom 25%			
a_i	0.643	0.126	-0.072	-0.007
	(0.085)	(0.098)	(0.041)	(0.046)
$1(rank_k(a_i)<0.25)$	-0.015	-0.064	0.072	-0.007
	(0.054)	(0.065)	(0.030)	(0.032)
$a_{-i,k}$	0.451	0.159	0.014	0.189
	(0.423)	(0.321)	(0.126)	(0.152)
p: Relative Age=0	0.78	0.33	0.02	0.83
p: a_i=Relative Age=0	0.00	0.03	0.00	0.98
N	3989	3084	3138	4359

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. See text for a complete description of subsamples. All models include school fixed effects and the full vector of controls listed in the notes to Table 3. Standard errors (in parentheses) are clustered on K class.

Table 7 - Robustness of TSLS Estimates to Direct Effect of Birthday on Outcomes

	Dependent Variable:			
	Math and Reading Average K	Math and Reading Average 1994	Below Grade, 1994	Took ACT/SAT
	(1)	(2)	(3)	(4)
	A. Regression Discontinuity Estimates, Age Only			
a_i	0.648	0.182	-0.153	0.000
	(0.139)	(0.162)	(0.076)	(0.072)
p: birthday	0.90	0.92	0.08	0.17
	B. Regression Discontinuity Estimates, Relative Age = Dummy for Bottom 25%			
a_i	0.649	0.073	-0.049	0.003
	(0.176)	(0.198)	(0.092)	(0.091)
$1(\text{rank}_k(ai)<0.25)$	-0.015	-0.095	0.084	-0.001
	(0.076)	(0.088)	(0.042)	(0.040)
$a_{-i,k}$	0.762	0.59	-0.259	0.193
	(0.400)	(0.287)	(0.125)	(0.122)
p: Relative Age=0	0.84	0.28	0.05	0.98
p: ai=Relative Age=0	0.00	0.26	0.02	1.00
p: birthday	0.92	0.85	0.90	0.32
N	5719	4436	4508	6248

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. Models are estimated on the full sample. All models include a quadratic in normalized birthday, entered directly and interacted with an indicator for having a birthday in the six months following September 30. Models use this indicator as an instrument instead of expected age at school entry; other instruments remain unchanged. (See text for more details.) All models also include school fixed effects and the full vector of controls listed in the notes to Table 3. Standard errors (in parentheses) are clustered on K class.

Table 8 - TSLS Estimates of the Effect of Absolute Age, Relative Age in Kindergarten Classroom, and Age of Kindergarten Classmates, by Gender

	Dependent Variable:							
	Read/Math Average K		Read/Math Average 1994		Below Grade		Took SAT/ACT	
	Males	Females	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Age Only				
a_i	0.635	0.700	0.341 ^	0.115 ^	-0.277 ^	-0.114 ^	0.059 ^	-0.051 ^
	(0.071)	(0.076)	(0.102)	(0.075)	(0.042)	(0.034)	(0.038)	(0.033)
				B. Relative Age=Dummy for Bottom 25%				
a_i	0.715	0.625	0.279	0.109	-0.126	-0.056	0.013	-0.056
	(0.114)	(0.111)	(0.140)	(0.113)	(0.060)	(0.048)	(0.057)	(0.055)
$1(rank_k(a_i)<0.25)$	0.058	-0.071	-0.054	-0.011	0.126	0.051	-0.039	-0.007
	(0.071)	(0.064)	(0.094)	(0.070)	(0.044)	(0.032)	(0.037)	(0.038)
$a_{-i,k}$	0.563	0.915	0.236	0.627	-0.264	-0.13	0.04	0.34
	(0.468)	(0.411)	(0.380)	(0.315)	(0.164)	(0.156)	(0.145)	(0.169)
P-value: Relative Age=0	0.41	0.27	0.57	0.88	0.00	0.11	0.29	0.85
P-value: a_i=Relative Age=0	0.00	0.00	0.00	0.26	0.00	0.00	0.18	0.33
N	2936	2783	2191	2245	2234	2274	3213	3035

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. All models are estimated using TSLS and include school fixed effects and the full vector of controls listed in the notes to Table 3. Standard errors (in parentheses) are clustered on K class. ^ designates that the male-female difference in coefficients is significant at the 10 percent level.

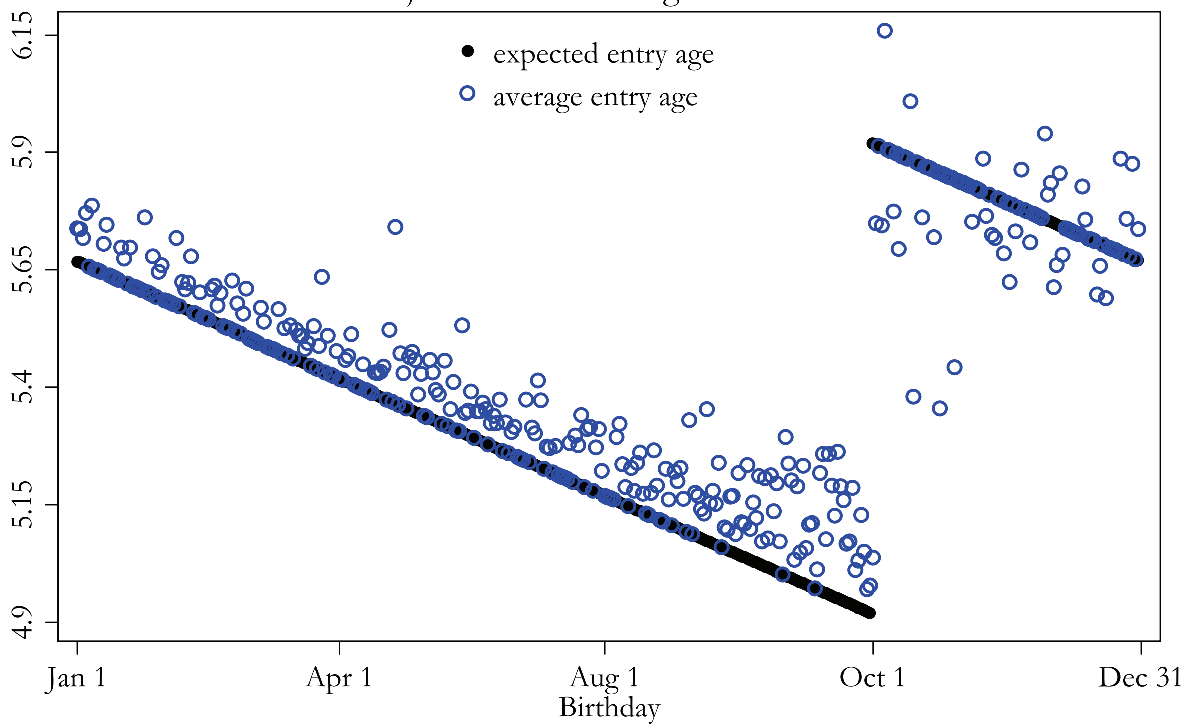
Table 9 - TSLS Estimates of the Effect of Absolute Age, Relative Age in Kindergarten Classroom, and Age of Kindergarten Classmates, by Free Lunch Status

	Dependent Variable:							
	Read/Math Average K		Read/Math Average 1994		Below Grade		Took SAT/ACT	
	Lunch	No Lunch	Lunch	No Lunch	Lunch	No Lunch	Lunch	No Lunch
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Age Only								
a_i	0.587 ^	0.767 ^	0.178	0.229	-0.186	-0.188	-0.005	0.029
	(0.072)	(0.081)	(0.095)	(0.086)	(0.044)	(0.034)	(0.038)	(0.036)
B. Relative Age=Dummy for Bottom 25%								
a_i	0.492 ^	0.858 ^	0.228	0.103	-0.136	-0.052	-0.097 ^	0.066 ^
	(0.114)	(0.108)	(0.133)	(0.115)	(0.065)	(0.047)	(0.054)	(0.061)
$1(rank_k(a_i)<0.25)$	-0.095	0.069	0.037	-0.104	0.049	0.110	-0.084 ^	0.030 ^
	(0.068)	(0.064)	(0.086)	(0.074)	(0.044)	(0.033)	(0.036)	(0.039)
$a_{-i,k}$	1.020 ^	0.482 ^	0.386	0.375	-0.377	-0.113	0.443 ^	-0.038 ^
	(0.499)	(0.423)	(0.472)	(0.307)	(0.207)	(0.134)	(0.155)	(0.160)
P-value: Relative Age=0	0.16	0.28	0.67	0.16	0.27	0.00	0.02	0.44
P-value: a_i=Relative Age=0	0.00	0.00	0.13	0.02	0.00	0.00	0.06	0.55
N	2768	2951	2038	2398	2088	2420	3031	3217

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. All models are estimated using TSLS and include school fixed effects and the full vector of controls listed in the notes to Table 3. Standard errors (in parentheses) are clustered on K class. ^ designates that the lunch-no lunch difference in coefficients is significant at the 10 percent level.

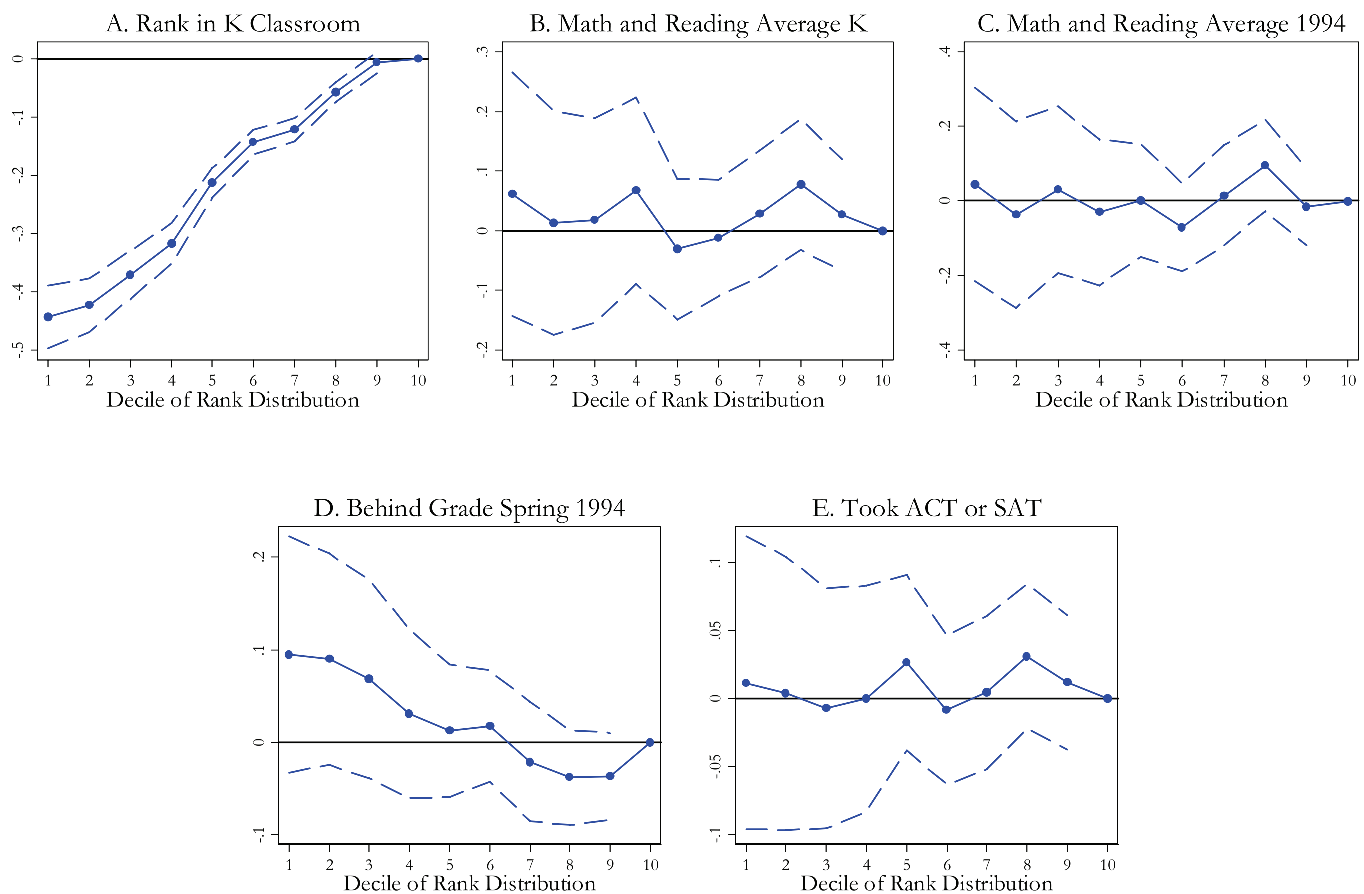
Figure 1. Expected and Average Kindergarten Entry Age by Birthday

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)

Figure 2. Sensitivity of Estimates to Parameterization of Relative Age Effect



Notes: Figures plot coefficients on indicators for deciles of the expected age rank distribution in kindergarten classroom (dots and solid lines) and 95 percent confidence intervals on these coefficients (dashed lines). The indicator for the oldest decile is omitted from the model. Standard errors are robust to heteroskedasticity and correlation of error terms within kindergarten classrooms. Underlying regressions include expected age at kindergarten entry, the average expected age of kindergarten classmates, school fixed effects and the vector of additional controls listed in the notes to Table 3.

Appendix Table A1 - First Stage Regressions for Table 3

	Dependent Variable:				
	a_i	$rank_k(a_i)$	$a_{-i,k}$	$1(rank_k(a_i)<0.25)$	$1(a_i<a_{-i,k}-0.25)$
	(1)	(2)	(3)	(4)	(5)
			A. Age Only		
ea_i	0.812	-	-	-	-
	(0.015)				
F-statistic: instrument	2933				
			B. Relative Age=Rank		
ea_i	0.735	-0.209	0.006	-	-
	(0.058)	(0.048)	(0.005)		
$rank_k(ea_i)$	0.075	0.971	-0.005	-	-
	(0.053)	(0.044)	(0.004)		
$ea_{-i,k}$	0.138	0.216	0.849	-	-
	(0.079)	(0.051)	(0.049)		
F-statistic: instruments	1025	1488	739		
			C. Relative Age=Dummy for Bottom 25%		
ea_i	0.812	-	0.000	-0.122	-
	(0.019)		(0.003)	(0.023)	
$1(rank_k(ea_i)<0.25)$	-0.001	-	-0.001	0.759	-
	(0.012)		(0.001)	(0.019)	
$ea_{-i,k}$	0.061	-	0.855	0.079	-
	(0.054)		(0.050)	(0.027)	
F-statistic: instruments	1015		726	2132	
			D. Relative Age=Dummy for 3 Months Below $a_{-i,k}$		
ea_i	0.821	-	0.000	-	-0.336
	(0.019)		(0.004)		(0.034)
$1(ea_i<ea_{-i,k}-0.25)$	0.008	-	0.000	-	0.623
	(0.012)		(0.002)		(0.025)
$ea_{-i,k}$	0.051	-	0.855	-	0.065
	(0.053)		(0.050)		(0.060)
F-statistic: instruments	1013		730		1865

Notes: Each column and panel of the table presents coefficients and standard errors from a different regression. All models include school fixed effects and the vector of additional controls listed in the notes to Table 3. Estimates are based on 5719 observations. Standard errors (in parentheses) are clustered on K class.