

**The Effects of Relative Age on Early Childhood Academic Achievement:  
How They Differ Between Gender and Change Across Time\***

May 2012

Stephanie Yuechen Zhong

Department of Economics  
Stanford University  
Stanford, CA 94305  
yuechenz@stanford.edu

Under the direction of  
Professor Caroline M. Hoxby

**ABSTRACT**

Gary Becker argues that a person should make investments in human capital as early as possible because he will thereby maximize the period over which the investment pays out. Also, the pay out period will be less discounted, and a person's opportunity cost of time is lowest when he is young. However, this theory does not provide much guidance on what is the optimal age to enroll children in formal education. At that point, there are two concerns: (i) whether the child is learning more at home than he would at school and (ii) whether his age *relative to his peers* is such that he would learn more given his aptitude. Relative age has recently become a controversial issue because some parents now deliberately enrolled their children a year after they are eligible to start school, in the hope of giving them a permanent learning advantage. This is known as redshirting. There is little robust quantitative evidence on how relative age affects learning and whether the effects differ by gender or parents' education. I examine how relative age affects achievement using the nationally representative ECLS-K longitudinal survey. Because relative age could be endogenous to parents' observations of their child's maturity, I construct instrumental variables for relative age based on state-level policies of kindergarten entrance age cut-off dates. I find that the relative age has a positive effect from kindergarten through grade five in reading and from kindergarten through grade eight in math. I find little evidence of the permanent effect for which redshirting families hope. The relative age effect is greater for females and for children whose parents are more educated. The latter result suggests that children with more educated parents learn more at home than do other children.

*Key words:* human capital investment, relative age effects, kindergarten entrance cut-off age policy, redshirting practice, early childhood academic achievement

---

\* I am indebted to Professor Caroline Hoxby for her inspiring insights, engaging discussions and thoughtful guidance through every step of the research and writing process. I feel extremely grateful for her time and energy. I am thankful to Haibo Zhao, Patricia Foo, Sean Mandell, Alan Guo, and Rebecca Sachs for the encouragement and the valuable comments. I want to thank my friends and family for their continuous support.

## I. INTRODUCTION

Economists view education as an investment — an investment to augment your *human capital*. Human capital can be thought of as a person's internal factory — consisting of things, such as knowledge, which are acquired through education — that combined with that person's unskilled labor can generate earnings and economic value. In this sense, human capital is like physical capital, which is created by an investment and produces a return over a number of years. If parents rationally plan their children's education to *maximize the return* on their investment, they will consider several factors, such as the return to a given year (or grade) of education and the return associated with different schools of varying quality. They will also consider the age at which their child should first enroll in school. It is the age of first enrollment decision that I analyze in this study.

The seminal work of Gary Becker (1993), *Human Capital (3<sup>rd</sup> Edition)*, argues that a person should make investments in education as early as possible. That is because, first of all, “with finite lifetimes, later investments cannot produce returns for as long as earlier ones, therefore, usually have smaller total benefits” (Becker, p. 114). In the second place, “later investments are less profitable than earlier ones because the present value of net benefits (or profits) is reduced merely by postponing them (and the reduction can be sizable, even for postponements of a few years)” (Becker, p. 114). A third consideration is “probably of greater importance,” as Becker suggested, which is that at younger ages, “the value of the time is small and probably even negative

because parents or other baby-sitting services must be employed if he is not in school (Becker, p. 114)". This translates into an increasing marginal cost of later investments compared with earlier ones, since "the former use more expensive time" (Becker, p.114).

It is fairly straightforward to apply Becker's argument in the case of — say — college education. By his logic, it would make little sense to take several years off between high school and working towards a baccalaureate degree unless there was an extenuating circumstance or a better non-school-based way of making human capital investments at that time. However, it is somewhat less straightforward to apply Becker's argument to initial enrollment in schooling. Presumably, we do not enroll newborn infants in formal schools because the human capital they most need to acquire (such as basic language) is best acquired through contact with their families. But, at some age, the human capital that a school instills becomes a necessary complement to what the family instills, if the child is to continue on the *optimal human capital investment path*. This age may be different for children from families with more or less well-educated parents. On the one hand, a well-educated family may have a greater ability to substitute for formal education. Well-educated parents may enhance a four, five, or six year-old's cognitive development more than formal schools. On the other hand, the opportunity costs of keeping a child at home — rather than in formal school — rise with parents' education. Therefore, well-educated parents may maximize their extended family's wealth by sending their child to school earlier, rather than later, than less-educated parents. Moreover, the age at which a

child should optimally begin formal schooling may vary from child to child, even among children with similar parents. Some children may mature faster than others and be ready for school earlier. Children may also differ in the extent to which they are helped or hurt by interacting with older children. In any case, economic theory does not give us strong guidance about the age at which students should initially enroll in school and whether this age varies by parents' education.

In practice, it is not immediately obvious what parents think is the optimal age at which to enroll their children in formal schooling. Some parents try to enroll their children before the usual starting age. Others deliberately keep their child in pre-school or at home for a year (or rarely, more) after he is eligible to attend school. This latter practice is known as “redshirting” in an analogy to college athletics, and it has become so common that it has received significant media attention and controversy (Gootman, 2006). Families that “redshirt” hope that their child will perform more highly if he is relatively older for his grade and among his classmates. They hope that this increase in performance will be so great that it more than makes up for the fact that their child is acquiring his education later, with the corresponding consequences (fewer and more highly discounted years of earnings on the human capital, and more years of costly child care). In other words, they hope their children can enjoy a consistently higher *rate of return* on their education, one that more than compensates for the extra year of waiting.

Parents often make the redshirting decision in the absence of robust quantitative evidence on whether it makes sense from the human capital investment point of view. In this paper, I attempt to fill the gap in evidence by answering four questions. First, do the data support a *causal* relationship between children's relative age and their initial performance in school? Second, if so, does this effect persist sufficiently into later grades so that a person would likely be better off as an adult if he were redshirted? Third, does the effect differ for children of different genders? Fourth, does the effect differ for children whose parents have low versus high educational attainment?

While the answer to the second question determines whether or not “redshirting” is a sound investment decision for parents, the answers to the third and fourth questions have immediate implications for policy makers. Up until now, most countries have set a minimum age for school enrollment without regard to gender or parents' background. For instance, in the United States, children in most states can enroll in kindergarten in the school-year associated with their 5th birthday regardless of gender (Saluja, Scott-Little and Clifford, 2000). California, for example, requires kindergartens to admit children at the beginning of the 2011-2012 school year, if they will be five years of age on or before December 2 during that academic year. However, this is not the same everywhere, and the specific entrance *cut-off date* associated with this 5th birthday rule differs significantly across the country: five states have cut-off dates on or before August 15, thirty-five states have cut-off dates between August 31 and October 16, four states (including California) and the District

of Columbia have kindergarten entrance cut-off dates between November 1 and January 1, and a few other states do not even have a state entrance age policy and instead they leave the decision to the local education agencies (LEA) (Gordon, 2011).

Not only do states have different entrance cut-off dates, but many states have also been revising their cut-off dates and striving to optimize their entrance age policy. If my results suggest that children of different genders or backgrounds are affected differently by their school starting age, governments might well be advised to make gender- or background-specific minimum age recommendations in order to boost overall academic performance. If, for instance, my study shows that boys are subject to a larger relative age effect (younger boys tend to underperform their peers by a larger margin than similarly younger girls), that would suggest that boys should be more carefully screened for maturity when they attempt to enroll in kindergarten. Another policy that might be considered is having narrower age ranges for classrooms. One kindergarten class might contain students aged 5.0 to 5.5 years at the beginning of the school year. The other kindergarten entry class might contain students aged more than 5.5 years.

Using the Early Childhood Longitudinal Study-Kindergarten cohort, a nationally representative and very thorough survey of children were entering kindergarten in the fall of 1998, I attempt to answer the aforementioned four questions. The main challenge that I face is that a child's actual relative age is potentially endogenous to his rate of maturation or other unobserved variables that

may affect his achievement. As a result, a naïve regression of a child's achievement on his relative age might not generate a coefficient that reflects the *causal* effect of relative age. To remedy this problem, I construct an instrumental variable: the child's *predicted relative entrance age*, which is derived from the interaction between the child's birthday and cut-off date for school enrollment in the child's state. This instrumental variable is highly correlated with actual relative age, but it is very plausibly uncorrelated with idiosyncratic, unobserved determinants of a child's maturity and aptitude. Therefore, my instrumental variable estimates credibly identify the causal effect of a child's relative age on his achievement.

This paper proceeds as follows. Section II presents relevant literature. I review different findings from the existing literature, describing the limits of the methodologies that have been used and how my approach can overcome these limitations. Section III describes the Early Childhood Longitudinal Study-Kindergarten (ECLS-K) cohort data set. In Section IV, I present my empirical strategies, and in Section V, I present the results from a rich array of instrumental variable regressions. Section VI finally concludes this study.

## II. LITERATURE REVIEW

There is an abundant amount of discussion, ranging from popular literature to academic research, about the effects of relative age on educational achievement and at what age children should first enroll in formal education. *Kindergarten: It Isn't What*

*It Used to Be*, an informative guide written by Susan Golant and Mitch Golant (1999) for parents with young children, stresses the importance of the timing for first entering kindergarten and suggests numerous ways to ensure child's readiness for school.

When is the right time to start school, and what are the trade-offs of being older or younger than the majority of children in a grade? *Outliers: The Story of Success*, written by Malcolm Gladwell (2008), discusses the benefit of being relatively older than peers and attributes the success of adult Canadian hockey players to their initial advantage of being the relative older ones in the childhood hockey leagues. The book *Outliers* states that 40% of the professional hockey players were born between January and March, and 30% between April and June, 20% between July and September and only 10% between October and December. Given that the age cut-off for youth hockey in Canada is January 1 and hockey leagues start around age 6, the argument is that the success of adult hockey players is largely the result of their initial relatively age advantage. The children who initially get selected for the best teams are the relatively older ones, and a half year age difference means a lot mentally and physically at a young age. The best teams have the best practices and coaches, and play better competition, thus making those children relatively older in the age 6 league more likely to get selected into the age 7 league as well. Hence this initial relative age effect builds up and persists into adulthood. Naturally, this popular book by Gladwell inspired many people, not only parents but also policy makers, to wonder if this phenomenon of relative age would hold not only in the sports world but in the academic field as well.



There are many empirical studies on the effects of relative age on children's academic achievement. Most studies report older children do better in early primary school (Cameron & Wilson, 1990; Bickel, Zigmond & Strayhorn, 1991; Stipek, 2002). Bickel, Zigmond and Strayhorn's study (1991), for example, surveyed 222 children in the same grade in Pittsburgh, treating their age as continuous variable over 12 months, and found that older students do tend to have higher math scores at the first grade level. However, a few studies, such as Dietz and Wilson's study (1985), found no relative age effect in some or all achievement tests, even in the early grade levels. Dietz and Wilson surveyed 177 students in the same grade from the Delaware school district, grouped them into three intervals based on age at school entry, and found no significant age effects on either K readiness scores or ITBS achievement scores. One limitation of the above findings is their small sample size. This is a problem I aim to avoid by using a national representative ECLS-K dataset, which surveyed more than 20,000 children entering kindergarten in the year 1998-1999 across the country from different geographic locations, types of schools, and family backgrounds. Another issue with above-mentioned methodologies is how to interpret their findings. It is likely that unobserved qualities that led parents to decide to delay their children's enrollment into kindergarten make these children older than peers of the same grade, also contribute to differences in academic achievement found later between these children and children who began school "on time." For instance, if some of the older children were older because their parents had recognized that they were developmentally slow at the usual kindergarten age, they would perform worse

all else equal. Thus, an omitted variable with negative consequences would offset a potentially positive *causal* effect of delayed entry. To remedy this selection problem, I construct an instrument variable based on state-specific entrance age laws, which will be further discussed in the empirical strategy section of this paper.

If the relative age does affect academic achievement, how long this relationship lasts is another important question, one I also explore in this study. A number of researchers whose findings support the initial relative age effects also find that older children persistently outperform younger children in the same entry cohort well into their school careers. This has been found not only in United States studies (Datar, 2006), but also in studies conducted in other countries such as Sweden (Fredriksson and Öckert, 2006), Chile (McEwan and Shapiro, 2006), and Germany (Puhani and Weber, 2005), among others. This body of research that has led to the increasing practice of “redshirting”— sending a child to school one year later than the law stipulates, making the child older rather than younger than his or her peers in the same school class. Gootman (2006) commented that in some communities this “redshirting” practice has become prevalent enough to be considered “an epidemic.”

There are, however, many researchers who disagree. Some studies show that the relative-age effect becomes smaller as children grow older and disappears by the end of primary school. Elder and Lubotsky (2009), for example, argue that the *relative age effect* is a manifestation of skills that were acquired prior to school entry but does not affect the rate of learning once in school. According to this view, older and

younger children within a grade tend to learn at the same rate. They are on the same “slope,” and so age-related differences in achievement will fade away as the children progress through school. Cascio and Schanzenbach (2007), in another example, estimate the effects of relative age using data from Project STAR, an experiment where children of the same biological age were randomly assigned to different classrooms at the start of school. They find “no evidence that relative age matters for test scores or the likelihood of taking a college-entrance exams in the population at large.” Yet a concern regarding their results is that the variation in relative age in their study comes from slight age differences in the *other* children who are in a child’s classroom, not variation in the child’s own age. One might thus argue that the Project STAR-based study does not answer questions like whether parents ought to redshirt their children. In contrast, my empirical strategy does exploit variation driven by a child’s *own* age, so that I can address questions like redshirting.

In the discussion of relative age effects, gender has occasionally been a concern. Devereux and Black’s paper (2011), for example, mention that boys are slightly more likely to be “redshirted,” but it does not actually explore whether relative age affects boys differently than girls. The conventional wisdom, given the relatively early cognitive maturity of girls, seems to suggest that relatively younger boys might suffer the most from relative age effects. Whether this holds true is the third question I explore in this paper. In addition, returning to Becker’s theory of human capital investment, parent educational attainment is directly related to the level of human capital a child can accumulate at home if being held out of school. It should be

interesting to see how the relative age effects differ for children whose parents have different levels of education, an issue I examine with policy implications regarding inequality in educational resources.

### III. DATA

My data come from the Early Childhood Longitudinal Study—Kindergarten Class of 1998-99 (ECLS-K). It is a large, nationally representative, longitudinal survey that follows children from kindergarten through eighth grade.

The National Center for Education Statistics (NCES) began collecting the first wave of ECLS-K data in the fall of 1998, tracking a nationally representative sample of 21,260 children who entered kindergarten in the fall of the 1988-1999 academic year. From 1998 to 2007, seven total waves of data were collected from the same children, their parents, their school teachers, and their school administrators. Table 1 presents the waves of data collection, the years the data were collected, and the sample size. It is worth noting that in fall of 1999 (the shaded row of the below Table 1), only a 30 percent subsample of children were surveyed, so I leave out data from this wave of collection from the fall of 1999 for my study, keeping the other 6 waves of data.

**Table 1. ECLS-K Waves of Data Collection: Beginning with Kindergarten and Ending with Eighth Grade**

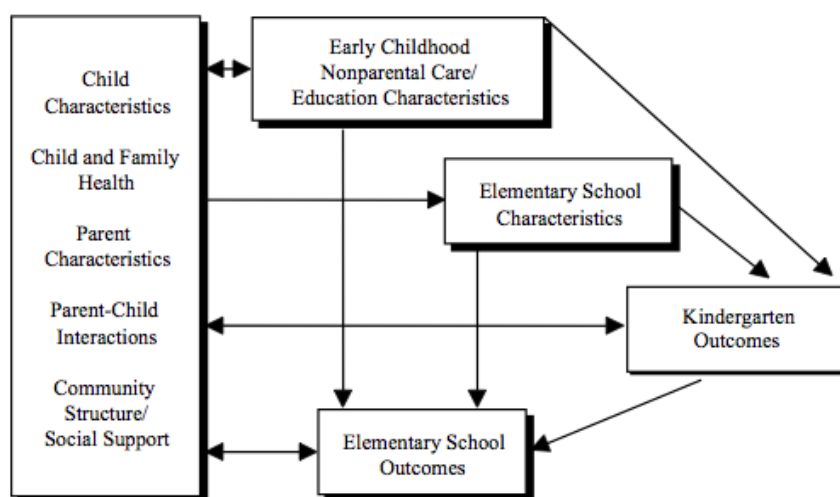
<b>Waves of Data Collection</b>	<b>Year Data Collected</b>	<b>Sample</b>
Kindergarten-fall	Fall 1998	Full Sample
Kindergarten-spring	Spring 1999	Full Sample

<u>First grade-fall</u>	<u>Fall 1999</u>	<u>30 percent subsample</u>
First grade	Spring 2000	Full Sample plus freshening
Third grade	Spring 2002	Full Sample
Fifth grade	Spring 2004	Full Sample
Eighth grade	Spring 2007	Full Sample

Note: Data from the third row, which is shaded in Table 1, is not used in this study, due to its partial sampling of the cohort.

Chart 1 shows the conceptual model of the ECLS-K survey. In addition to conducting academic assessment tests and interviews from each wave of data collection as noted in the Table 1, ECLS-K also collected descriptive information on children's personal information at school entry, their transition into school, children's characteristics, health conditions, parent characteristics, parent-child interactions, social support, as well as other information such as whether they have moved at some point during the years of data collection.

Chart 1. ECLS-K Conceptual Model



A very useful feature of the ECLS-K survey is that it is large in scale. I use a base sample of 13,551 observations on math scores collected in the kindergarten year

and 13,004 reading scores. Owing to its sampling plan, the ECLS-K has a smaller number of has a smaller number of students in each wave of data collection. However, each wave is designed to be nationally representative, as the base year (Kindergarten) wave was. I drop a child from my study if he or she did not take an assessment tests in *any* grade (as opposed to only one or two grades). In addition, I also drop students whose relative age is not reported (this is rare) or for whom I cannot compute the instrumental variable. For instance, I must drop children from states that did not have a state-level law regulating the kindergarten entrance cut-off date during the academic year of 1998-1999.

I use Item Response Theory (IRT) math and reading test scores from the ECLS-K to measure children's academic achievement in this study. Although the ECLS-K did occasionally test students in other subjects, math and reading are the only two subjects that were tested in every wave of the survey. The IRT method of test scoring takes into account the difficulty level of exams.

Regarding my instrument, I use not the *school-level* kindergarten entrance cut-off dates (reported by the school administrators in ECLS-K) but the state-level kindergarten cut-off laws. I do this because school-level cut-offs, especially those of private schools, are potentially correlated with the socioeconomic status of parents and the children's ability level.

The ECLS-K data set has a few limitations. First, the ECLS-K academic assessment tests take place at most twice a year. However, during the crucial early

grades, it would be interesting to track children's academic performance more regularly, for example on a monthly basis.

The ECLS-K data set does not survey the entire population of any school. Usually, the chosen school has only a handful of children who are surveyed, and this means that we cannot do extensive analysis of classroom peers' effects on one another.

An alternative dataset that I considered is the National Education Longitudinal Study of 1988 (NELS: 88). NELS:88 is another national, large-scale longitudinal study, and it surveyed 1032 schools, each of which contributed as many as 26 eighth grade students to the base year. Thus, NELS:88 contains more information on peers in the same school. However, the NELS:88 is focused on students' secondary school experiences, and it does not ask where a student lived or how old he or she was at kindergarten entry. This early childhood stage information is crucial for my study of relative age effect, because the effect is likely to be most significant in the beginning grades.

In short, despite its minor limitations, I use the ECLS-K survey. It is a large-scaled, comprehensive, and nationally representative.

#### **IV. EMPIRICAL STRATEGY**

In this study, I address four main questions:

- (1) Do empirical data support a *causal* relationship between children's relative age and their initial academic performance?
- (2) If so, does this effect *persist* sufficiently into later grades?
- (3) Does the effect differ for children of different gender?
- (4) Does the effect differ for children whose parents have low versus high educational attainment?

As discussed in the previous section, Datar (2006) suggests a persistent effect from the early childhood relative age advantage. However, under Elder's hypothesis (2009), the relative age effect from early childhood would not persist. I examine the first question and the second question by regressing the relative age effect on math and reading scores in each of the six waves of data collection (from kindergarten through eighth grade). Since the children in each wave are the same people, the relative age effect will only change with their grade if the *effect* changes.

For my third question, I examine whether the relative age effect differs by the gender of the student. Specifically, I run regressions for the two gender groups separately.

Similarly, to answer the fourth question, I look for heterogeneity by parents' education in the relative age effect. Again, I divide the entire sample into two sub groups: one group consisting of children whose parents' highest educational degree is a bachelor's degree or above; and another group consisting of children whose parents' highest educational degree is below a bachelor's degree.



In summary, I run a total of 60 separate instrumental variables regressions:

- Full sample set from six different waves of data collection, by 2 subjects
- Female group from six different waves of data collection, by 2 subjects
- Male group from six different waves of data collection, by 2 subjects
- High Parent's Education Level Group from six different waves of data

collection, by 2 subjects

- Low Parent's Education Level Group from six different waves of data

collection, by 2 subjects.

The *relative age* variable is my key explanatory variable, and it is measured in months. I use months as the unit rather days in order to flexibly divide relative age into brackets.

Numerically, *relative age* is computed by subtracting the median age of *all* children in the same grade in ECLS-K from each child's individual age, measured in months. For instance, a child's relative age could be +7 months or -2 months compared with the median age of the full sample.

Secondly, I control for socio-demographic variables.  $X_i$  in the naïve regression function below represents a set of control variables including child gender, child race, child health condition, the number of siblings at home, parent education attainment, and household income.  $X_i$  also includes three indicators of family income: Low Income ( $\leq \$30K$ ), Middle Income ( $\$30K-75K$ ), and High Income ( $\geq \$75K$ ). These variables control for household income more flexibly than a linear control for

household income measured continuously. I omit the Middle Income indicator variable to avoid co-linearity. Finally,  $X_i$  includes seven indicator variables for a child's race. The indicator for white race is omitted to avoid co-linearity, which means that I use the white race as a benchmark.

A naïve regression would be

$$(1) \quad \text{Outcome}_i = \beta_0 + \beta_1 \text{RELAGE}_i + \mathbf{X}_i' \delta + \varepsilon_i$$

where

$\text{Outcome}_i$  is the test score of a child

$\text{RELAGE}_i$  is the key explanatory variable *relative age*

$\mathbf{X}_i$  is a vector of control variables

However, there would be two major problems with the regression form above:

#### ***A. Different Assessment Test Time***

Even though children in the ECLS-K cohort are taking the same exam every year, the exact testing date varies a lot owing to practical constraints regarding testing. That is, the ECLS-K staff cannot arrange for every child to be tested on the same day.

For instance, the assessment date varies from March 2002 to July 2002 during the third grade data collection process. That is to say, if Mike is 1 month older than George, but Mike happens to take the test 4 months earlier than George does, the positive relative age advantage on Mike (if any) might very likely be offset by the

disadvantage of taking the test earlier much earlier. Hence, I create a new control variable, *relative assessment time*, to control for differences in the assessment data across students in the survey.

*Relative assessment time* is measured in months (for example, 123 months), which is consistent with other time-related variables in this study. Numerically, I compute the *relative assessment time* variable by subtracting *relative age* (for example, +3months) from *age at assessment* (for example, 120 months).

### ***B. Unobserved Qualities Motivating Family to Redshirt a Child***

One can argue that parents from different socio-economic backgrounds have different propensities to start their child in kindergarten at an early age. For instance, research has shown that schools and teachers serving low-income populations have different preferences about age-at-entry than those serving high income populations. Thus, some of the variation in a child's relative age could be endogenous to a child's intellectual maturity or correlated with omitted aspects of his or her family background. As a result, a naïve regression of a child's achievement on his relative age as shown in formula (1), might not generate a coefficient that reflects the *causal* effect of relative age. To remedy this problem, I construct an instrumental variable: the child's *predicted relative entrance age* which is derived from the interaction between the child's birthday and cut-off date for school enrollment in the child's state

back in the year 1998-1999. This instrumental variable is measured in months as well.

For instance, if a child is born in August and the school entrance cut-off is September, then this child's *predicted relative entrance age* would be 1 month on the cut-off date when he is entering the school. However, if another child who is born in October and obeys the school entrance law, then this child will have to wait until next year to enter school, thus his *predicted relative entrance age* will be 11 months.

This instrumental variable is highly correlated with actual relative age, but it is very plausibly uncorrelated with idiosyncratic, unobserved determinants of a child's maturity and aptitude because state laws are set at a high level compared to an individual student. That is, even if state laws were somehow to reflect policy makers' long term view about the usual maturity of students in their state, the laws would be arbitrary with respect to the small number of students surveyed by the ECLS-K. Therefore, my instrumental variables estimates credibly identify the causal effect of a child's relative age on his achievement.

In the first stage regression, I use the instrument *predicted relative entrance age* to predict my key explanatory variable *relative age*, and then in the second stage, I plug the predicted relative age from previous regression into the new second stage regression.

The revised regression is:

$$(2) \quad \text{RELAGE}_i = \beta_0 + \beta_1 \text{ENTAGE}_i + \mathbf{X}_i' \delta + \varepsilon_i \quad \rightarrow \text{Predicted RELAGE}_i$$

$$(3) \quad \text{Outcome}_i = \alpha_0 + \alpha_1 \text{RELAGE}_{i\text{predicted}} + \mathbf{X}_i' \gamma + \varepsilon_i \quad \rightarrow \text{Predicted Outcome}_i$$

where

ENTAGE<sub>i</sub> is the new instrument variable: *predicted relative entrance age*

Outcome<sub>i</sub> is the test score of a child

RELAGE<sub>i</sub> is the key explanatory variable *relative age*

$\mathbf{X}_i$  is a vector of control variables

I use this instrumental variable specification on the full sample, on groups defined by gender, and by groups defined by parents' education. Each regression is specific to a survey wave as well as by academic subject. That is, I conduct 60 separate regressions. I also show tests for weak instruments.

Tables 2, 3 and 5 in the Results section below present the relative age estimates. In the Table 3, column (1) and column (4) are drawn from Table 2 to serve as a comparison to the gender-specific effects.

Table 4 and Table 6 in the Results section show the estimated relative age effect in standard deviation units. That is, I divide each of the estimated coefficients in Table 3 and Table 5 by the standard deviation of scores in the relevant wave and subject in the full sample. Expressing coefficients in this standardized way is customary in education research. Note that even for an estimate from the regression on females, I still divide that estimate by the standard deviation of the full sample's

scores (rather than the standard deviation of females' scores). In this way, I present the magnitude of the female-specific effect relative to the performance of the all children. One could alternatively divide a female-specific coefficient by the standard deviation of females' scores. However, that magnitude would not provide direct information to parents who are wondering whether to hold their little girl out of school for a year. What they care about, presumably, is how much their child's test score would rise relative to all children in her grade in the nation.

## V. RESULTS

In this section, I examine the relative age effects on childhood academic achievement (reading and math) using the ECLS-K cohort's kindergarten-grade up to eighth-grade scores. Before showing results of gender-specific effects (as per my third question) and parent educational background-specific effects (as per my fourth question), I first show results regarding the relative age effect's causality (as per my first question) and its persistency (as per my second question) from the full sample regressions. These results are the basis for the third and fourth research questions of my study, and also give a sense of how good the *instrumental variable regressions* as well as the constructed instrument variables are.

### 1. Results of Overall Trends, Statistical Significance, and Test of Weak Instrument

Column (1) and column (3) of Table 2 present the estimated coefficients of the relative age effect on IRT scores. Column (2) and (4) show the magnitudes of the effect by dividing those coefficients in column (1) and (3) using the standard deviation of scores from their respective tests.

**Table 2: Coefficient of Relative Age from Full Sample IV Regressions**

*Each cell in column (1) and column (3) excluding the last row contains the estimate from a separate regression (with its standard error in parentheses and its iv regression F-statistic in brackets).*

Different Waves of Tests on Same ECLS-K Cohort Children	Dependent Variable: Full Sample Reading Score (1)	Magnitude of Full Sample Relative Age Effect: Reading (2)	Dependent Variable: Full Sample Math Score (3)	Magnitude of Full Sample Relative Age Effect: Math (4)
K-fall	<b>0.3831 ***</b> (.0673) [859.0]	<b>0.0439</b>	<b>0.4605 ***</b> (0.0490) [983.1]	<b>0.0639</b>
K-spring	<b>0.5178 ***</b> (0.0783) [986.8]	<b>0.0473</b>	<b>0.6151 ***</b> (0.0575) [1067.9]	<b>0.0704</b>
First Grade	<b>0.5482 ***</b> (0.0941) [1040.8]	<b>0.0395</b>	<b>0.5049 ***</b> (0.0601) [1086.7]	<b>0.0548</b>
Third Grade	<b>0.6550 ***</b> (0.1322) [1052.7]	<b>0.0323</b>	<b>0.6110 ***</b> (0.1186) [1048.5]	<b>0.0281</b>
Fifth Grade	<b>0.3017 *</b> (0.1743) [801.6]	<b>0.0129</b>	-	-
Eighth Grade	-	-	-	-

\* Statistically significant at 10% level

\*\* Statistically significant at 5% level

\*\*\* Statistically significant at 1% level

Note: If an estimated coefficient is not significant at the 10 percent level, then it respective cell is noted as “-”.

As shown in column (1) and column (3), the coefficients decrease in significance over time. The effects of relative age for both math and reading turn insignificant by the time of the eighth grade. The relative age effect disappears

somewhere between the third and the fifth grade for math scores, whereas the effect is still mildly significant (at the 10% level) at the fifth grade level for reading scores.

The estimate from the first row (K-fall) and the first column (Full Sample Reading Score) of Table 2 is 0.3831, representing the fact that if a child's age relative to other kindergarteners increases by 10 months then his kindergarten-fall reading score is predicted to rise by 3.831 points. Reading down the first column, the estimate from the third row (First Grade) is 0.5482, which implies that if this child's relative age increases by 10 months then his first grade reading score is predicted to increase by 5.482 points. It is worth noting that even though 5.482 points is larger than 3.831 points, this does not necessarily suggest a growing magnitude of the relative age effect from kindergarten-fall to first grade, because raw score points from separate tests should not be compared directly in terms of effect magnitude. Instead, one should compare the ratio of a coefficient over its respective standard deviation of test scores to put the magnitude of the effects into perspective.

Column (2) and column (4) of Table 2 present the magnitude of the relative age effect at a particular point in time. The magnitude is calculated as the coefficient divided by the standard deviation of the scores from the full sample. For instance, the standard deviation for the full sample reading score in kindergarten-fall is 0.0439, which equals to 0.3831 (coefficient) over 8.7336 (standard deviation of scores), hence the magnitude of the relative age effect on reading in kindergarten-fall is 4.39 percent of a standard deviation. Put in another way, if a child's family holds him out of



school for an extra year—so that his relative age goes up by 12 months—then his reading score in kindergarten-fall is predicted to go up by 52.7 percent of a standard deviation (0.0439 times 12). Comparing column (2) with column (4) horizontally, one will find that the effect has a much larger magnitude on math from kindergarten up through the first grade. Following that, the effect nearly halved in size from first grade to third grade for math. This results in a smaller magnitude, as described above, on reading relatively to math, in the third grade as well as from the third grade onwards.

Reading down column (2) or column (4) one will see how the magnitude of relative age effects on academic achievement changes over time. The relative age effects on both reading and math increase from kindergarten-fall to kindergarten spring, and then decrease from kindergarten-spring onwards. This first-increase-and-then-decrease trend at kindergarten-spring is a new finding in this study.

Since this paper adopted instrumental variable regressions as its empirical strategy, it is crucial to examine whether the instrument variable *predicted relative entrance age* is adequately correlated with the endogenous variable, *relative age*. I use stata post-estimation command *estat firststage* following *ivregress*. *Observed* test results are far higher than the threshold F statistics in all the IV regressions. The minimum F statistic on the Table 2 is 859.0, which is shown in the first row and column (1) of Table 2, is from the regression of relative age effects on the kindergarten-fall reading scores).

## 2. Results from Gender Specific-Effect Regressions

I divide the full sample set into two gendered subsets, and perform separate regressions for each gender group. Columns (2) (3) (5) and (6) of the Table 3 below present the heterogeneous effects by gender.

**Table 3: Coefficients From Instrumental Variable Regressions for Relative Age Effect:**

### Full Sample Regressions versus By Gender Regressions

*Each cell contains the estimate from a separate regression (with its standard error in parentheses).*

Different Waves of Tests on Same Children (with its number of observations in brackets)	Dependent Variable: Reading Score			Dependent Variable Math Score		
	Full Sample	By Gender		Full Sample	By Gender	
		Male	Female		Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
K-fall	<b>0.3831 ***</b> (0.0673)	<b>0.2346 *</b> (0.1128)	<b>0.4915 ***</b> (0.0812)	<b>0.4605 ***</b> (0.0490)	<b>0.3922 ***</b> (0.0811)	<b>0.5101 ***</b> (0.0591)
K-spring	<b>0.5178 ***</b> (0.0783)	<b>0.4177 ***</b> (0.1274)	<b>0.5979 ***</b> (0.4562)	<b>0.6151 ***</b> (0.0577)	<b>0.6152 ***</b> (0.0951)	<b>0.6121 ***</b> (0.0693)
First Grade	<b>0.5482 ***</b> (0.0941)	<b>0.6531 ***</b> (0.1527)	<b>0.4562 ***</b> (0.1151)	<b>0.5049 ***</b> (0.0601)	<b>0.5067 ***</b> (0.0982)	<b>0.5178 ***</b> (0.0730)
Third Grade	<b>0.6550 ***</b> (0.1322)	<b>0.5649 ***</b> (0.2165)	<b>0.7325 ***</b> (0.1618)	<b>0.6110 ***</b> (0.1186)	<b>0.5723 **</b> (0.1925)	<b>0.6322 ***</b> (0.1466)
Fifth Grade	<b>0.3017 *</b> (0.1743)	-	<b>0.4918 **</b> (0.2060)	-	-	-
Eighth Grade	-	-	<b>0.7578 ***</b> (0.2701)	-	-	<b>0.5207 **</b> (0.2262)

\* Statistically significant at 10% level

\*\* Statistically significant at 5% level

\*\*\* Statistically significant at 1% level

Note: Regressions in column (1) and (4) all use the full sample, which are the same from Table 2 and serve as a comparison to the gender specific-effects here. If an estimated coefficients is not significant at the 10 percent level, then its respective cell is noted as “-”.

Column (2) and column (5) of the Table 3 present estimates of the male-specific relative age effect. These two columns show that the male-specific effects disappear in fifth grade. In contrast, column (3) and column (6) of Table 3 show that female-specific effects persist longer and are still significant by the time of the eighth grade at the 95% level for both math and reading.

Take the estimate 0.5101 in the first row and column (6) and the estimate 0.3922 in the first row and column (5) of Table 3 for example. These two estimates suggest that for female children, 10 months of relative age advantage raises kindergarten-fall math scores by 5.101 points, and a same 10 months of relative age advantage raises a male child's kindergarten-fall math scores by 3.922 points only, much smaller as a score improvement compared with what females could gain. The below Table 4 also present a larger relative age effect on females in beginning grades, in standard deviation units.

**Table 4: The Magnitudes of Relative Age Effect:**

**Full Sample Regressions versus By Gender Regressions**

*Each cell contains the magnitude ratio of an estimated coefficient (taken from Table 2) from a separate regression divided by the standard deviation of the score from the respective test of respective testing year (or respective wave of test, since there are two waves of test in the kindergarten year: K-fall, and K-spring).*

Different Waves of Tests on Same Children	Dependent Variable: Reading Score			Dependent Variable: Math Score		
	Full Sample (1)	<i>By Gender</i>		Full Sample (4)	<i>By Gender</i>	
		Male (2)	Female (3)		Male (5)	Female (6)
K-fall	<b>0.0439</b>	<b>0.0269</b>	<b>0.0563</b>	<b>0.0639</b>	<b>0.0544</b>	<b>0.0708</b>

K-spring	<b>0.0473</b>	<b>0.0382</b>	<b>0.0546</b>	<b>0.0704</b>	<b>0.0704</b>	<b>0.0701</b>
First Grade	<b>0.0395</b>	<b>0.0470</b>	<b>0.0328</b>	<b>0.0548</b>	<b>0.0550</b>	<b>0.0562</b>
Third Grade	<b>0.0323</b>	<b>0.0278</b>	<b>0.0361</b>	<b>0.0281</b>	<b>0.0319</b>	<b>0.0352</b>
Fifth Grade	<b>0.0129</b>	-	<b>0.0210</b>	-	-	-
Eighth Grade	-	-	<b>0.0274</b>	-	-	<b>0.0236</b>

Note: if the numerator of the magnitude ratio—the estimated coefficient from its respective regression as shown in Table 3—is not significant at the 10 percent level, then the ratio's respective cell is noted as "-".

Table 4 gives the magnitude of the gender-specific effect in terms of test scores' standard deviation. The comparison of column (2) and (3) of Table 4 reveals a magnitude trend consistent with the coefficient trend in Table 3 – the relative age effect persists longer with females, and it is of a larger magnitude for females in the earlier years. The ratio 0.0563 from the first row and column 3 suggests that if a girl's relative age rises up by 10 months, her kindergarten-fall reading score is expected to rise by 56.3 percent of a standard deviation. However, if a boy's relative age rises by up 10 months then his score improvement is only a 26.9 percent of a standard deviation. The comparison of column (5) and (6) reveals a similar pattern for female and male students regarding the math scores – the relative age effect is more pronounced for female children in the beginning years and also persists all the way into female children's eighth grade.

### ***3. Results from Parent Educational Attainment Specific-Effect Regressions***

Table 5 and Table 6 show how the effects differ among students whose parents have high educational attainment versus low educational attainment. The comparison of columns (2) and (3) as well as the comparison of columns (5) and (6) in Table 5

show that the relative age effect is more pronounced in families with high parent educational attainment.

**Table 5: Coefficients From Instrumental Variable Regressions for Relative Age Effect:**

**Full Sample Regressions versus By Parent Educational Attainment Regressions**

*Each cell contains the estimate from a separate regression (with its standard error in parentheses).*

Different Waves of Tests on Same Children	Dependent Variable: Reading Score			Dependent Variable: Math Score		
	<i>By Parent Educational Attainment</i>			<i>By Parent Educational Attainment</i>		
	Full Sample (1)	High Parent Education (2)	Low Parent Education (3)	Full Sample (4)	High Parent Education (5)	Low Parent Education (6)
K-fall	<b>0.3831 ***</b> (0.0673)	<b>0.6269 ***</b> (0.1531)	<b>0.2717 ***</b> (0.0673)	<b>0.4605 ***</b> (0.0490)	<b>0.7076 ***</b> (0.1127)	<b>0.3581 ***</b> (0.0504)
K-spring	<b>0.5178 ***</b> (0.0783)	<b>0.8078 ***</b> (0.8232)	<b>0.3926 ***</b> (0.0819)	<b>0.6151 ***</b> (0.0577)	<b>0.9213 ***</b> (.1261)	<b>0.4852 ***</b> (0.0623)
First Grade	<b>0.5482 ***</b> (0.0941)	<b>0.8232 ***</b> (0.1854)	<b>0.4295 ***</b> (0.1080)	<b>0.5049 ***</b> (0.0601)	<b>0.7350 ***</b> (0.1146)	<b>0.4114 ***</b> (0.0707)
Third Grade	<b>0.6550 ***</b> (0.1322)	<b>1.1529 ***</b> (0.2433)	<b>0.4518 ***</b> (0.1576)	<b>0.6110 ***</b> (0.1186)	<b>0.8394 ***</b> (0.2240)	<b>0.5187 ***</b> (0.1394)
Fifth Grade	<b>0.3017 *</b> (0.1743)	<b>1.0553 ***</b> (0.3227)	-	-	<b>0.8936 ***</b> (0.2878)	-
Eighth Grade	-	<b>0.9263 **</b> (0.3628)	-	-	<b>0.6904 **</b> (0.2983)	-

\* Statistically significant at 10% level

\*\* Statistically significant at 5% level

\*\*\* Statistically significant at 1% level

Note: Regressions in column (1) and (4) all use the full sample, which are the same from Table 2 and serve as a comparison to the gender specific-effects here. If an estimated coefficient is not significant at the 10 percent level, then its respective cell is noted as “-”.

The effect of relative age advantage for children of parents with limited education disappears as early as in the third grade, yet it is still significant at a 95% level in the eighth grade for children of parents with high educational attainment.

This persistent effect in the eighth grade on children with highly educated parents is not present when we look at the entire sample as a whole (column (1) and (4) of both Table 5 and Table 6). It only emerges when we conduct separate regressions for background-specific subgroups.

**Table 6: The Magnitudes of Relative Age Effect:**

**Full Sample Regressions versus By Parent Educational Attainment Regressions**

*Each cell contains the magnitude ratio of an estimated coefficient (taken from Table 4) from a separate regression divided by the full sample standard deviation of the score from the respective test of respective testing year (or respective wave of test, since there are two waves of test in the kindergarten year: K-fall, and K-spring).*

Different Waves of Tests on Same Children	Dependent Variable: Reading Score			Dependent Variable: Math Score		
	<i>By Parent Educational Attainment</i>			<i>By Parent Educational Attainment</i>		
	Full Sample (1)	High Parent Education (2)	Low Parent Education (3)	Full Sample (4)	High Parent Education (5)	Low Parent Education (6)
K-fall	<b>0.0439</b>	<b>0.0718</b>	<b>0.0311</b>	<b>0.0639</b>	<b>0.0981</b>	<b>0.0497</b>
K-spring	<b>0.0473</b>	<b>0.0738</b>	<b>0.0359</b>	<b>0.0704</b>	<b>0.1055</b>	<b>0.0555</b>
First Grade	<b>0.0395</b>	<b>0.0593</b>	<b>0.0309</b>	<b>0.0548</b>	<b>0.0798</b>	<b>0.0447</b>
Third Grade	<b>0.0323</b>	<b>0.0568</b>	<b>0.0222</b>	<b>0.0281</b>	<b>0.0468</b>	<b>0.0289</b>
Fifth Grade	<b>0.0129</b>	<b>0.0451</b>	-	-	<b>0.0417</b>	-
Eighth Grade	-	<b>0.0335</b>	-	-	<b>0.0313</b>	-

Note: if the numerator of the magnitude ratio—which is the coefficient as shown in Table 5—is not significant at the 10 percent level, then the magnitude ratio is also noted as “-” in this Table 6.

Table 6 shows the magnitudes of parent education level-specific effects.

Children who are 10 months relatively older and have highly educated parents achieve a 98.1 percent standard deviation score improvement, but there is only a 49.7 percent

standard deviation improvement for children of low education parents, roughly only half of the effect magnitude of the former group. This not only suggests that children coming from well-educated families gain more if they are redshirted, but also suggest that they will be hurt more academically if they enter kindergarten at a relatively young age.

## VI. CONCLUSIONS

In this study, I conduct instrumental variable regressions for the relative age effect on math and reading scores from the ECLS-K children cohort (data collections started with kindergarten-fall and ended with eighth grade).

My approach has two advantages. First, I mitigate the concern that a child's actual relative age (the difference between a child's real age and the ages of his or her peers in the same ECLS-K cohort across the country in different types of schools) is potentially endogenous to his or her rate of maturation or other unobserved variables that may affect achievement. For this, I construct an instrumental variable: the child's *predicted relative entrance age*, which is derived from the interaction between the child's birthday and the cut-off date for school enrollment in the child's state. This instrumental variable is highly correlated with actual relative age, but it is very plausibly uncorrelated with idiosyncratic, unobserved determinants of a child's maturity and aptitude since it is derived from state laws.

Secondly, I take into consideration that in such a large-scale and geographically spread-out national survey as the ECLS-K, annual or bi-annual assessment tests are usually not conducted on exactly the same day. To illustrate, if a 5 year-old child takes the test 4 months earlier (or later) than another similarly-aged child, the two will score differently, on average, due to the discrepancy in age at testing date. Hence I construct a new control variable *relative assessment time*, controlling for the difference in data collection and testing dates across the nation.

#### ***A. Overall Causality and Longitudinal Trend***

There is a statistically significant causal relationship between relative age and academic achievement in early grades. This positive effect persists until the fifth grade for reading and it persists until the third grade for math.

The magnitude of the overall relative age effect first grows from kindergarten-fall to kindergarten-spring, and then it decreases from kindergarten-spring onwards. This overall decreasing longitudinal trend supports Elder's (2009) argument that relative age effects diminish and disappear in the long run. Thus, there appears to be little or no evidence of the permanent achievement effect upon which redshirting strategies are predicated.

#### ***B. Gender Specific-Effect***

The relative age effect persists longer and has a much larger impact on female children (especially in the beginning grades). For instance, if a girl's relative age



decreases by 10 months, her kindergarten-fall reading score is expected to drop by 56.3 percent of a standard deviation. However, a boy is expected to lose only 26.9 percent of a standard deviation for a 10 months drop in relative age. This suggests to policy makers the potential importance of gender-specific kindergarten screening policies. Interestingly, this finding contradicts the popular belief that younger boys might be hurt more if they start school on time because girls tend to mature earlier. What is worth even more attention is that this disadvantage will accompany female children through their early childhood into later years –the coefficients for female-specific relative age effects are still statistically significant at a 95% level in their eighth grade, for both reading and math.

This finding supports the hypothesis (Cascio and Schanzenbach, 2007) that boys and girls experience different developmental trajectories in early childhood, but it raises a new possibility: a larger and more persistent relative age effect for female students.

### ***C. Parent Educational Attainment Specific-Effect***

The effect for children of parents with limited education disappears as early as in the third grade, yet it is still significant at a 95% level in the eighth grade for children with highly educated parents. In addition, the effect has a much larger magnitude for children with better educated parents than for children with parents of low educational attainment level.

That is to say, if a child from a well-educated family background is being held out of school to accrue more relative age advantage, he or she is likely to achieve a larger and more persistent academic gain, compared with redshirted children whose families lack sufficient educational resources at home. Those parents coming from a rich and well-educated background may be able to provide a greater amount of human capital to children at home, and make up for or even substitute for formal education. This could be the reason for the increasing popularity of redshirting, as discussed in the introduction section of this paper. Furthermore, well-educated parents who redshirt their children may have decided that keeping children home for an extra year (rather than sending them to school) can provide the most beneficial human capital to the child, hence for these families redshirting could be a practice in line with the optimal path of human capital investment.

However, as Becker suggests, one should always keep in mind the opportunity cost associated with redshirting. The opportunity cost of keeping a child at home rises with parents' education level and salary level. An extremely well-educated mother would give a great deal of income to stay home with her child. This is another cost-benefit question that parents need to consider.

On the policy side, my study suggests that policy makers should be particularly concerned about underprivileged families which adopt the practice of redshirting, because, all else equal, children of parents with low educational attainment gain much less from entering school later, as compared with peers whose parents are highly

educated. Moreover, if state-level policy moves kindergarten entrance date earlier, which would mean increasing the average age of children in kindergarten, children of parents with low educational attainment might fall even more behind due to the lack of proper educational resources at home during the extra time spent waiting outside of school. Also, if the kindergarten entrance age goes up, underprivileged children would likely be at greater risk of dropping out of school when they reach the legal age of school exit, while children from privileged families would be more likely to continue with schooling, as research has already shown. All this will result in a bigger achievement gap between children from different social-economic backgrounds, which is in line with the findings of Deming and Dynarki (2008).

In short, this study maps out longitudinal trends of relative age effects on two different academic subjects (math and reading). It finds distinct gender-specific effects and parent educational attainment-specific effects. These findings shed light on potential revisions of future policies regarding school entrance age.

## REFERENCE

1. Becker, G. *Human Capital*, 3<sup>rd</sup> edition (University of Chicago, 1993)
2. Bickel, D., Zigmond, N., and Strayhorn, J. *Chronological Age at Entrance to First Grade: Effects on Elementary School Success* (Early Childhood Research Quarterly 1991)
3. Black, S. and Devereux, P. *Too Young to Leave the Nest? The Effects of School Starting Age* (2011)
4. Cascio, E. and Schanzenbach, D. *First in the class? Age and education production function* (National Bureau of Economic Research, 2007)
5. Deming, D. and Dynarski, S. *The Lengthening of Childhood* (National Bureau of Economic Research, 2008)
6. Elder, T. *Kindergarten Entrance Age and Children's Achievement: Impact of State Policies, Family Background, and Peers* (Michigan State University, 2008)
7. Gladwell, M. *Outliers: The Story of Success* (2008)
8. Golant, S. and Golant, M. *Kindergarten: It Isn't What It Used to Be* (1999)
9. Gootman, E. *Preschoolers Grow Older as Parents Seek an Edge* (The New York Times, October 19, 2006)
10. Gordon, A. *OLR Backgrounder: Narrowing Kindergarten Entrance Age* (OLR Research Report, 2011)
11. Lincove, J. and Painter, G. *Does the age that children start kindergarten matter? Evidence of Long-Term Educational and Social Outcomes* (University of Southern California, 2004)
12. Malone, L. *The early reading and mathematics achievement of children who repeated kindergarten or who began school a year late* (U.S. Department of Education, 2006)
13. Maurin, E. and Ouazad, A. *Peer Effects in English Primary Schools* (2006)

14. Reaney, M., West, J., and Denton, L. *Is time a gift? Cognitive performance of kindergarten children who repeat kindergarten or whose entry is delayed* (2001)
15. Saluja, G., Scott-Little, C., and Clifford, R. *Readiness for School: A Survey of State Policies and Definitions* (Early Childhood Research and Practice, 2000)
16. Stipek, D. *At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents* (Social Policy Report, 2002)