working paper

11 26R

When Should Children Start School?

Dionissi Aliprantis

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 11-26R March 2012*

When Should Children Start School?

Dionissi Aliprantis

Kindergarten-entrance-age effects are difficult to identify due to the nonrandom allocation of entrance-age and simultaneous relative-age effects. This paper presents evidence that instrumental variable frameworks do not identify age effects for the youngest children of a cohort using the results of statistical tests for essential heterogeneity in initial enrollment decisions. Restricting attention to the oldest children in a cohort yields a sample with quasirandom variation in entrance and relative ages. This variation is used to identify the parameters of education production functions in which both entrance and relative ages are inputs for achievement. Estimates of entrance-age parameters from the ECLS-K data set are positive, large, and persist until the spring of third grade. Relative-age parameters are smaller, tend to be negative, and fade-out for math achievement by third grade. For the average child in our sample these estimates imply that both an earlier entrance cutoff date and an earlier birthdate will increase achievement if the child remains eligible. There is extreme heterogeneity in effects by gender and home environment, and these results are likely to be the most relevant for policy.

Keywords: Kindergarten Entrance Age, Redshirting, Early Childhood Education, Essential Heterogeneity, Strong Ignorability.

JEL Classification Numbers: H40, I21, I28, J18.

*Original version October 2011.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland, and he can be reached at dionissi.aliprantis@clev.frb.org. The author thanks Becka Maynard, Francisca G.-C. Richter, Ken Wolpin, Petra Todd, Fabienne Doucet, Guhan Venkatu, Russ Cole, Rob Grunewald, Kimber Bogard, Tim Dunne, Julie Cullen, and participants at the Penn IES Seminar and Penn Empirical Micro Club for helpful comments. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C050041-05 to the University of Pennsylvania. The opinions expressed are those of the author and do not represent views of the U.S. Department of Education.

1 Introduction

Investing in early childhood education appears to be good policy from several perspectives. Empirical evidence from studies such as the Perry Preschool Experiment and the Abecedarian Project indicate that investments in early childhood can significantly improve outcomes over the life-cycle (Currie (2001), Heckman et al. (2010b)). Even if one ignores the normative arguments that such interventions could help move our society towards greater equality of opportunity, there are still compelling arguments for investing in early childhood education. There is reason to believe educational investments made at early ages yield higher returns than similar ones made later in the life cycle (Cunha et al. (2005)), and researchers have calculated the returns to society from such investments to be extremely high (Heckman et al. (2010a), Rolnick and Grunewald (2003)).

Policy-makers have responded to this research by increasing investments in early childhood education during recent years. A related concern that has emerged is how to best maintain gains made from investments in early childhood as children transition to kindergarten and beyond (Bogard and Takanishi (2005), Jacob et al. (2010)). One policy that can affect the persistence of these gains is the age at which children enter kindergarten. Children in the United States are eligible to enroll in kindergarten if they turn five before a specific date set by their state of residence, known as an entrance cutoff date. In recent decades many state governments have chosen to move entrance cutoff dates earlier, increasing the average age at which children start kindergarten in the US.[1] Parents and schools have also made choices during recent decades to increase the average entrance age (Elder and Lubotsky (2008), Weil (2007)).

It is difficult to judge the merits of entrance age policies and choices because it is difficult to identify entrance age effects.[2] This paper presents evidence that entrance age effects are likely to be unidentified for the youngest children of a given cohort in an instrumental variables framework. The paper does, however, present a related identification framework that separately identifies entrance and relative age effects for the oldest children in an entering class.

The allocation of entrance age is a key challenge to the identification of entrance age effects. Many analyses in the literature use date of birth and entrance age cutoffs to instrument for entrance age. However, adults choose whether to enroll children when first eligible, with the decision to delay a child's initial enrollment referred to as redshirting. There is likely to be a correlation between the unobservable factors determining the redshirting decision and the unobservable factors determining age effects.

Redshirting creates a violation of the monotonicity assumption necessary to identify parameters in settings where these unobservables are correlated (ie, under essential heterogeneity (EH); see Aliprantis (2011) and Barua and Lang (2009).). When the monotonicity assumption fails to hold under EH the entire instrumental variable framework breaks down, and no interpretable parameter

[1]For example, 22 states moved their cutoff dates to an earlier point in the school year between 1975 and 2000 (Stipek (2002)). See Figure 1 in Elder and Lubotsky (2008) for a summary of the evolution of these laws during the past 40 years.

[2]This is true despite the plethora of studies showing a positive correlation between entrance age and school performance. See Datar (2006) for an extensive discussion of this literature.

can be recovered from the data (Heckman and Vytlacil (2005)). The paper first presents statistical tests indicating that redshirting decisions are made in settings of essential heterogeneity rather than strong ignorability. An implication is that the identification strategy using date of birth and entrance cutoff dates to instrument for entrance age does not identify age effects on children for whom redshirting is prevalent.

The paper shows next that only a small share of the relatively oldest children in a given cohort are redshirted. The entrance and relative ages of these children are not determined by enrollment choices. This finding is used to select a sample from the ECLS-K data set for whom variation in age is generated by differences in individual-level birth dates and state-level entrance cutoff dates alone.

Another challenge to the identification of entrance age effects is related to the fact that changing a child's birth date changes both their entrance age and the age relative to classmates at which he or she is exposed to educational inputs.[3] As a result many treatment effect estimates in the literature are combinations of entrance and relative age effects, and thus their policy implications are ambiguous even if identified. The relationship between entrance and relative ages is addressed in this paper by specifying education production functions (EPFs) in which both entrance and relative ages are inputs for achievement. Parameters related to each of these ages are identified using the quasi-random variation in ages in the restricted sample from the first step. Careful consideration is given to this step since correctly specifying EPFs is a non-trivial task (Todd and Wolpin (2003), Rothstein (2010), Andrabi et al. (2011)). An important part of this step is identifying ability using achievement before school inputs have been applied to children.

The separate identification of entrance and relative age parameters allows for specific counterfactual questions to be addressed. Estimated EPF parameters imply that a state moving its entrance cutoff to an earlier date improves the scores of children in our sample. This result may be a surprise; it suggests that on average younger children catch up to older peers rather than being left behind them. Estimated parameters also imply that an earlier birth date increases test scores, but these effects are smaller than those from changing entrance cutoff dates.

The paper discusses several challenges to determining optimal entrance age policy. Among others, one challenge is that effects from changing the entrance cutoff date are remarkably heterogeneous by gender and home environment, with effects on boys being much stronger than those on girls. Viewing the empirical results together with these challenges indicates that attention should be given to designing policies that can accommodate the significant heterogeneity of entrance and relative ages in any cohort of children.

The remainder of the paper is organized as follows: Section 2 briefly discusses related policy and literature. Section 3 presents empirical evidence that redshirting choices are made under essential heterogeneity and discusses the implications of this evidence. Both the ECLS-K data set used in the analysis and some variable definitions are introduced in this section. Section 4 presents

[3] It also determines later outcomes such as the years of schooling students must complete before becoming eligible to drop out of school (Angrist and Krueger (1991)) and the age at which they first enter the labor market (Deming and Dynarski (2008)).

the identification strategy, with 4.1 first describing the selection of the sample used in estimation. Sections 4.2 and 4.3 specify the education production functions to be estimated and the variation in the data used to identify their parameters. Estimation results are presented in Section 5 along with discussions of their counterfactual and policy implications. Section 6 concludes.

2 Policy Background and Related Literature

The average age of children entering kindergarten has increased in recent decades in the United States (Elder and Lubotsky (2008)). This increase can be seen as the result of state-level policy changes moving entrance cutoff dates to an earlier point in the year. Since an older entrance age is empirically associated with better performance in school, these changes in policy can first be seen as a response to legislation such as 1994's "Goals 2000: Educate America Act" and 2002's "No Child Left Behind Act of 2001," which give schools and state governments increasing incentives to raise children's test scores.

However, high stakes testing and policy changes are not the only cause of the increasing kindergarten entrance age, as the average kindergarten entrance age began to increase before the introduction of high stakes testing (Deming and Dynarski (2008)). Recent changes in entrance age can also be attributed to increasing concerns about the "readiness" of children for school (Graue and DiPierna (2000), Graue (1993)). Evidence of these concerns is seen in parents and schools delaying the enrollment of many children until the year after they are first eligible to enroll or else having children repeat kindergarten. The prevalence of this practice, known as redshirting, is 8 percent for children in the ECLS-K, with increasing frequency for the relatively youngest in any cohort.[4]

Overall, the choices made by parents, teachers, and schools appear to be more important in explaining changes in the increasing average age at enrollment than do changes in entrance cutoff dates. Deming and Dynarski (2008) find that only a quarter of the decrease in attainment of six-year-olds in recent decades has come from changes in state entry laws, with the rest of the decrease coming from choices made by parents, teachers, and schools. Of those chosen to be delayed by third grade, Bedard and Dhuey (2006) find that 41 percent are behind due to delayed enrollment and 59 percent are behind due to retention.

One reason for increasing concerns about readiness is research showing the importance of the childhood environment. There is a line of research showing that investments in children and their parents before the age of five can have profound impacts on adult outcomes (Almond and Currie (2010)). Another line of research shows that outcomes determined by age 15 or 16 can explain a great deal of subsequent education and labor market outcomes (Eckstein and Wolpin (1999), Keane and Wolpin (1997), Neal and Johnson (1996), Keane and Wolpin (2000), Cameron and Heckman (2001)).

The evidence from this related research, together with recent policy changes, has led to a huge

[4]Other estimates of the prevalence of redshirting include 16 percent in the US (Deming and Dynarski (2008)) and 15 percent in Australia (Edwards et al. (2011)). The data used to construct this estimate will be discussed in depth in Sections 3.2 and 4.

increase in research investigating the effects of kindergarten entrance age. The majority of the evidence points to positive effects on academic achievement and grade progression from increased entrance age.[5] Using the international TIMSS data set, Bedard and Dhuey (2006) report findings on the effects of age across several countries. They find large net relative age effects on test scores at the 4th grade level. Although they find these effects persist into the 8th grade, they are considerably reduced by that time. Datar (2006) finds large effects of entrance age on test scores in kindergarten and first grade for children in the ECLS-K data set. Using Chilean data, McEwan and Shapiro (2008) find that delaying enrollment has significant positive effects on outcomes such as retention and standardized test scores. And data from Germany and Sweden indicate higher relative age is associated with higher academic attainment (Puhani and Weber (2005) and Fredriksson and Öckert (2005)).

A smaller body of evidence finds smaller or even negative effects from increased entrance or relative age. An example is Black et al. (2011), which examines data from Norway and finds little evidence of long-term entrance age effects on IQ or earnings. As well, Elder and Lubotsky (2008) find evidence that age-related differences in academic performance dissipate as children advance in school, attributing most of the initial differences to the accumulation of skills before children enter kindergarten. And while both Elder and Lubotsky (2008) and Dobkin and Ferreira (2010) find that relatively younger children are more likely to be held back, Elder and Lubotsky (2008) find that younger children have higher achievement and Dobkin and Ferreira (2010) find that younger children have higher academic attainment.[6] Finally, Cascio and Schanzenbach (2007) examine data from Project STAR and find no effects on achievement from relative age.

3 Redshirting Choices and Treatment Effect Heterogeneity

3.1 Model and Test of Strong Ignorability

We begin our analysis by investigating whether redshirting decisions are made in a setting of Strong Ignorability (SI) or Essential Heterogeneity (EH). It has been shown that due to redshirting, an exogenous change in birthdate would increase entrance age for some individuals while decreasing it for other individuals (Aliprantis (2011), Barua and Lang (2009)). This violation of the monotonicity assumption from Imbens and Angrist (1994) is crucial if redshirting decisions are made under EH, because in this scenario the entire instrumental variable framework breaks down, leaving parameters of interest unidentified (See Aliprantis (2011) and Heckman and Vytlacil (2005) for related discussions.).

We can study these alternative assumptions using a joint model of selection into treatment and potential outcomes. One specific age treatment of interest is the effect of delaying entrance by an

[5] In addition to the literature on the effects entrance age has on academic outcomes, there is also a literature documenting the important role of relative age in competitive sports (Muscha and Grondin (2001)) and leadership (Dhuey and Lipscomb (2008)).

[6] Dobkin and Ferreira (2010) also find that relative age has little if any effect on adult outcomes such as employment, wages, or home ownership.

entire year. Define a binary treatment variable as

$$D_i = \begin{cases} 1 & \text{if child } i \text{ is redshirted;} \\ 0 & \text{if child } i \text{ enters when first eligible.} \end{cases} \tag{1}$$

Assume potential outcomes in each treatment state, $Y(0)$ and $Y(1)$, are functions of observable characteristics X_D and some treatment level specific unobservable component U_j for $j \in \{0,1\}$:

$$\begin{aligned} Y(0) &= \mu_0(X_0) + U_0 \\ Y(1) &= \mu_1(X_1) + U_1. \end{aligned} \tag{2}$$

We are interested in identifying features of the distribution of the treatment effect $\beta_i \equiv Y_i(1) - Y_i(0)$, and in this analysis the outcome of interest is achievement.

We further suppose that individuals select into treatment based on a latent index:

$$D = \begin{cases} 1 & \text{if } D^* \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Assuming $X_0 = X_1$, each individual's latent index D^* depends on observable characteristics X, an instrument Z, and some unobserved component V as follows:

$$D^* = \mu_D(X,Z) - V \tag{4}$$

$$= \mu(X) + \gamma Z - V, \tag{5}$$

Assigned relative age is used as an instrument in this analysis, and its application is discussed in detail in the following sections. We define the propensity score conditional on Z to be $\pi^Z(X) \equiv F_V(\mu_D(X,Z)) \equiv Pr(D = 1|X,Z)$, with the conditional quantiles of the unobservable characteristics determining selection into treatment denoted by $u_D = F_{V|X}(V|X)$. We assume A1-A6 from Heckman and Vytlacil (2005).

As developed in Imbens and Angrist (1994), Heckman and Vytlacil (2005), and Heckman et al. (2006), models assuming EH allow selection into treatment to be determined in part by anticipated gains from treatment that depend on individual level characteristics that are unobserved by the econometrician. To be precise, Heckman et al. (2006) define Essential Heterogeneity (EH) as the scenario in which treatment effect heterogeneity is correlated with selection into treatment even conditional on observables,

EH: $COV(U_1 - U_0, U_D) \neq 0.$

A stronger assumption often made in the statistics and economics literature is that agents select into treatment based only on the observable characteristics measured in X. This assumption, typically referred to as Strong Ignorability (SI), can be stated as follows:

SI: $U_1 - U_0 \perp\!\!\!\perp U_D$.

Note that EH can alternatively be written as $COV(\beta, D) \neq 0 \,|\, X$, and SI can alternatively be written as $\beta \perp\!\!\!\perp D \,|\, X$.

The case in which the unobserved characteristics determining selection are not related to treatment effect heterogeneity, as assumed under SI, has a testable implication that all instruments identify a homogeneous parameter. Heckman et al. (2010) show that the null hypothesis

$$H_0 : \beta \perp\!\!\!\perp D \,|\, X,$$

can be stated equivalently as

$$H_0^{IV} : \beta_{IV, [\underline{p}_1, \overline{p}_1]} = \beta_{IV, [\underline{p}_2, \overline{p}_2]} \;\mid\; X \tag{6}$$

where

$$\beta_{IV, [\underline{p}_j, \overline{p}_j]} \;\mid\; X = \frac{Cov\Big(\; Y \, , \, P(Z) \;\mid\; P(Z) \in [\underline{p}_j, \overline{p}_j] \, , \, X \; \Big)}{Var\Big(\; P(Z) \;\mid\; P(Z) \in [\underline{p}_j, \overline{p}_j] \, , \, X \; \Big)} \quad \text{for} \quad j = 1, 2. \tag{7}$$

The Wald test statistic of H_0^{IV} is asymptotically distributed as a χ_1^2 random variable, and Heckman et al. (2010) characterize the properties of this test under a variety of scenarios.

The null hypothesis in Equation 6 is based on the fact that under SI all instruments identify the same homogeneous parameter. Using the propensity score as an instrument over different intervals of its support should thus identify the same parameter in the absence of EH. That is, under SI exogenous variation in the propensity score should have the same impact on the outcome variable, regardless of where in the support of the distribution of propensity scores this variation occurs.

That SI implies H_0^{IV} can be seen clearly in the Marginal Treatment Effect (MTE) framework developed in Heckman and Vytlacil (2005) and originally introduced by Björklund and Moffitt (1987). Full discussions of the implications of EH and SI can be found in Heckman and Vytlacil (2005), Heckman et al. (2006), and Aliprantis (2012). Here we will only define the Marginal Treatment Effect (MTE) as

$$\triangle^{MTE}(x, u_D) \equiv E[Y(1) - Y(0) \,|\, x, u_D], \tag{8}$$

to allow us to illustrate the contrast between SI and EH in Figure 1.

3.2 Data

The analysis presented in this paper utilizes data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) data set, which is a nationally representative sample of over 22,000 children enrolled in over 1,200 schools who started kindergarten in the fall of 1998. Data were collected during the the fall and the spring of kindergarten (1998-99), the fall and spring of 1st grade (1999-2000), the spring of 3rd grade (2002), 5th grade (2004), and 8th grade (2007) from the children, their parents/guardians, teachers, and school administrators.[7] Since observations in Round 3 (Fall of 1st grade) were only gathered from about 30 percent of the sample, this round will not be considered for the analysis. To account for the sampling scheme used to collect the ECLS-K, weights will be utilized in all estimation.

3.2.1 Variables

Following the terminology in Bedard and Dhuey (2006), we refer to the relative age at which a child would be observed if they entered kindergarten when first eligible as *assigned relative age*, and the child's actual age relative to their school's cutoff date as *observed relative age*. We define monthly groups for each of these relative ages. For the cohort of children we observe entering school in the fall of 1998 we will observe children entering who are too young to be eligible, but who enter anyway, children who enter when first eligible, and children who were eligible to begin school in the fall of 1997, but whose parents decided to enter them in kindergarten in the fall of 1998. We define the youngest of these children to be in the relative age group M_1, and the oldest to be in the relative age group M_{36}. These relative age groups can be seen in Figure 2. We also define birth cohorts $\{B_1, B_2, \ldots, B_{12}\}$ using these assigned relative age groups:

$$B_j = \begin{cases} M_j & \text{if } i \in M_j \text{ and } M_j \in \{M_1, \ldots, M_{12}\}; \\ M_j - 12 & \text{if } i \in M_j \text{ and } M_j \in \{M_{13}, \ldots, M_{24}\}; \\ M_j - 24 & \text{if } i \in M_j \text{ and } M_j \in \{M_{25}, \ldots, M_{36}\}. \end{cases}$$

In order to assign children in the ECLS-K to these relative age groups, we first construct their assigned entrance and relative ages. Let T(Calendar Date) denote a function of dates from the calendar to a time line in which day 1 is January 1, 1990. We define each child's entrance age EA to be their age in months over 5 years at the start of the school year when they began school, and we implement this definition empirically as the child's age over 5 on September 1st, 1998. We also define a child's assigned relative age RA for the entire sample to be the amount of time before the entrance cutoff date they turned five years of age:

$$EA = T(\text{Sept. 1, 1998}) - T(\text{5th BDay}) \tag{9}$$

$$RA = T(\text{Cutoff Date}) - T(\text{5th BDay}) \tag{10}$$

Considering children who enter on time when first eligible, the youngest children will have a relative age of 0, and the oldest children will have a relative age of 12 months. Redshirted children will have relative ages between 12 and 24 months, and children who enter early will have relative ages between -12 and 0 months. These variables are shown in Figure 2 and will be discussed in depth in Section 4.

We use school-level entrance cutoff dates to construct entrance and relative age variables. In

[7] Eighth grade will be the last round of data collection due to sample attrition.

the ECLS-K 6 percent of school-level cutoffs are not ascertained. For an additional 1 percent of children a school-level cutoff date is reported to be not applicable, and an additional 7 percent of children have an implausible cutoff date (1995 or earlier). We consider the data to be missing for all of these children, a total of 14 percent of the children in the ECLS-K. The outcome measures used in this analysis are math and reading Item Response Theory (IRT) test scores.

3.3 Empirical Implementation and Test Results

In order to test H_0^{IV} with data from the ECLS-K, we follow Heckman et al. (2010) and first estimate a probit model to obtain predicted propensity scores for redshirting. We estimate this probit on the sample of kindergarteners in relative age groups $\{M_{13}, M_{14}, M_{15}, M_{25}, M_{26}, M_{27}\}$ where the instrument Z is assigned relative age and its square. Time-invariant demographic characteristics X are also used to estimate propensity scores. Included in X are the number of children's books at home, whether the child ever received benefits from the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and whether the child's mother is present and works 35 hours or more as measured in the spring of kindergarten; the socio-economic status (SES) quintile of the child and the mother's education level measured in the fall of kindergarten; whether the father was present in the household during the fall of first grade; and gender and race dummies.

Figure 4 shows a histogram of the estimated propensity scores $\widehat{p}(z, x)$. We see the expected pattern that children who were redshirted tend to have higher estimated propensity scores.

The distinction between Z and X in our model is an extremely important one. Even in the absence of an instrument the treatment effect and MTE parameters we have defined all exist, but these parameters are only identified through an instrument. Examining our application, if relative age impacts outcomes not only through its effects on redshirting decisions, but also directly, then assigned relative age belongs in X, and under EH we cannot identify any parameters of our model. It is important to remember that all of the hypothesis tests that follow are related to the effect of redshirting, which changes both entrance age and relative age. A central focus of the paper is understanding the direct effects of relative age on achievement, and in later sections it will be assumed that relative age is an input in the production of achievement. Nevertheless, for the sake of identification the following hypothesis testing abstracts from the effects of relative age acting in isolation.

It is also important to note that the test in Section 3.1 is related to parameters that are defined conditional on observable characteristics X. Figure 3 shows the importance both of conditioning on observables and of the variation in treatment induced by the instrument. H_0^{IV} is based on the implication from SI that the MTE is constant in the unobservable component of selection (ie, that $E[Y(1) - Y(0)|x, u_D] = E[Y(1) - Y(0)|x]$ for all u_D). It is possible that SI will hold for some regions of observables but not for others. As well, SI might hold over some intervals of u_D but not over others. The variation in treatment induced by the instrument determines which MTEs are identified by the instrument. Consider that in our empirical estimates in Figure 5 the instrument induces fewer individuals into treatment at lower values of $\widehat{\mu}(X)$ than at higher values.

Determining how to condition on observables introduces the need to make arbitrary assumptions. One approach taken in Carneiro et al. (2011) is to estimate effects at the mean observables of the sample and then to perform hypothesis testing on these parameters. An alternative approach, implemented here, is to estimate effects and conduct tests in several regions of the distribution of observables.

Given the empirical distribution of observables shown in Figure 6, we proceed by creating subsamples defined by five equally-spaced intervals of the quantiles in which observable characteristics fall, $\theta = 1 \times \mathbf{1}\left\{\widehat{\mu}(X) \in [Q_{\tau=.125}, Q_{\tau=.275})\right\} + \cdots + 5 \times \mathbf{1}\left\{\widehat{\mu}(X) \in [Q_{\tau=.725}, Q_{\tau=.875})\right\}$. Next for each θ we separately estimate regressions of test scores on the predicted propensity score ($\widehat{p}(z, x)$) as well as the covariates in X with separate slopes and separate intercepts for the coefficients on $\widehat{p}(z, x)$ for those whose estimated propensity score is below the midpoint of the support of $\widehat{p}(z, x)$ conditional on θ and for those with $\widehat{p}(z, x)$ above the midpoint. The test of H_0^{IV} is empirically implemented as a Wald test of the hypothesis that these slope coefficients are equal over these two different intervals of $\widehat{p}(z, x)$.

Table 1 reports the results of these tests together with the estimates of the two coefficients. Overall 40 tests are conducted, with 8 being rejected at the 5 percent level and one more rejected at the 10 percent level. There is very strong evidence of EH for math, as 7 of 20 tests reject the null of SI at the 5 percent level. The evidence for reading is quite different, as only 1 of 20 tests is rejected at the 5 percent level.

Alternatively, Figures 7 and 8 show these data graphically. These Figures show estimates of mean test scores as quadratic functions of the estimated propensity score conditional on θ. Under SI and H_0^{IV}, MTEs should be constant in U_D conditional on θ, so each of these functions should be linear. Comparing math and reading at each point in time, these Figures visually replicate the statistical results reported in Table 1, that there is stronger evidence of EH for math achievement than reading achievement.

This evidence in favor of a model of EH rather than one of selection under SI is a rather negative result. EH and SI are fundamentally assumptions about how much the econometrician can observe. The ECLS-K is an extremely rich data set, and in many data sets less information will be observable. Combined with the evidence on monotonicity (Aliprantis (2011), Barua and Lang (2009)), these results indicate that date of birth (ie, assigned relative age) used as an instrumental variable does not identify age effects in the presence of redshirting.

4 Identification Strategy

4.1 Sample Selection

The preceding evidence discourages us from using assigned relative age as an instrument for entrance age in the presence of redshirting. However, this evidence does not preclude the possibility of estimating entrance age effects using an identification scheme inspired by Angrist and Krueger (1991) if we can find a sample that is not affected by redshirting. If we assume parents' decision rule for determining observed entry age does not change over time, cutoff dates stayed the same between 1997 and 1998, and that any seasonal patterns in number of births are repeated every year, then we may use the number of children in each birth cohort to estimate the percentage of children in each relative age group who enter early, when first eligible, or after redshirting. These estimates of the distribution of observations in each relative age group are presented in Table 2a, and Table 2b shows these estimates aggregated to the level of quarters.

The estimates presented in Table 2a guide the selection of a sample relatively unaffected by redshirting. Note that over a quarter of children who turned 5 within one month of their school's cutoff date, 26 percent, were redshirted. While this rate does decline as children become relatively older, it is still 11 percent for children who turned 5 in the fourth month before their school's cutoff date. Also note that a negligible number of children enter early, except for 6 percent in the oldest assigned relative age group. Finally, turning our attention to children who turned five between 6 and 11 months before their school's entrance cutoff date, over 96 percent of each cohort entered kindergarten when first eligible.

The sample used for the remaining analysis in this paper is composed of those kindergarteners in the relative age groups M_{18} through M_{23} because they are the least effected by redshirting. Furthermore, the sample is restricted to children living in states which set an entrance cutoff date between August 31st and January 1st. Children from states with Local Education Authority (LEA) options (Colorado, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, and Vermont), from the three states with entrance cutoff dates before August 31st (Alaska, Indiana, and Missouri), and from states with fewer than 30 respondents (Arkansas, Delaware, DC, Hawaii, Idaho, Montana, Nebraska, Nevada, North Dakota, Oregon, South Carolina, South Dakota, West Virginia, and Wyoming) are omitted from our sample. The 27 states included in our sample are: Alabama, Arizona, California, Connecticut, Florida, Georgia, Illinois, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Michigan, Minnesota, Mississippi, New Mexico, North Carolina, Ohio, Oklahoma, Rhode Island, Tennessee, Texas, Utah, Virginia, Washington, and Wisconsin.

This analysis will not consider outcomes after the third grade due to the severe attrition in the ECLS-K data set. Table 3 shows the sample size for math and reading IRT test scores by survey round. The Table reveals heavy attrition in grades 5 and 8 in the ECLS-K. In our sample 41% of initial respondents do not have reading IRT scores by the 5th grade, and 51% attrit by the 8th grade. The problem is even worse for math IRT scores.

4.1.1 Descriptive Statistics of the Sample

Figure 10 shows the distributions of math and reading IRT test scores of children in the sample. Figures 10a and 10b show how these distributions changed for all children between the Fall of kindergarten and the Spring of third grade. Math scores are more smoothly distributed than reading scores.

The remaining Figures in 10 show the distributions of test scores by state-level entrance cutoff dates. For math scores we see the expected relationship where the distribution of test scores shifts in a near monotonic fashion as entrance cutoff dates move earlier. Surprisingly, we do not observe such a relationship for reading test scores, as their distributions do not monotonically increase with a class's average entrance age. Reading test scores tend to be larger for the cohort of children facing the September 1st cutoff than the younger children facing an entrance cutoff date of October 1st. However, the even younger cohort facing an entrance cutoff date of December 1st, two months earlier, tends to have distributions that dominate those of the October 1st cohort.

In addition to state-level differences in outcomes, it is also important to remember that the distributions shown in Figure 10 are of cohorts with different age compositions, and the analysis in this paper is being conducted at the individual level. Monotonicity assumptions on the changes in individuals' experiences due to changes in entrance cutoff dates do not hold. Changing entrance cutoff dates changes the experiences of different individuals differently, some through both entrance and relative age effects, and others only through relative age effects. These distributions and related counterfactuals will be discussed in greater depth in Section 5.

4.2 Education Production Functions

Given the sample free from redshirting just defined, we now proceed to specify education production functions (EPFs) to be estimated on that sample. We follow Todd and Wolpin (2003) and assume the test score of student i in school j at time t is a function of the entire history of family and school inputs up until time t ($F_{ij}(t)$ and $S_{ij}(t)$), age at time t ($a(t)$), innate ability (μ_{ij0}), and an unobserved factor (ϵ_{ijt}):

$$Y_{ijt} = Y_t[F_{ij}(t), S_{ij}(t), a(t), \mu_{ij0}, \epsilon_{ijt}].$$

For the sake of tractability we assume a linear, separable production function:

$$Y_{ijt} = \alpha_t F_{ij}(t) + \beta_t S_{ij}(t) + \gamma_t a(t) + \delta_t \mu_{ij0} + \epsilon_{ijt}. \tag{11}$$

The unobservability of ability has been one of the key obstacles to empirically estimating EPFs like those in 11. Since it is likely that family and school inputs are correlated with ability, grouping ability into the error term and estimating Equation 11 as

$$Y_{ijt} = \alpha_t F_{ij}(t) + \beta_t S_{ij}(t) + \gamma_t a(t) + \eta_{ijt}$$

by OLS will yield biased parameter estimates. As discussed in Todd and Wolpin (2003), there are several specifications of EPFs that resolve this problem, but each makes a set of identifying assumptions that may or may not be appropriate for the application at hand. The way we will resolve this problem in the ensuing analysis is to assume ability is fixed over time, and then to include estimates of this ability in the specification of EPFs.

We obtain estimates of ability μ_{ij0} by assuming that in the fall of kindergarten ($t = 1$) school inputs have not yet impacted the production of educational achievement:

$$Y_{ij1} = \alpha_1 F_{ij}(1) + \gamma_1 a(1) + \delta_1 \mu_{ij0} + \epsilon_{ij1}. \tag{12}$$

The parameters of the EPF in Equation 12 are estimated using race as an instrumental variable that is correlated with family inputs but uncorrelated with ability. Noting that age $a(1)$ and ability should be uncorrelated in our sample and assuming μ_{ij0} and ϵ_{ij1} are uncorrelated with

$$\mu_{ij0} \sim N(0, \sigma_\mu^2) \qquad \text{and} \qquad \epsilon_{ij1} \sim N(0, \sigma_{\epsilon_1}^2),$$

2SLS estimation of

$$Y_{ij1} = \alpha_1 F_{ij}(1) + \gamma_1 a(1) + \eta_{ij1} \tag{13}$$

yields estimates of the residuals

$$\eta_{ij1} = \delta_1 \mu_{ij0} + \epsilon_{ij1} \sim N(0, \sigma_\mu^2 + \sigma_{\epsilon_1}^2),$$

that can be used as a measure of ability in the estimation of later production functions.

Race is an observable characteristic that is informative of many experiences that are unobservable to the econometrician. Thus even if race is uncorrelated with ability as assumed, it is likely that race will be correlated with the η_{ij1} from Equation 13 through the ϵ_{ij1}. Although this poses a problem for the proposed identification of ability, we proceed with this strategy because it appears to be the most palatable option. The alternative strategy is to simply use fall of kindergarten test scores as measures of ability, and this requires the assumption that age and family inputs between conception and the fall kindergarten test have no effects on test scores.

For the remaining points in time $t > 1$ (spring of kindergarten, spring of 1st grade, and spring of 3rd grade), we assume $F_{ij}(t) = F_{ij}$ and $S_{ij}(t) = S_{ij}$ for all t and estimate the following production function:

$$\begin{aligned} Y_{ijt} &= \alpha_t F_{ij}(t) + \beta_t S_{ij}(t) + \gamma_t a(t) + \delta_t \mu_{ij0} + \epsilon_{ijt} \\ &= \alpha_t F_{ij} + \beta_{t1} EA + \beta_{t2} RA + \beta_{t3} EA\, RA + \beta_{t4} s_{ijk} + \cdots + \beta_{tk} s_{ijk} + \delta_t \widehat{\eta}_{ij1} + \epsilon_{ijt}. \end{aligned} \tag{14}$$

The short time horizon we study makes an assumption of constant inputs reasonable despite evidence that both contemporaneous and past inputs impact test scores (Todd and Wolpin (2007)). Note that the β_{t1} and β_{t3} parameters we estimate are combinations of effects from age when receiving educational instruction and age at test effects.

There are several reasons for choosing this specification of the education production function. First, this specification interprets entrance age EA_i and relative age RA_i as elements of the vector of school inputs $S_{ij}(t)$. These ages make sense as elements of S_{ij} because they determine the age at which one is exposed to educational instruction in school, and also help determine the type of classroom and teacher interactions a student will have as a result of their age relative to their classmates.

The chosen specification also allows inputs to impact achievement differently at different times. It is important to allow these parameters to change over time, as their evolution over time is an open empirical question. There are theoretical reasons to think these effects might grow or shrink over time. On one hand, age differences as a share of total age become increasingly smaller as children age, and thus age effects might be dwarfed by other effects over time (Bogard and Berkley (2011), Magnuson et al. (2007)). On the other hand, institutional features such as tracking, or features of the developmental process itself, may mean that small effects at early ages are amplified at older ages (Cunha and Heckman (2010)).

4.3 Identifying Variation

Figure 2 helps to illustrate the variation in the sample used to identify the parameters of the specified education production functions. Moving from point b to a is the same as being born at an earlier date, all else being constant. We can see in Table 4 that this increases a child's entrance age EA as well as their relative age RA. Thus, date of birth alone is not enough to separately identify the effects of these variables. Moving from point b to d is the same as moving a child to a different state, all else being constant. In this case the child will have the same entrance age but will be a different age relative to her classmates. Thus variation in state entrance cutoff dates identifies relative age effects, $\{\beta_{2t}\}$. Now consider moving a child from point a to d. This is the same as moving the child's birthdate to a later date, but also moving the child to a state whose entrance cutoff date is earlier by the same amount of time. In this case, the child will be the same age relative to her classmates, but she will enter kindergarten at a younger age. Thus variation in birthdate, together with variation in state entrance cutoff dates, identifies the effects of entrance age, $\{\beta_{1t}\}$.

Although entrance age and relative age are linearly dependent within individuals, this is not the case between individuals. The variation in these two ages generated by differences in date of birth across individuals and entrance cutoff dates across states is shown in scatterplots in Figure 11. Figure 11a shows the variation in the entire ECLS-K in comparison to the variation in the sample as shown in Figure 11b. Note that due to the sample selection conducted to ensure that these ages vary randomly, we only observe children between the relative ages of 5 and 11 months. It must be remembered that the subsequent analysis can only speak to this sample of children, and is uninformative about age effects on children starting kindergarten at other relative or absolute ages.

If the identification scheme just described is to provide unbiased estimates of the entrance and relative age effects using the variation in ages shown in Figure 11, date of birth must be random conditional on observables (family inputs, school inputs, school fixed effects, and state fixed effects). However, considerable evidence indicates that date of birth is not random, but rather the product of parents' choices, biological factors, or both. Seasonal birth patterns have been demonstrated to be related to maternal characteristics such as age, marital status, education, and birth order

(Bobak and Gjonca (2001), Buckles and Hungerman (2008)), demographic characteristics such as income and race (Bound and Jaeger (2000)), and even tax schedules (Dickert-Conlin and Chandra (1999)). It is currently unclear how these socio-economic characteristics interact with biological factors (Rizzi and Dalla-Zuanna (2007)), temperature (Lam and Miron (1996)), and geographic location (Bobak and Gjonca (2001)) in determining seasonal birth patterns. Although this process has been formally modeled (Lam et al. (1994)), the parameters of such a model of seasonal birth patterns have yet to be identified. This leaves seasonality of births as a serious concern for our identification strategy.

Most studies of the effects of entrance age have dealt with concerns related to seasonal birth patterns by simply assuming that date of birth is exogenous. In support of this approach, Dickert-Conlin and Elder (2009) present compelling evidence that while birth date is clearly manipulated within short windows for nonmedical purposes, this does not appear to be the case in the windows around school entrance cutoff dates. Similar evidence is found in McEwan and Shapiro (2008). However, these studies focus on the manipulation of birth timing at the level of days or weeks through cesarean deliveries and the inducement of labor. Hence it is still possible that birth timing is planned around a less precise period such as month or quarter, leaving the identification strategy presented here open to the problems arising if parents make non-random choices, whether consciously or unconsciously, about when to have children.

We examine some evidence from the ECLS-K related to the seasonality of birth in Table 5. This Table presents results of F-tests on null hypotheses of equal means across seasonality of birth for several characteristics. Apart from race and living with one's father, there does not appear to be strong evidence of birth seasonality in our sample. Thus it may be reasonable to follow convention and assume that date of birth is exogenous, while at the same time acknowledging that the current analysis will be greatly improved by further research into seasonal birth patterns.

In addition to seasonal birth patterns, the endogeneity of state entrance cutoff dates is also a concern for our identification strategy, and we use 2000 Census data from the National Historical Geographic Information System (NHGIS, Minnesota Population Center (2004)) to investigate this issue. Although the EPFs we estimate all include state and school fixed effects, unobserved characteristics within states with particular entrance cutoff dates could still be driving results. Strong correlations between entrance cutoff dates and demographic characteristics would increase this concern. Figure 12 eases this concern by showing the correlations between entrance cutoff date and state demographic characteristics such as the state poverty rate, male labor force participation rate, and high school graduation rate are small and statistically insignificant.

5 Estimation Results and Discussion

Before estimating the EPF in Equation 14 we must first obtain estimates of math and reading ability, which we do by estimating the fall of kindergarten EPF in Equation 13. We assume there are three inputs into the production of achievement in the fall of kindergarten; age, mother's educational attainment, and household income. Assuming race is uncorrelated with ability but correlated with the fall test scores through the family inputs, we use race indicators as instruments for family inputs. The ability estimates obtained as residuals from 2SLS estimation of Equation 13 are shown in Figure 13.

These ability estimates allow us to estimate the EPFs in Equation 14. The home inputs included in estimation are income, mother's education, whether a foreign language is spoken at home, whether enrolled in a pre-kindergarten program, mother's employment between birth and kindergarten, whether the father was living in the household, and whether the child ever received WIC benefits. School inputs include entrance age, relative age, the interaction of the two ages, school and state fixed effects, gender, and race.

Estimates of the entrance age and relative age parameters in Equation 14 are displayed in Tables 6a and 6b. The first impression we might take away from the estimates in these Tables is that entrance age parameters grow between kindergarten and third grade in both math and reading. All else equal, one month of extra entrance age increases math IRT scores 0.11 standard deviations in spring of kindergarten, and this grows to 0.18 standard deviations in spring of third grade. The entrance age parameter for reading test scores similarly grows from 0.14 standard deviations in the spring of kindergarten to 0.17 in the spring of third grade.

A second impression from Tables 6a and 6b is that the parameters related to being relatively older than your classmates are quite different than the parameters for entering at an older age. These parameters tend to be small and are negative in the cases they are statistically significant. The relative age coefficient for math is strongly negative in the spring of first grade, but zero in the spring of kindergarten and the spring of third grade. And although the relative age coefficient for reading starts in the spring of kindergarten at –0.13 standard deviations, this shrinks to a statistically insignificant –0.07 by the spring of third grade.

In light of the evidence of essential heterogeneity provided in Section 3, we also estimate Equation 14 at different times by subgroup. Tables 7a-8b show separate effect estimates for boys and girls. Tables 9a-10b show similar estimates for children with different home environments as measured by the number of children's books they have at home.

The estimates indicate that entrance age effects are extremely different for boys and girls. By the spring of third grade the entrance age parameter for math scores is only 0.05 standard deviations for girls, but for boys it is 0.50 standard deviations! Although the trends in earlier grades are different for reading, we see a very similar pattern in the entrance age effect by the spring of third grade: increasing entrance age has very large and beneficial effects on boys, but barely impacts girls at all.

Relative age effects are also quite different by gender. By the spring of third grade, the relative age parameters for boys are negative and large at –0.21 and –0.22 standard deviations for math and reading, respectively. Just as they are for entrance age, relative age effects are also quite modest for girls, falling to 0.09 and –0.01 standard deviations for math and reading scores by the spring of third grade.

Finally, we investigate heterogeneity in effects by one measure of the home environment, the number of children's books the child has at home. We define children to have few books if they have fewer than 35 books, which is the 33rd percentile of the sample distribution, and we define children to have many books at home if they have more than 100 books, the 66th percentile of the sample distribution.

Entrance age and relative age have the largest effects on children with few books at home. For every month older these children are when they enter school, all else constant, these children score 0.31 and 0.42 standard deviations worse on their third grade math and reading tests. Relative age effects are of the opposite sign and in the spring of third grade are of even larger magnitude for these children, at 0.35 and 0.63 standard deviations for math and reading test scores.

Entrance age and relative age have smaller parameters for children with many books at home, and these parameters tend to be of the opposite sign of those for children with few books at home. Entering one month older increases test scores in the spring of third grade on average by 0.26 standard deviations for children with many books, but entering relatively older decreases test scores at this time by 0.11 standard deviations. The effects on reading test scores are similar but of smaller magnitudes; the entrance age parameter in spring of third grade is 0.13 and the relative age parameter is –0.02 standard deviations.

5.1 How Would an Earlier Entrance Cutoff Date Affect an Eligible Child's Achievement?

One issue of interest for policy makers is understanding how states making changes to their entrance cutoff dates would impact children. That is, we might be interested in understanding how an earlier entrance cutoff date would affect a child's achievement. We can use the estimated EPF parameters to investigate the effects of this counterfactual policy change.

It is important to note that changing the entrance cutoff date does not change all children's experiences in the same way. For those children whose eligibility would not change, their entrance age would stay the same (This assumes they are not redshirted.). For those children whose eligibility would change, their entrance age would change by 12 months. And although this policy change would change every child's relative age, it would do so differently for those who remain eligible versus those whose eligibility changes.

Our EPF parameter estimates only allow us to make inference about those children who are in our sample and whose eligibility would not be changed by a new entrance cutoff date. It is important to remember that our sample is only comprised of the oldest children in a given cohort. Furthermore, we cannot speak to the overall effects of such a policy change because our EPF parameters are not informative about the large, discontinuous changes in entrance and relative ages experienced by children whose eligibility would be altered.

We assume that changing the entrance cutoff date will not change any family inputs and will only change one school input, age relative to classmates, RA, for children in our sample. Returning to the EPF in Equation 14, asking how an earlier entrance cutoff date would affect an eligible child's achievement is the same as asking how decreasing a child's relative age impacts their achievement:

$$-\frac{\partial Y_{ijt}}{\partial RA} = -\beta_{t2} - \beta_{t3} EA. \tag{15}$$

The relative age parameter β_{t2} from the EPF determines the level of the effect from changing the entrance cutoff date, and the interaction parameter β_{t3} determines how much this varies at different entrance ages.

When interpreting these results, it is useful to consider possible mechanisms through which relative age functions as a school input. One such mechanism has lead to concerns about readiness. As a child becomes older relative to other children in their classroom, the child may receive additional inputs at school, perhaps in the form of addition attention or encouragement from their teacher or from tracking with a higher achieving peer group. Stimulated from the extra input, the child's achievement goes up. This mechanism could also work in reverse for children receiving fewer inputs due to being relatively younger than their classmates. If this mechanism were to dominate we would expect relatively older children to perform better on achievement tests, or for there to be a negative effect from decreasing a child's relative age.

A second mechanism through which relative age might serve as a school input is also related to peer effects. If a teacher teaches to the average level of achievement in their classroom, school inputs will be tailored towards the average age in a classroom. Inputs aimed at higher achieving peers could cause children to accelerate their learning to keep pace with their peers, raising their test scores.[8] In this case we would expect the effect of decreasing relative age to have a positive sign.

Table 11 shows the implied effects in the spring of third grade of an earlier entrance cutoff date based on our EPF parameter estimates at the median entrance age in our sample, 5 years and 7.00 months. The effects on all children of changing the entrance cutoff date are large, especially considering that this date typically varies across a four month interval. The estimates imply the average child would experience effects of 0.28 and 0.36 standard deviations on math and reading test scores if their state moved from the latest to the earliest cutoff date in our sample.

It is perhaps surprising that the effects of an earlier cutoff date are positive, because this indicates that children perform better when relatively *younger* than their classmates. This strong positive effect indicates the second mechanism through which relatively younger children speed up their learning to match the level of their peers dominates the harm done by the first mechanism. And while it might be surprising that children have higher achievement when relatively younger, similar results have also been reported in the literature. Elder and Lubotsky (2008) document that relatively younger children have higher achievement and Dobkin and Ferreira (2010) find that younger children have higher academic attainment.

The heterogeneity in effects by gender and home environment is remarkably strong. The effects on boys are much larger than the effects on girls. If it is surprising that children perform better

[8] Alternatively, it might be the case that if children cannot keep pace, such inputs might also leave children behind, decreasing their test scores.

when relatively younger than their classmates, it might be even more surprising to see this result for boys. Together with the well documented practice of redshirting boys more often than girls (Aliprantis (2011), Deming and Dynarski (2008)), this heterogeneity suggests a topic for future research to investigate is whether early childhood trends might be contributing to the gender gap in attainment (Goldin et al. (2006)).

Children with few books at home perform much worse when relatively younger than their classmates. Thinking about the two mechanisms discussed earlier, one hypothesis is that such children's family inputs work in favor of the first mechanism dominating the second mechanism. Children with low family inputs who are also relatively young may simply be too far behind to catch up with their peers.

5.2 How Would an Earlier Birthdate Affect an Eligible Child's Achievement?

An issue of interest for parents is understanding how their child being born at a different point in time would impact their achievement. This might be posed by parents as the question, "How would an earlier birthdate affect an eligible child's achievement?"

We can also use the estimated EPF parameters to investigate the effects of this counterfactual scenario. Inserting EA and RA from Equations 9 and 10 into the production function in Equation 14, we can see that all else constant, moving a child's birthday to an earlier date is the same as increasing both the child's entrance and relative ages. Thus an earlier birthdate results in a combination of entrance and relative age effects:

$$-\frac{\partial Y_t}{\partial \text{BDay}} = \beta_{t1} + \beta_{t2} - \beta_{t3}(EA + RA).$$

These are the net kindergarten entrance age effects typically identified in the literature.

Table 12 shows estimates of these effects evaluated at the median value of $EA + RA$ in our sample, 15. Since these effects are the combination of several parameters, the standard errors are so large we cannot make any meaningful statements about the point estimates. As a result we also estimate specifications of the EPF in Equation 14 in which there are not interaction terms between entrance and relative ages. The results from these regressions are shown in Table 13.

Several features of the effects in Table 13 merit consideration. First, while the effects of changing birthdate by one month are smaller than those of changing entrance cutoff dates by one month, the possible effects of changing birthdate are still large. Birthdates within any state in our sample vary by up to six months, and moving the average child's birthdate earlier by this amount would increase their third grade achievement in math and reading by 0.18 and 0.42 standard deviations.

Second, while the effects of changing a child's birthdate do exhibit heterogeneity by gender and home environment, these differences are much smaller than the heterogeneity in effects from changing entrance cutoff dates. Finally, the positive effects in Table 13 help to illustrate that entrance age parameters are typically of the opposite sign and of a larger magnitude than their corresponding relative age parameters.

In addition to the relative age mechanisms discussed in Section 5.2, it is now relevant to consider mechanisms related to entrance age. A first entrance age mechanism one might imagine is that increasing the age at which a child enters school increases the time they remain in the home or preschool environment. This mechanism can be seen as a change to the allocation of children's time (Fiorini and Keane (2011)). A second mechanism would be that children have more time for their biological development process to unravel. A large literature in child psychology has focused on understanding the relationship between these two mechanisms; that is, how children's environments and activities interact with the process of their biological development (Tudge et al. (2009)). The heterogeneity in effects by gender indicates that this biological development process is not uniform for all children, and the heterogeneity in effects by home environment is evidence the development process is not independent of children's environments.

5.3 Policy Discussion

The analysis conducted in this paper is ultimately aimed at helping to determine the optimal age for children to start school. There are at least three hurdles to assessing when states should set their entrance cutoff dates.

The first hurdle is that deciding when children should start school requires we understand how changing the entrance cutoff date would effect the achievement of an entire cohort of children. As discussed in Section 5.1, the changes to children's entrance and relative ages induced by changing the entrance cutoff date do not satisfy monotonicity assumptions (ie, The change in these ages is in different directions for different children.). The evidence presented here only pertains to the monotonic impacts for the oldest children in a cohort.

Although it should be possible to estimate the effects from different types of changes in entrance and relative ages, it must be noted that such effects, like the EPF parameters estimated in this paper, will typically not be structural in the sense of being policy invariant. The parameters in this analysis describe the technology producing educational achievement for a specific subset of children in the US entering kindergarten in the fall of 1998. It is entirely possible that the parameters of the production function have changed since then as the organization and behaviors of families, schools, and communities have themselves changed. In fact, one key reason for doing research such as that conducted in this paper is to inform policy with the goal of improving the technology of the production function. It seems unlikely that any of the estimated parameters represent inherent or unaltering relationships between age and achievement.

Finally, even if we did fully understand the achievement production technology at all ages and under all policies, we would still need to use some objective function in order to weigh the effects of changing entrance cutoff dates on achievement against the concurrent effects to society from changing child care costs (Cascio (2009), Fitzpatrick (2010)), attainment (Dobkin and Ferreira (2010), Angrist and Krueger (1991)), and foregone labor force participation and government revenue (Deming and Dynarski (2008)).

These hurdles make it difficult to imagine a single answer to the question, "When should children start school?" Furthermore, the evidence of strong heterogeneity presented in this analysis indicates that from a policy perspective, answering this question may be of less importance than answering "How can schools best meet the diverse needs of new students?" These questions need not be independent of each other, but children will continue to enter school with large age differences no matter when the entrance cutoff date is set.

6 Conclusion

This paper presented evidence that redshirting renders instrumental variables frameworks unable to identify age effects for the youngest children of a given cohort. Statistical tests on the nature of the redshirting decision yielded evidence against strong ignorability and in favor of essential heterogeneity. Identification when entrance and relative ages are determined in a setting of essential heterogeneity requires a monotonicity assumption on the impact of assigned relative age, and it has been shown that this assumption is not met in the presence of redshirting (Aliprantis (2011), Barua and Lang (2009)).

Despite this negative result, the paper showed that the relatively oldest children in a given cohort are largely unaffected by by redshirting. This fact was used as the basis for an identification strategy using quasi-random variation in ages to separately estimate entrance and relative age effects. Both entrance and relative ages were specified to be inputs into education production functions, and the separate identification of these production parameters allowed for the consideration of specific policy counterfactuals.

One of the most surprising results in the paper was that a state moving its entrance cutoff to an earlier date improves the scores of children in our sample. This indicates that on average younger children catch up to older peers rather than being left behind them. This result holds on the sample of relatively oldest children in a cohort, so it cannot be extended to relatively older children. However, such a result indicates that relative and entrance ages are likely to have complicated relationships with achievement. The extreme heterogeneity in effects by gender and home environment found in this paper further support such a conclusion.

References

Aliprantis, D. (2011). Redshirting, compulsory schooling laws, and educational attainment. *Journal of Educational and Behavioral Statistics*. Forthcoming.

Aliprantis, D. (2012). Assessing the evidence on neighborhood effects from Moving to Opportunity. *Federal Reserve Bank of Cleveland Working Paper 11-22r*.

Almond, D. and J. Currie (2010). Human capital development before age five. *NBER Working Paper 15827*.

Andrabi, T., J. Das, A. I. Khwaja, and T. Zajonc (2011). Do value-added estimates add value? Accounting for learning dynamics. *American Economic Journal: Applied Economics 3*(3), 29–54.

Angrist, J. D. and A. B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics 106*(4), 979–1014.

Barua, R. and K. Lang (2009). School entry, educational attainment and quarter of birth: A cautionary tale of LATE. *NBER Working Paper 15236*.

Bedard, K. and E. Dhuey (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics 121*(4), 1437–1472.

Björklund, A. and R. Moffitt (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics 69*(1), pp. 42–49.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011). Too young to leave the nest? The effects of school starting age. *The Review of Economics and Statistics 93*(2), 455–467.

Bobak, M. and A. Gjonca (2001). The seasonality of live birth is strongly influenced by socio-demographic factors. *Human Reproduction 16*(7), 1512–1517.

Bogard, K. and T. Berkley (2011). Understanding fade out from the early years to the early grades. *Manuscript, New York University*.

Bogard, K. and R. Takanishi (2005). PK-3: An aligned and coordinated approach to education for children 3 to 8 years old. *Social Policy Report 19*(3), 3–23.

Bound, J. and D. A. Jaeger (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In S. W. Polachek (Ed.), *Worker Well Being*, Volume 19, pp. 83–108. Research in Labor Economics.

Buckles, K. and D. M. Hungerman (2008). Season of birth and later outcomes: Old questions, new answers. *NBER Working Paper 14573*.

Cameron, S. V. and J. J. Heckman (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy 109*(3), 455–499.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–2781.

Cascio, E. and D. W. Schanzenbach (2007). First in the class? Age and the education production function. *NBER Working Paper 13663*.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into American public schools. *The Journal of Human Resources 44*(1), 140–170.

Cunha, F. and J. J. Heckman (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2005). Interpreting the evidence on life cycle skill formation. *IZA Discussion Paper 1675*.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives 15*(2), 213238.

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review 25*, 43–62.

Deming, D. and S. Dynarski (2008). The lengthening of childhood. *Journal of Economic Perspectives 22*(3), 71–92.

Dhuey, E. and S. Lipscomb (2008). What makes a leader? Relative age and high school leadership. *Economics of Education Review 27*(2), 173–183.

Dickert-Conlin, S. and A. Chandra (1999). Taxes and the timing of births. *Journal of Political Economy 107*(1), 161–177.

Dickert-Conlin, S. and T. Elder (2009). Suburban legend: School cutoff dates and the timing of births. *Mimeo., Michigan State University*.

Dobkin, C. and F. Ferreira (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review 29*(1), 40–54.

Eckstein, Z. and K. I. Wolpin (1999). Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica 67*(6), 1295–1339.

Edwards, B., M. Fiorini, and M. Taylor (2011). Who gets the 'gift of time' in Australia? Exploring delayed primary school entry. *Australian Review of Public Affairs 10*(1), 41–60.

Elder, T. E. and D. H. Lubotsky (2008). Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers. *Journal of Human Resources*.

Fiorini, M. and M. Keane (2011). How the allocation of children's time affects cognitive and non-cognitive development. *Mimeo., University of Technology Syndey*.

Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten. *Journal of Labor Economics 28*(1), 51–85.

Fredriksson, P. and B. Öckert (2005). Is early learning really more productive? The effect of school starting age on school and labor market performance. *IZA Discussion Paper No. 1659*.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives 20*(4), 133–156.

Graue, M. E. (1993). *Ready for What? Constructing Meanings of Readiness for Kindergarten.* State University of New York.

Graue, M. E. and J. DiPierna (2000). Redshirting and early retention: Who gets the "gift of time" and what are its outcomes? *American Educational Research Journal 37*(2), 509–534.

Heckman, J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010a). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics 1*(1), 1–46.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010b). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94*(1-2), 114 – 128.

Heckman, J. J., D. Schmierer, and S. Urzua (2010). Testing the correlated random coefficient model. *Journal of Econometrics 158*(2), 177 – 203. Specification Analysis in Honor of Phoebus J. Dhrymes.

Heckman, J. J., S. Urzúa, and E. Vytlacil (2006). Understanding Instrumental Variables in models with essential heterogeneity. *The Review of Economics and Statistics 88*(3), 389–432.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–475.

Jacob, B. A., L. Lefgren, and D. P. Sims (2010). The persistence of teacher-induced learning. *Journal of Human Resources 45*(4), 915–943.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

Keane, M. P. and K. I. Wolpin (2000). Eliminating race differences in school attainment and labor market success. *Journal of Labor Economics 18*(4), 614–652.

Lam, D. A. and J. A. Miron (1996). The effects of temperature on human fertility. *Demography 33*(3), 291–305.

Lam, D. A., J. A. Miron, and A. Riley (1994). Modeling seasonality in fecundability, conceptions, and births. *Demography 31*(2), 321–346.

Magnuson, K. A., C. Ruhm, and J. Waldfogel (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review 26*(1), 33–51.

McEwan, P. J. and J. S. Shapiro (2008). The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates. *Journal of Human Resources 43*(1), 1–29.

Minnesota Population Center (2004). *National Historical Geographic Information System* (Prerelease Version 0.1 ed.). Minneapolis, MN: University of Minnesota. http://www.nhgis.org.

Muscha, J. and S. Grondin (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review 21*(2), 147–167.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*(5), 869–895.

Puhani, P. A. and A. M. Weber (2005). Does the early bird catch the worm? Instrumental Variable estimates of educational effects of age of school entry in Germany. *IZA Discussion Paper No. 1827*.

Rizzi, E. L. and G. Dalla-Zuanna (2007). The seasonality of conception. *Demography 44*(4), 705–728.

Rolnick, A. and R. Grunewald (2003). *Early Childhood Development: Economic Development with a High Public Return*. Federal Reserve Bank of Minneapolis.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics 125*(1), 175–214.

Stipek, D. (2002). At what age should children enter kindergarten? A question for policy makers and parents. *Social Policy Report XVI*(2), 1–16.

Todd, P. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal 113*(485), F3–F33.

Todd, P. and K. I. Wolpin (2007). The production of cognitive achievement in children: Home, school and racial test score gaps. *Journal of Human Capital 1*(1), 91–136.

Tudge, J. R. H., L. B. L. Freitas, and F. Doucet (2009). The transition to school: Reflections frmo a contextualist perspective. In H. Daniels, H. Lauder, and J. Porter (Eds.), *Educational Theories, Cultures, and Learning: A Critical Perspective*, pp. 117–133. London: Routledge.

Weil, E. (2007, June 3). When should a kid start kindergarten? *The New York Times Magazine*.

Figures

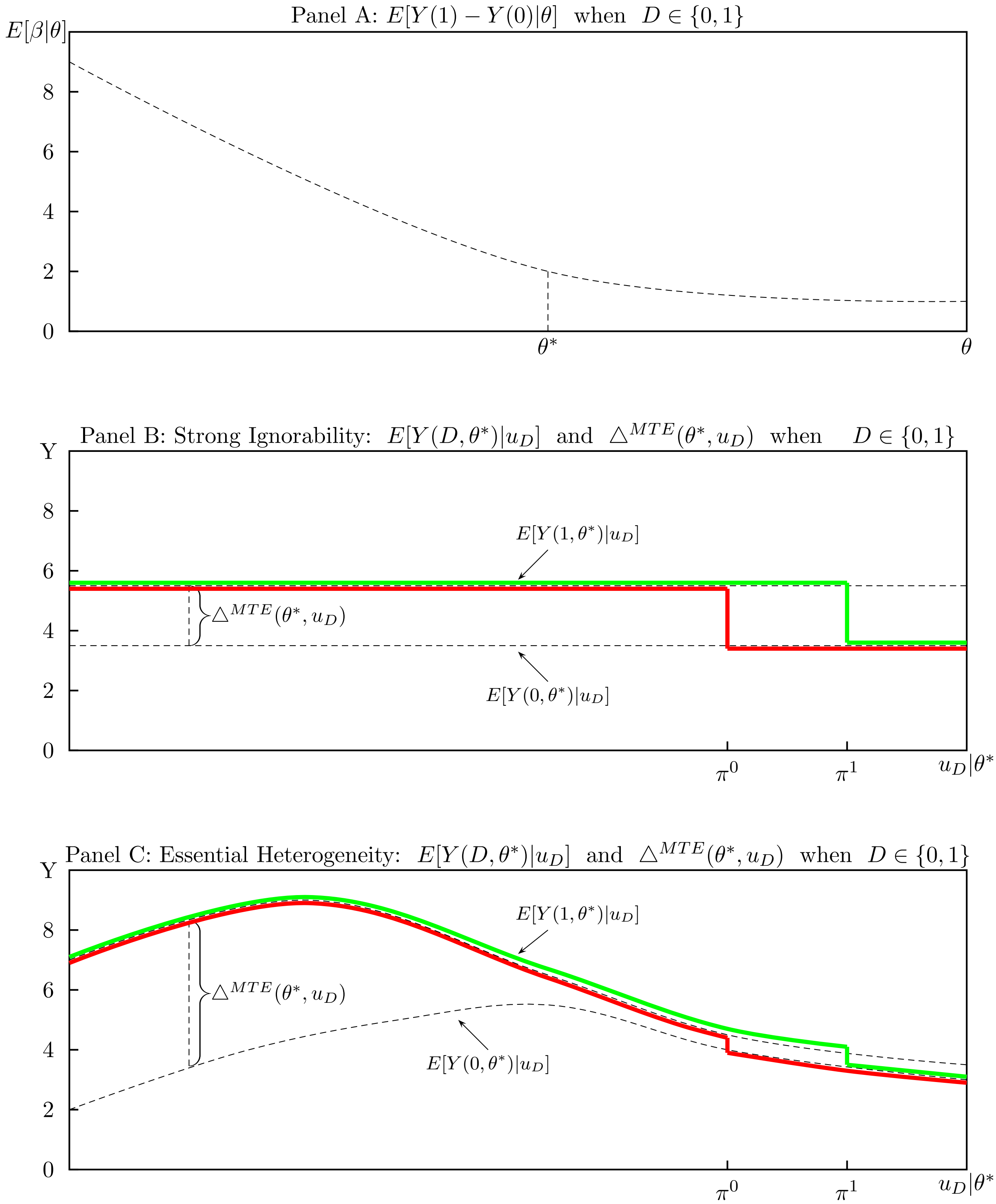


Figure 1: Example I: Strong Ignorability and Essential Heterogeneity

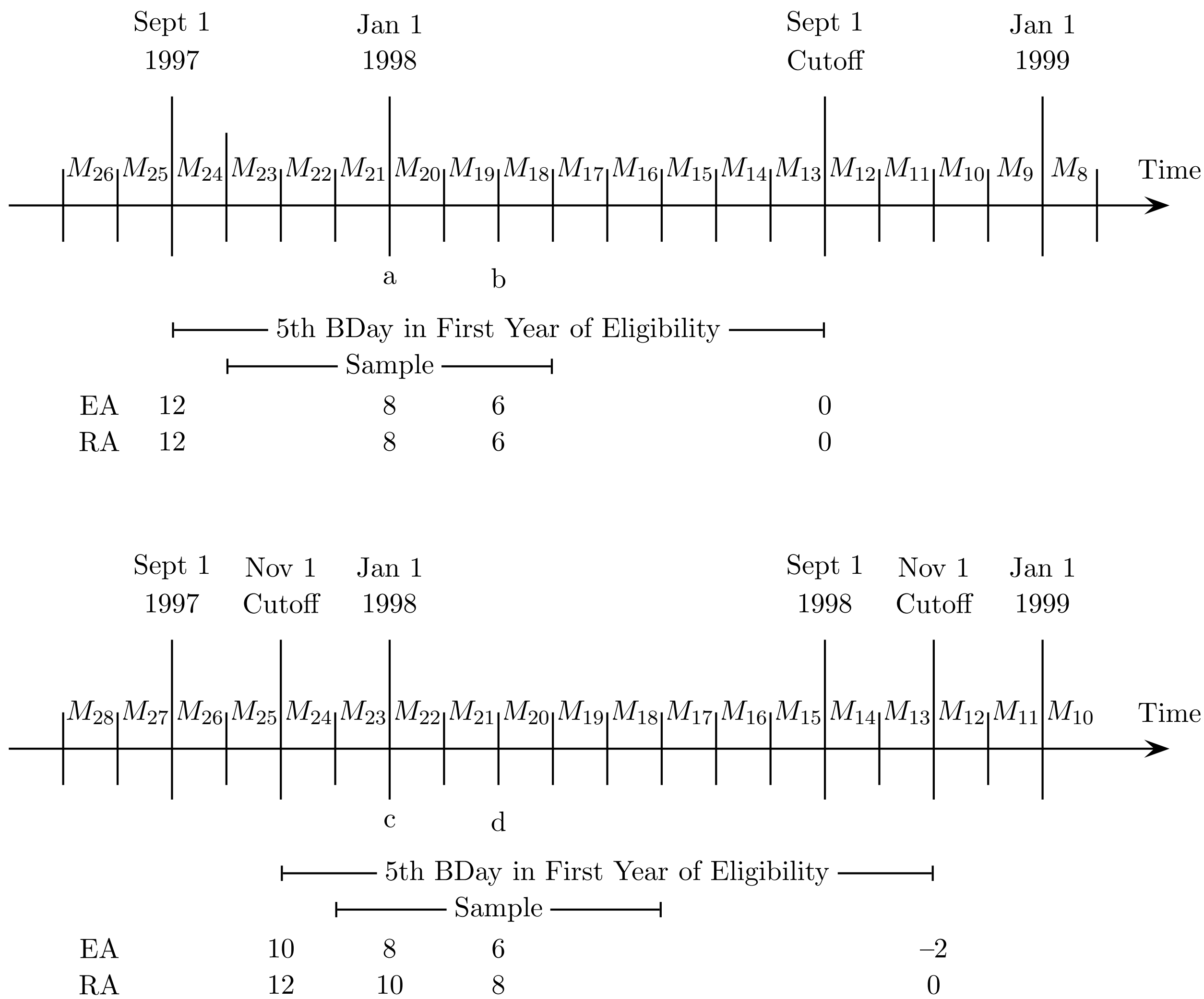


Figure 2: Variation in Entrance Age and Relative Age

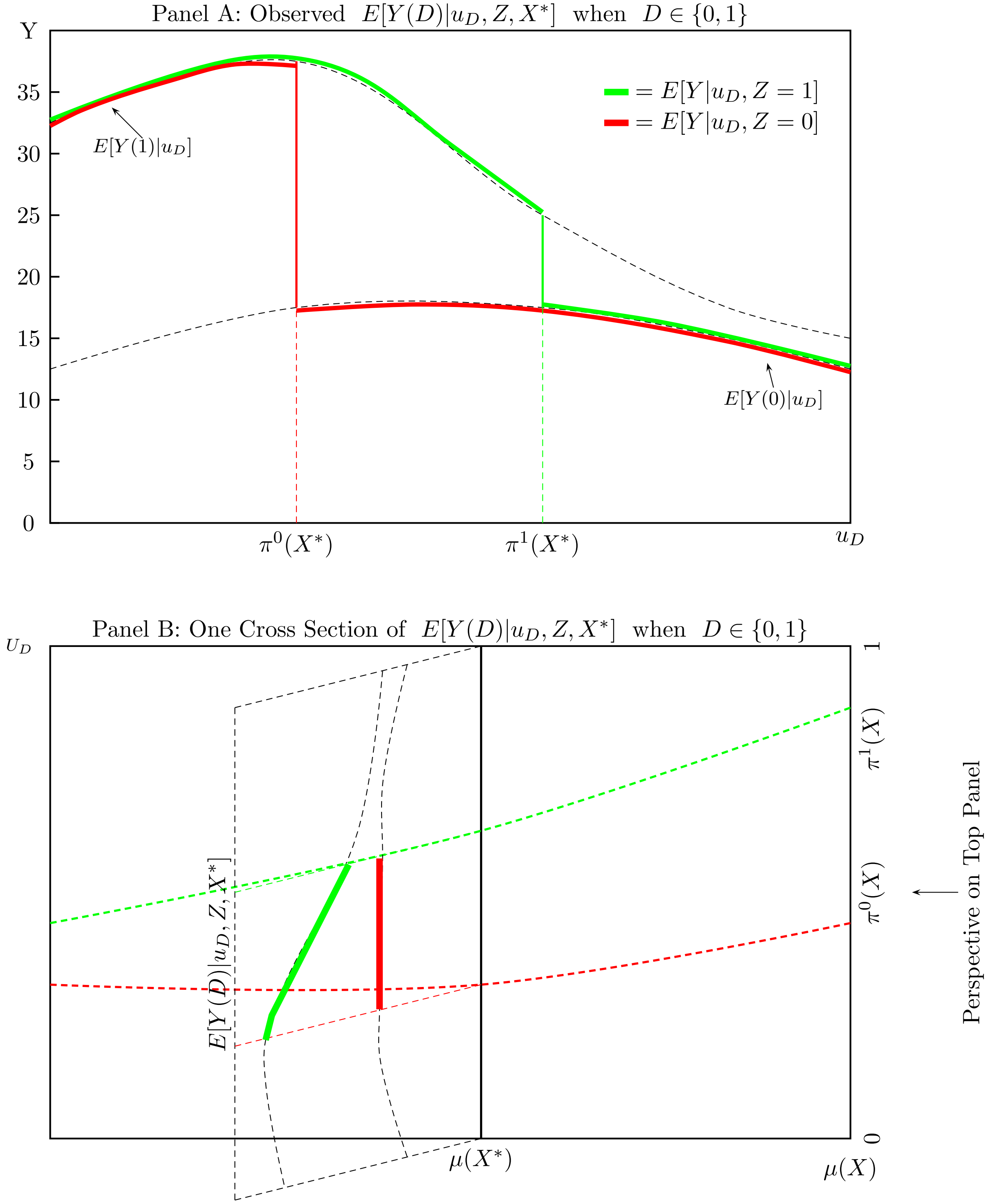


Figure 3: Example II: Potential Outcomes and Marginal Treatment Effects Given $\mu(X)$

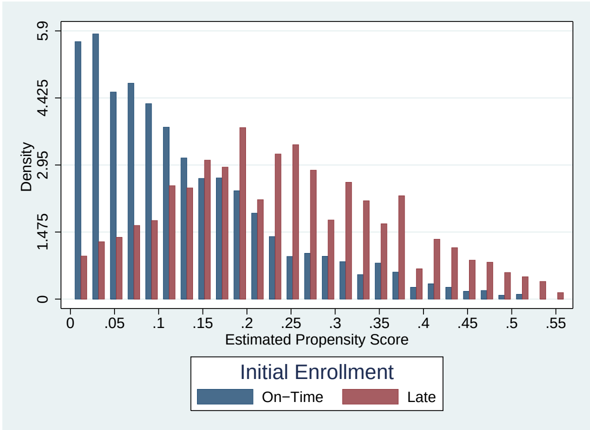


Figure 4: Distribution of Estimated Propensity Scores (By Initial Enrollment Decision)
Note: Propensity scores are predicted using the parameters of a probit model estimated on assigned relative age and its square, the number of children's books at home, whether the child ever received WIC benefits, and whether the child's mother is present and works 35 hours or more as measured in the spring of kindergarten; the SES quintile of the child and the mother's education level measured in the fall of kindergarten; whether the father was present in the household during the fall of first grade; and gender and race dummies.

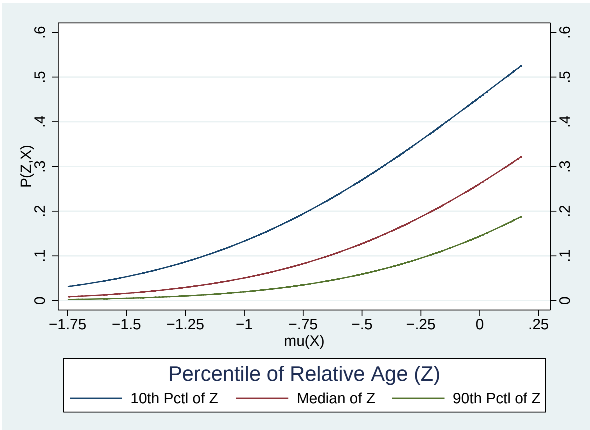


Figure 5: Predicted Probability of Redshirting (Conditional on $\mu(X)$ and Percentiles of Z)

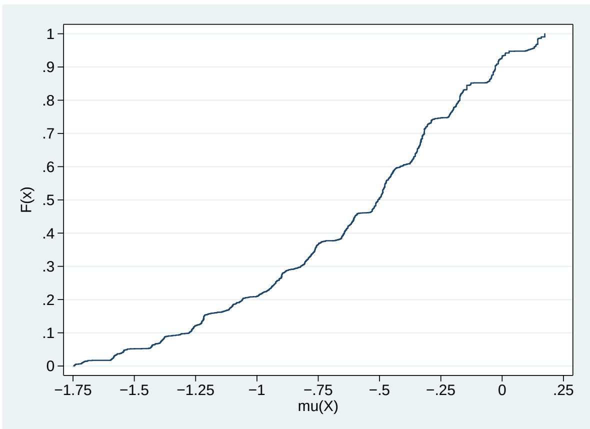
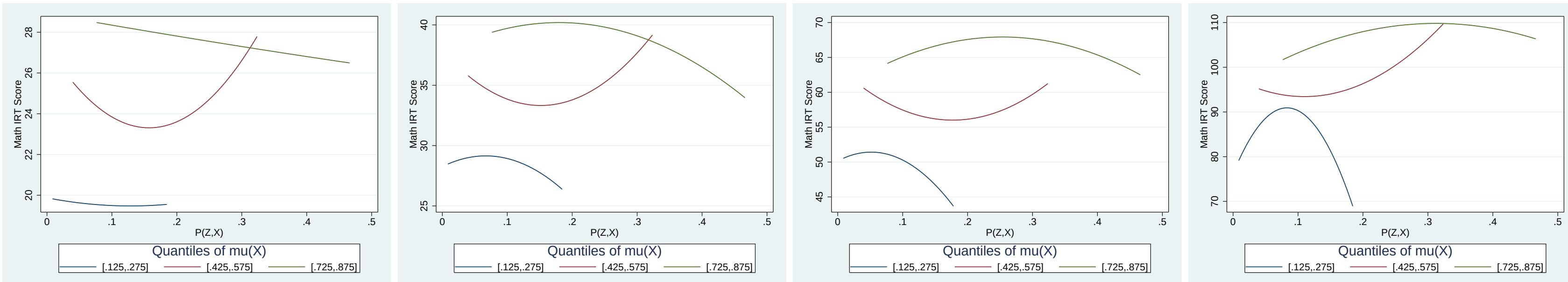
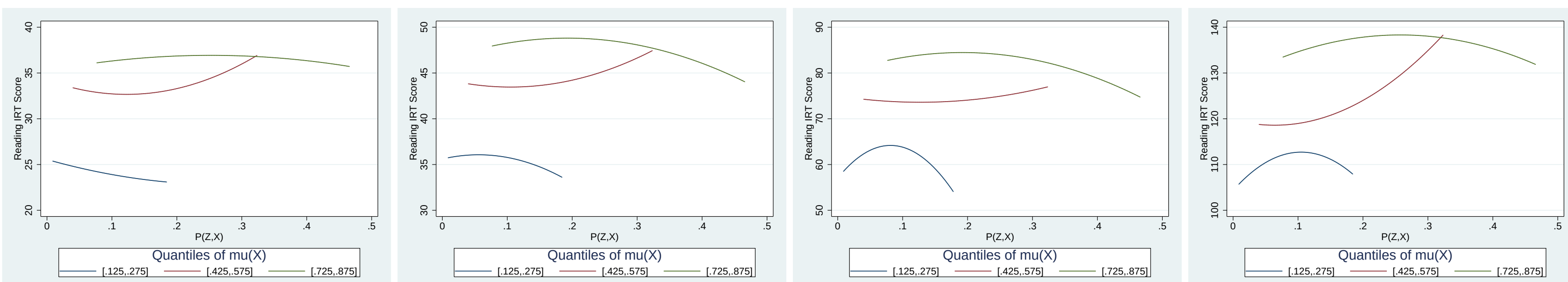


Figure 6: CDF of Observables (Using the Summary Measure $\mu(X)$)



(a) Fall K (b) Spring K (c) Spring 1st (d) Spring 3rd

Figure 7: Quadratic Model Prediction of $E[Y(P(z,x)) \mid \mu(x)]$, Math



(a) Fall K (b) Spring K (c) Spring 1st (d) Spring 3rd

Figure 8: Quadratic Model Prediction of $E[Y(P(z,x)) \mid \mu(x)]$, Reading

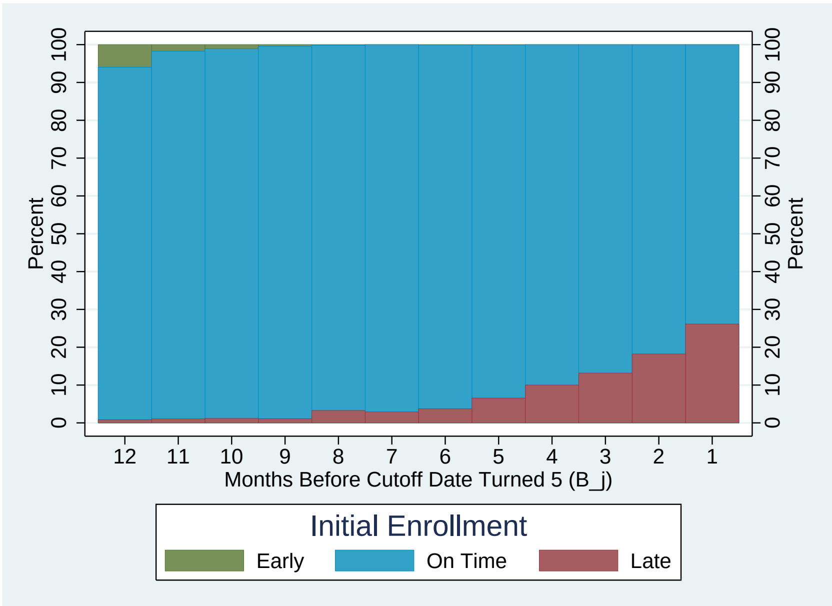
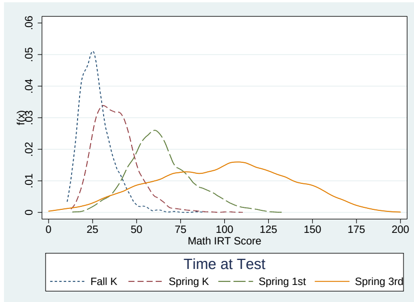
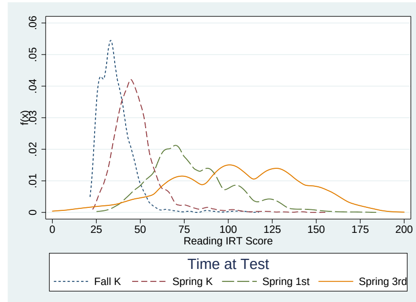


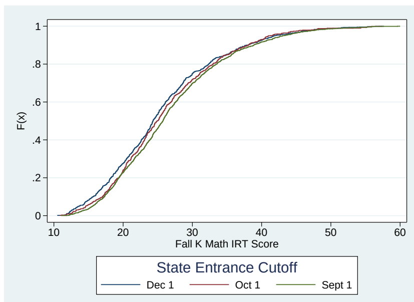
Figure 9: Initial Enrollment Decisions (By Assigned Relative Age Cohorts B_j)



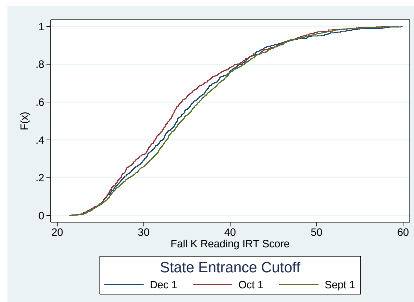
(a) PDFs of Math IRT Scores



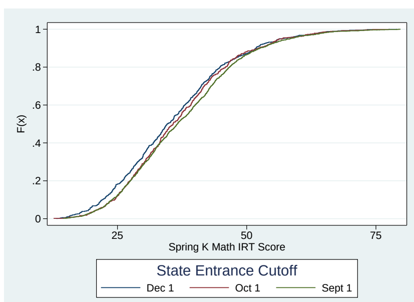
(b) PDFs Reading IRT Scores



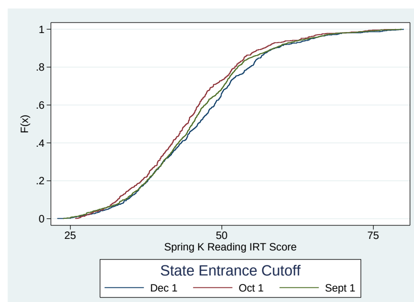
(c) CDFs of Math in Fall K



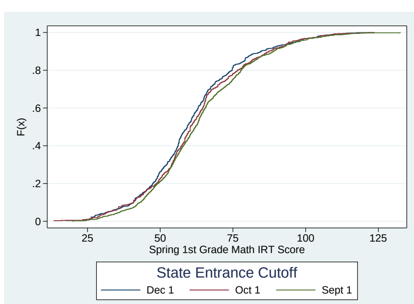
(d) CDFs of Reading in Fall K



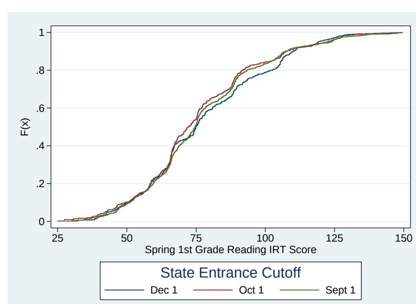
(e) CDFs of Math in Spring K



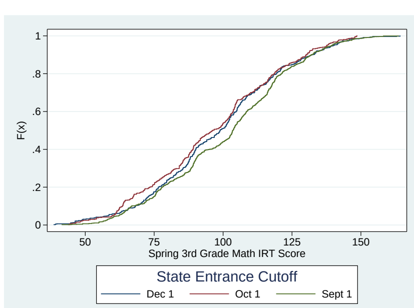
(f) CDFs of Reading in Spring K



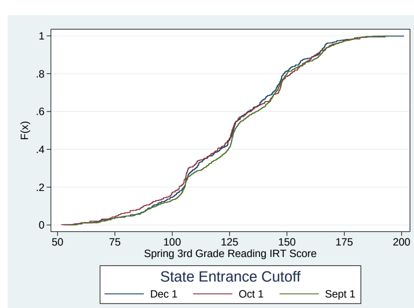
(g) CDFs of Math in Spring 1st



(h) CDFs of Reading in Spring 1st

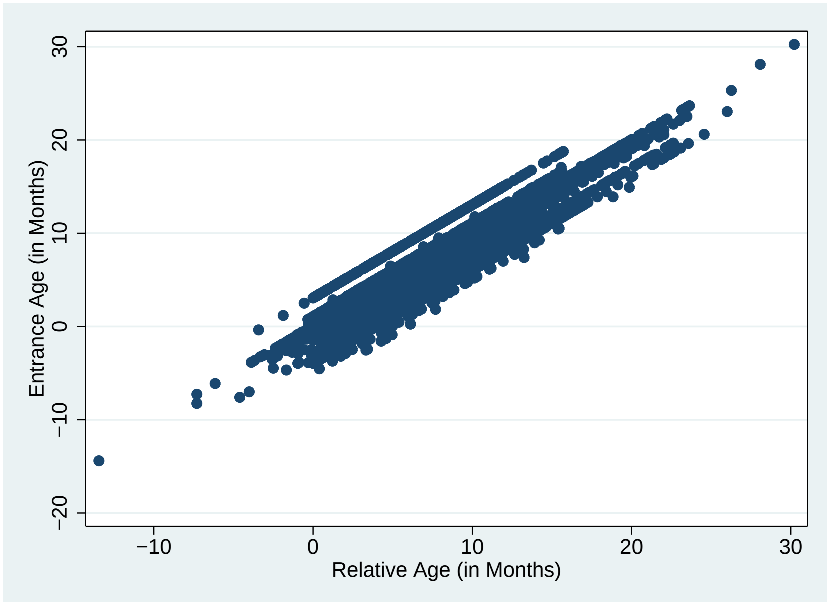


(i) CDFs of Math in Spring 3rd

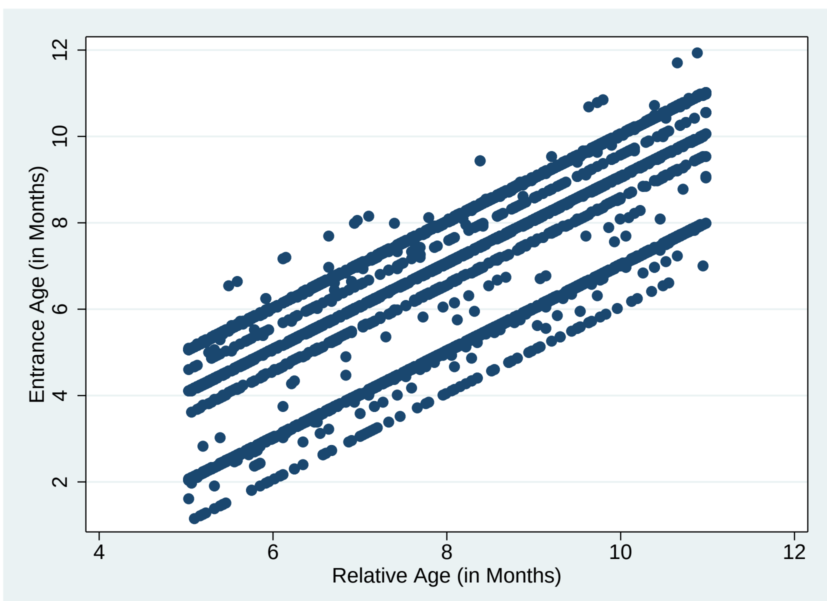


(j) CDFs of Reading in Spring 3rd

Figure 10: The Distributions of Math and Reading Test Scores

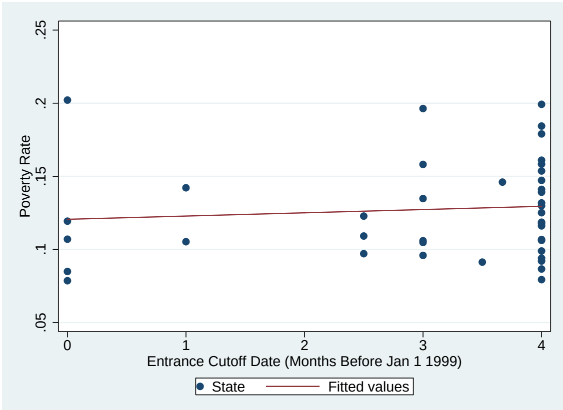


(a) Entrance Age and Relative Age in the Entire ECLS-K

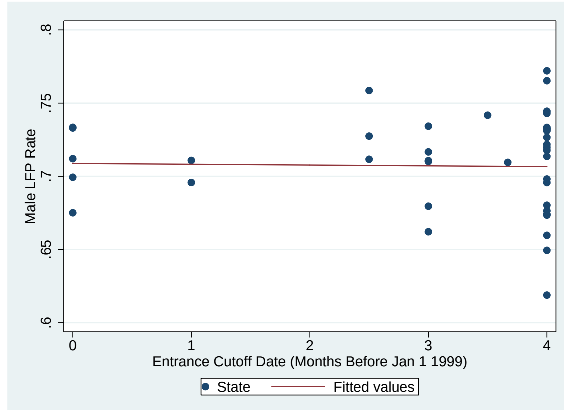


(b) Entrance Age and Relative Age in the Sample

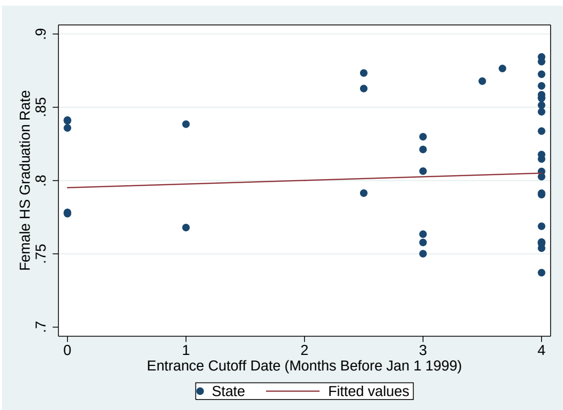
Figure 11: Variation in Entrance Age and Relative Age



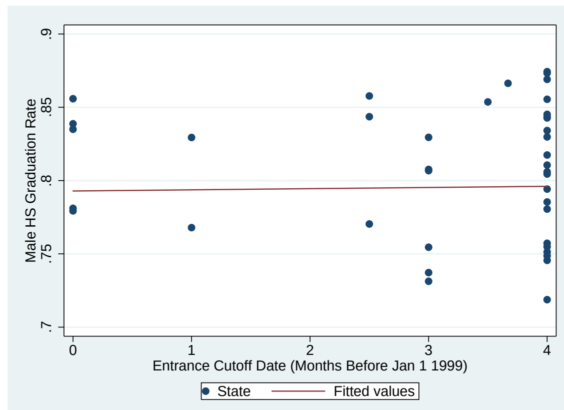
(a) Poverty and Entrance Cutoff Date



(b) Male LFP and Entrance Cutoff Date

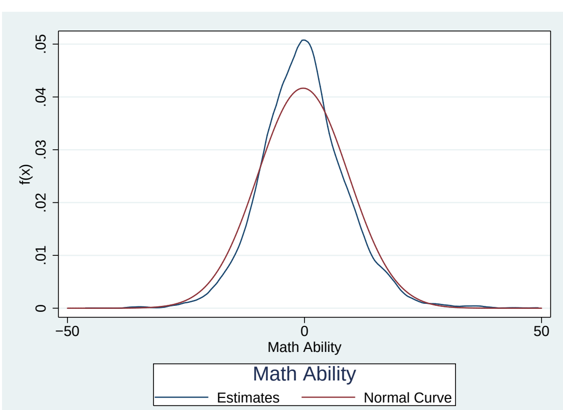


(c) Female HS Grad Rate and Entrance Cutoff Date

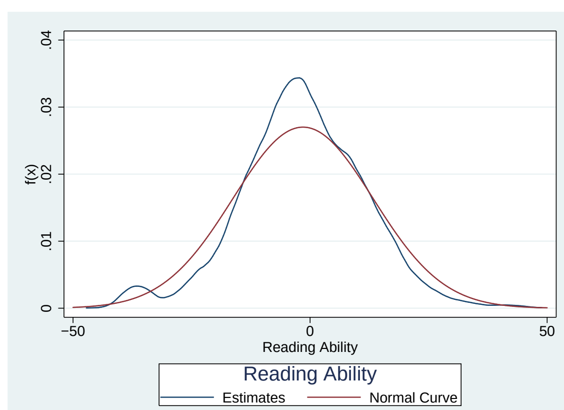


(d) Male HS Grad Rate and Entrance Cutoff Date

Figure 12: State Characteristics and 1998 Kindergarten Entrance Cutoff Dates
Note: State level data are from the US Census/NHGIS, and entrance cutoff dates are taken from Elder and Lubotsky (2008).



(a) Math Ability Estimates



(b) Reading Ability Estimates

Figure 13: The Distributions of Math and Reading Ability Estimates
Note: Ability estimates obtained via 2SLS using race as an instrument for family inputs. Family inputs included in estimation are mother's educational attainment and household income. Outcomes are IRT test scores in the fall of kindergarten.

Tables

Table 1: Testing $H_0^{IV}: \beta_{IV,[\underline{p},p_{med}]} = \beta_{IV,[p_{med},\overline{p}]}$

(a) Math

Observables	Fall K			Spring K			Spring 1st			Spring 3rd		
	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$
$\mu(X) \in [Q_{\tau=.125}, Q_{\tau=.275})$	0.84	-11.7	-5.2	0.56	-12.3	-40.5	0.27	31.1	-60.4	**0.00**	117.4	-231.5
		(16.7)	(28.3)		(24.4)	(41.6)		(40.7)	(72.3)		(68.6)	(97.6)
$\mu(X) \in [Q_{\tau=.275}, Q_{\tau=.425})$	0.11	1.2	38.2	**0.01**	-15.3	60.6	0.29	8.4	65.7	0.66	71.8	36.3
		(16.0)	(17.1)		(20.2)	(21.9)		(39.2)	(40.2)		(54.0)	(62.4)
$\mu(X) \in [Q_{\tau=.425}, Q_{\tau=.575})$	**0.02**	-24.3	26.6	0.07	-9.5	47.0	**0.01**	-104.8	28.0	**0.02**	-88.4	86.7
		(19.3)	(11.8)		(26.3)	(16.2)		(45.7)	(27.6)		(63.5)	(39.9)
$\mu(X) \in [Q_{\tau=.575}, Q_{\tau=.725})$	0.33	-26.4	-5.6	0.91	-1.7	1.1	0.13	-52.1	15.7	0.97	-32.9	-30.5
		(19.5)	(9.1)		(23.5)	(11.0)		(40.9)	(18.3)		(64.7)	(27.6)
$\mu(X) \in [Q_{\tau=.725}, Q_{\tau=.875})$	0.87	-2.1	1.2	**0.02**	74.1	-1.9	**0.01**	139.4	15.4	0.16	111.2	18.0
		(19.7)	(7.8)		(29.1)	(11.6)		(45.3)	(17.9)		(61.4)	(25.0)

(b) Reading

Observables	Fall K			Spring K			Spring 1st			Spring 3rd		
	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$	p-Value	$\widehat{\beta}_B$	$\widehat{\beta}_A$
$\mu(X) \in [Q_{\tau=.125}, Q_{\tau=.275})$	0.85	-11.0	-23.2	0.42	-45.7	-116.3	0.16	51.6	-139.5	0.70	23.2	-39.0
		(33.7)	(57.3)		(44.3)	(75.4)		(67.6)	(120.1)		(93.1)	(132.5)
$\mu(X) \in [Q_{\tau=.275}, Q_{\tau=.425})$	0.78	-1.5	7.5	0.36	-27.9	11.1	0.25	-27.1	58.4	0.22	40.6	-84.4
		(22.1)	(23.6)		(29.2)	(31.5)		(53.3)	(54.6)		(68.4)	(78.9)
$\mu(X) \in [Q_{\tau=.425}, Q_{\tau=.575})$	0.24	4.5	35.2	0.45	10.2	38.9	0.10	-83.5	38.3	0.10	-110.3	47.0
		(22.2)	(13.6)		(32.3)	(19.9)		(64.1)	(38.7)		(82.8)	(52.0)
$\mu(X) \in [Q_{\tau=.575}, Q_{\tau=.725})$	0.63	24.0	10.0	0.56	-7.9	12.6	0.84	47.2	33.1	0.21	109.5	5.6
		(26.7)	(12.5)		(32.4)	(15.2)		(65.7)	(29.5)		(76.2)	(32.5)
$\mu(X) \in [Q_{\tau=.725}, Q_{\tau=.875})$	0.33	31.7	10.7	**0.02**	77.9	7.4	0.28	72.9	7.9	0.17	115.4	17.2
		(20.1)	(8.0)		(28.8)	(11.5)		(56.0)	(22.1)		(66.0)	(26.9)

Note: $\beta_B = \beta_{IV,[\underline{p}_B,\overline{p}_B]}$ where $\underline{p}_B$ is the lowest value of the propensity score conditional on observables θ and $\overline{p}_B$ is the midpoint of the support of the propensity score conditional on θ. β_A is defined analogously for the values of the propensity score above the midpoint. Instruments in Z include assigned relative age and its square. Covariates in X also used to estimate the propensity score include the number of children's books at home, whether the child ever received WIC benefits, and whether the child's mother is present and works 35 hours or more as measured in the spring of kindergarten; the SES quintile of the child and the mother's education level measured in the fall of kindergarten; whether the father was present in the household during the fall of first grade; and gender and race dummies. IV parameters are estimated by regressing test scores on predicted propensity scores and covariates X on the subsample of kindergarteners in relative age groups $\{M_{13}, M_{14}, M_{15}, M_{25}, M_{26}, M_{27}\}$.

Table 2: Entering Rates by Cohort in the ECLS-K

(a) All Children (By Month)

	Month Before Cutoff Turned 5											
	12 (%)	11 (%)	10 (%)	9 (%)	8 (%)	7 (%)	6 (%)	5 (%)	4 (%)	3 (%)	2 (%)	1 (%)
Entering:												
Early	6	2	1	0	0	0	0	0	0	0	0	0
On-Time	93	97	98	98	97	97	96	93	90	87	81	74
Waiting	1	1	1	1	3	3	4	6	11	14	19	26

(b) All Children (By Quarter)

	Quarter Before Cutoff Turned 5			
	4 (%)	3 (%)	2 (%)	1 (%)
Entering:				
Early	3	0	0	0
On-Time	96	97	93	80
Waiting	1	3	7	20

Table 3: Attrition in the ECLS-K

(a) Math IRT Score Present

	Sample		All ECLS-K	
Round	Attrition (%)	n	Attrition (%)	n
Fall K	0	3,950	5	18,640
Spring K	3	3,850	0	19,650
Fall 1st	–	–	–	–
Spring 1st	17	3,270	15	16,630
Spring 3rd	28	2,830	27	14,370
Spring 5th	44	2,230	43	11,270
Spring 8th	54	1,830	53	9,290

(b) Reading IRT Score Present

	Sample		All ECLS-K	
Round	Attrition (%)	n	Attrition (%)	n
Fall K	0	3,740	7	17620
Spring K	1	3,710	0	18,940
Fall 1st	–	–	–	–
Spring 1st	14	3,230	14	16,340
Spring 3rd	25	2,820	25	14,280
Spring 5th	41	2,220	41	11,270
Spring 8th	51	1,820	51	9,230

Table 4: Entrance and Relative Ages as a Function of Birthdate and State Cutoff

Birthdate/State Combination	EA	RA
a	8	8
b	6	6
c	8	10
d	6	8

Table 5: P-Values of F-Tests of Equality of Characteristic Means by Quarter and Month

	All Children		Sample	
Characteristic	Quarter	Month	Quarter	Month
Race				
Black	0.98	0.07	0.00	0.00
White	0.41	0.42	0.00	0.00
Hispanic	0.44	0.79	0.00	0.00
Asian	0.38	0.91	0.03	0.22
Gender				
Female	0.71	0.18	0.89	0.92
Mom's Education Level				
Mom's HGC < 12	0.03	0.31	0.73	0.09
Mom's HDR $=$ HS Diploma	0.42	0.61	0.33	0.64
Mom's HDR $\geq$ BA	0.81	0.82	0.63	0.40
Home Characteristics				
# of Books at Home	0.30	0.31	0.99	0.52
Live with Father	0.65	0.91	0.00	0.04
Birth Characteristics				
Parents Married at Birth	0.18	0.37	0.62	0.26
Birth Weight	0.30	0.07	0.11	0.12
Socio-Economic Status				
SES=1	0.03	0.08	0.47	0.11
SES=2	0.54	0.69	0.75	0.15
SES=3	0.96	0.73	0.39	0.48
SES=4	0.64	0.91	0.78	0.19
SES=5	0.14	0.59	0.08	0.11

Table 6: Parameter Estimates for All Children

(a) Math Test Scores

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	1.36	0.11	5.83	0.32	4.30	0.18
	(0.48)**		(0.76)**		(1.39)**	
Relative Age ($\widehat{\beta}_{t2}$)	0.39	0.03	–4.76	–0.26	0.03	0.00
	(0.33)		(0.81)**		(1.14)	
n	3,220		2,780		2,410	

(b) Reading Test Scores

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	1.87	0.14	4.67	0.20	4.73	0.17
	(0.57)**		(1.16)**		(1.99)*	
Relative Age ($\widehat{\beta}_{t2}$)	–1.75	–0.13	–2.56	–0.11	–2.00	–0.07
	(0.46)**		(1.13)*		(1.55)	
n	3,100		2,740		2,400	

Outcomes are IRT test scores. Home inputs included in estimation are income, mother's education, whether a foreign language is spoken at home, whether enrolled in a pre-kindergarten program, mother's employment between birth and kindergarten, whether the father was living in the household, and whether the child ever received WIC benefits. School inputs include entrance age, relative age, the interaction of the two ages, school and state fixed effects, gender, and race. Ability estimates as described in Sections 4.2 and 5 are also included in estimation.

$^{*}p < 0.05$.

$^{**}p < 0.01$.

Table 7: Math Parameter Estimates By Gender

(a) Math Test Scores (Boys)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	2.29 (.83)**	0.19	4.17 (1.30)**	0.23	12.22 (2.41)**	0.50
Relative Age ($\widehat{\beta}_{t2}$)	.35 (.54)	0.03	–1.25 (1.96)	–0.07	–5.01 (2.22)*	–0.21
n	1,620		1,390		1,200	

(b) Math Test Scores (Girls)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	0.25 (.90)	0.02	1.03 (1.42)	0.06	1.19 (2.58)	0.05
Relative Age ($\widehat{\beta}_{t2}$)	1.55 (.67)*	0.13	0.75 (1.39)	0.04	2.30 (1.61)	0.09
n	1,600		1,390		1,210	

Table 8: Reading Parameter Estimates By Gender

(a) Reading Test Scores (Boys)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	1.66 (.83)*	0.12	–0.88 (1.92)	–0.04	13.62 (2.87)**	0.50
Relative Age ($\widehat{\beta}_{t2}$)	–1.14 (.55)*	–0.08	5.64 (2.34)*	0.24	-5.92 (3.63)	–0.22
n	1,560		1,370		1,190	

(b) Reading Test Scores (Girls)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	3.13 (1.09)**	0.23	5.27 (1.91)**	0.23	1.48 (3.73)	0.05
Relative Age ($\widehat{\beta}_{t2}$)	-3.24 (.89)**	–0.24	-4.90 (1.98)*	–0.21	–0.36 (2.65)	–0.01
n	1,550		1,370		1,210	

Table 9: Math Parameter Estimates By Books at Home

(a) Math Test Scores (Few Books ($\leq$ 35) at Home)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	–1.49	–0.12	–8.08	–0.44	-7.58	–0.31
	(.83)		(2.16)**		(4.01)	
Relative Age ($\widehat{\beta}_{t2}$)	2.18	0.18	7.11	0.39	8.62	0.35
	(.75)**		(1.48)**		(2.56)**	
n	980		830		700	

(b) Math Test Scores (Many Books ($\geq$ 100) at Home)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	2.47	0.21	7.89	0.43	6.32	0.26
	(1.00)*		(1.53)**		(2.47)*	
Relative Age ($\widehat{\beta}_{t2}$)	–1.92	–0.16	-5.10	–0.28	–2.70	–0.11
	(.74)**		(1.35)**		(2.13)	
n	1,190		1,050		920	

Table 10: Reading Parameter Estimates By Books at Home

(a) Reading Test Scores (Few Books ($\leq$ 35) at Home)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	-12.89	–0.95	-2.20	–0.09	-11.41	–0.42
	(1.80)**		(3.76)		(5.10)*	
Relative Age ($\widehat{\beta}_{t2}$)	15.22	1.12	4.83	0.21	17.05	0.63
	(1.34)**		(2.61)		(3.66)**	
n	880		790		700	

(b) Reading Test Scores (Many Books ($\geq$ 100) at Home)

	Spring K		Spring 1st		Spring 3rd	
	Raw	σ	Raw	σ	Raw	σ
Entrance Age ($\widehat{\beta}_{t1}$)	0.41	0.03	-13.78	–0.59	3.51	0.13
	(.95)		(3.60)**		(3.78)	
Relative Age ($\widehat{\beta}_{t2}$)	–1.00	–0.07	12.65	0.54	–0.65	–0.02
	(.75)		(2.30)**		(3.73)	
n	1,190		1,050		920	

Table 11: Counterfactual 1: Effects from Changing to an Earlier Entrance Cutoff Date
Spring of 3rd Grade ($-\frac{\partial Y_{ijt}}{\partial RA} = -\widehat{\beta}_{t2} - \widehat{\beta}_{t3} EA\big|_{EA=7}$)

	Math		Reading	
Sample:	Raw	σ	Raw	σ
All Children:				
All Children	1.68	0.07	2.40	0.09
	(0.53)		(0.98)	
Gender:				
Boys	7.87	0.32	8.47	0.31
	(2.11)		(3.35)	
Girls	–0.84	–0.03	0.30	0.01
	(2.11)		(4.57)	
Home Environment:				
Few Books	–8.56	–0.35	–15.70	–0.58
	(5.69)		(9.11)	
Many Books	4.15	0.17	1.42	0.05
	(0.76)		(6.85)	

Table 12: Counterfactual 2: Effects from an Earlier Birth Date
Spring of 3rd Grade ($-\frac{\partial Y_t}{\partial \text{BDay}} = \widehat{\beta}_{t1} + \widehat{\beta}_{t2} - \widehat{\beta}_{t3}(EA + RA)\big|_{EA+RA=15}$)

	Math		Reading	
Sample:	Raw	σ	Raw	σ
All Children:				
All Children	7.99	0.33	3.57	0.13
	(17.46)		(34.61)	
Gender:				
Boys	13.33	0.55	13.19	0.49
	(57.09)		(125.56)	
Girls	6.62	0.27	0.99	0.04
	(39.68)		(92.41)	
Home Environment:				
Few Books	1.15	0.05	8.53	0.31
	(91.32)		(174.9)	
Many Books	6.71	0.28	4.49	0.17
	(73.42)		(122.28)	

Table 13: Counterfactual 2 with No Interaction Term: Effects from an Earlier Birth Date
Spring of 3rd Grade ($-\frac{\partial Y_t}{\partial \text{BDay}} = \widehat{\beta}_{t1} + \widehat{\beta}_{t2}$)

	Math		Reading	
Sample:	Raw	σ	Raw	σ
All Children:				
All Children	0.66	0.03	1.87	0.07
	(0.09)		(0.16)	
Gender:				
Boys	1.05	0.04	2.19	0.08
	(0.26)		(0.46)	
Girls	0.35	0.01	1.25	0.05
	(0.22)		(0.41)	
Home Environment:				
Few Books	0.91	0.04	2.70	0.10
	(0.43)		(0.99)	
Many Books	0.70	0.03	1.22	0.05
	(0.52)		(0.46)	