ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 20-203

Impacts of School Entry Age on Academic Growth through 2nd Grade: A Multi-State Regression Discontinuity Analysis

Angela Johnson
NWEA

Megan Kuhfeld
NWEA

The belief that additional time allows children to become more ready for school has affected public policy and individual practices. Prior studies estimated either associations between school entry age and academic growth or causal effects on achievement measured at one or two points. This paper contributes novel causal evidence for the impacts of kindergarten entry age on academic growth in the first three years of school. We embed regression discontinuity into a piecewise multilevel growth model and apply it to rich assessment data from three states. Being a year older leads to higher initial achievement and higher kindergarten growth rates but lower growth rates during 1st and 2nd grades. Effects do not differ by gender or race.

VERSION: February 2020

Suggested citation: Johnson, Angela, and Megan Kuhfeld. (2020). Impacts of School Entry Age on Academic Growth through 2nd Grade: A Multi-State Regression Discontinuity Analysis. (EdWorkingPaper: 20-203). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/d4kt-nv59

Impacts of School Entry Age on Academic Growth through 2nd Grade:

A Multi-State Regression Discontinuity Analysis

The belief that additional time allows children to become more ready for school has affected both public policy and individual practices. Many states have shifted cutoff dates to increase the age at which children are permitted to enroll (Education Commission of the States, 2018; Elder & Lubotsky, 2009). Parents are also choosing to keep their children who have reached the legal age of school entry out of school for one or more years, a practice commonly known as academic redshirting (Bassok & Reardon, 2013). Unsurprisingly, the age at which children enroll in school has steadily increased (Deming & Dynarski, 2008).

The prevailing assumption behind raising the age of school entry, whether through policy changes or family choice, is that being older and more mature helps children derive more benefit from schooling (Meisels, 1999; Morrison, Alberts, & Griffith, 1997). Researchers posit that older entry age carries absolute and relative advantages (Jenkins & Fortner, 2019). The skills children gain during the extra time spent before school allow them to start at a higher level of achievement. This theory has been a driving force for changes in state policies on school entry age (Jenkins & Fortner, 2019). Parents may perceive that being older relative to one's peers might present additional advantages, as school staff may allocate academic opportunities based on students' relative achievement and behaviors (Schanzenbach & Larson, 2017). However, delaying school entry also has associated costs. In the absence of public pre-kindergarten programs, parents bear the cost of childcare during the additional year of out-of-school time. The quality of childcare during this hold-out year could attribute to achievement-level differences at school entry between children from high- and low-income families (Jenkins & Fortner, 2019). If entering school at an older age indeed leads to more learning in the long run, investments in the hold-out year would be worthwhile. If, however, older children enter with higher achievement level but grow at lower rates once in school, then initial differences would fade over time, resulting in limited long-term benefits (Elder & Lubotsky, 2009).

Research has tested these theories by examining the relation between school entry age and short- and long-run student outcomes. Descriptive studies show that older kindergarten entry age is associated with better academic and social outcomes, such as early literacy measures (Fortner & Jenkins, 2017; Herbst & Strawinski, 2016; NICHD Early Child Care Research Networks; 2007; Stipek & Byler, 2001). However, quasi-experimental studies present mixed findings. Some identified positive effects of older entry age on mental health and academic achievement and a reduction in juvenile crime and incarceration (e.g., Dee & Sievertsen, 2018; Depew & Eren, 2016; Jenkins & Fortner, 2019; McAdams, 2016). Others found negative effects on cognitive skills (e.g., Suziedelyte & Zhu, 2015; Zhang, Zhong, & Zhang, 2017).

Thus far, studies that focused on the relation between age of entry and academic outcomes fall into two categories with little overlap: (a) studies that applied causal evidence to achievement measured at one or two points in time; and (b) studies that used vertically-scaled measures to predict growth over time in the absence of a causal design. To our best knowledge, no research has examined the causal effect of school entry age on academic growth. Furthermore, studies that provided causal estimates used data from one US state (e.g., Cook & Kang, 2016; Depew & Eren, 2016). No research has compared effects across states or across schools within a state. Thus, the generalizability of existing findings is limited.

This study addresses these key gaps in the literature by answering two research questions:

(1) What is the impact of being a year older at kindergarten entry on students' academic growth during the first three years of school?

(2) Does the impact of being a year older at kindergarten entry on academic growth vary by student characteristics (e.g., gender, race/ethnicity) or by state?

We apply a regression discontinuity design to vertically-scaled academic achievement measures collected up to nine times between the fall of kindergarten and the spring of 2nd grade for a large analytic sample of students (N=30,552). Our findings will inform early education policies at the local and state levels and offer cross-state comparisons not found in prior research.

Research on School Entry Age

Scholars posit that the age at which students enter school influences downstream academic and labor outcomes (Elder & Lubotsky, 2009). A child's maturity at school entry could affect her growth rate and academic opportunities. In the short run, older children might benefit more than younger children from schooling in the early grades. In the long run—assuming all children get the same the number of years of schooling but older entrants learn at a faster rate and have better academic opportunities than younger entrants while in school–entering school later could result in greater human capital accumulation and higher wages. Perceiving these potential benefits to outweigh the cost of additional childcare, many parents decide to redshirt their children. Children who are white, male, and from economically advantaged families are more likely to enter school at an older age (Bassok & Reardon, 2013; Huang, 2015). However, research has found mixed results on the effect of school entry age student outcomes.

Earlier observational studies have tended to find that children who enter school at an older age scored higher on cognitive, academic, and behavioral measures compared to children who entered school at a younger age (see Stipek, 2002 for a review), but the differences diminish in higher grades. A related line of inquiry focused specifically on redshirting by parent choice. These studies used logistic regressions to predict the likelihood of redshirting and found small associations between redshirting and downstream outcomes (e.g., Fortner & Jenkins, 2017; Graue & DiPerna, 2000). Results from these correlational and predictive studies are informative but do not constitute credible causal effects. Older entrants and younger entrants can differ in systematic ways, and the direction of the bias is ambiguous. Parents might redshirt a child because the family is socioeconomically advantaged and can afford an extra year of out-of-school care; or they might redshirt the child because the child faced a developmental challenge. These factors are unobservable by the researcher but likely to influence downstream outcomes.

To deal with the issue of omitted variable bias, a line of quasi-experimental research used national or state cutoff dates to identify exogenous variations in school entry age and isolate the effect of being older on test scores and incarceration (e.g., Bedard & Dhuey, 2006, 2012; Bedard, Figlio, Karbownik, & Roth, 2019; Fletcher & Kim, 2016; Lenard & Peña, 2018; McAdams, 2016). Some of these studies use an instrumental variables approach, relying on the season or month of birth as an instrument for age measured at school entry. This method would effectively address omitted variable bias if the instrument is uncorrelated with other unmeasured determinants of outcome. However, as seasonal birth rates have been shown to vary based on family background characteristics (Buckles & Hungerman, 2013), this approach may still produce biased estimates.

Regression Discontinuity (RD) Studies

Recent studies have privileged the use of the regression discontinuity (RD) design, which improves upon other quasi-experimental methods by leveraging the jump in the average age of students at the state's cutoff date. States set cutoff dates on or before which a child must turn five to enter school. Students born on or just before the cutoff date are presumably similar on average to students born just after the cutoff date in background characteristics. But the former group, having just turned five, are permitted to enter school while the latter must wait until the following year when they are close to six years old. Any differences in downstream outcomes between the two groups can be interpreted as the causal effect of being a year older at school entry. The state policy creates exogenous variation in school entry age at a single cutoff date. This makes the RD less vulnerable to biases due to seasonal birth endogeneity since the RD estimates the local average treatment effect at a single point.

Results from RD studies on the effects of delayed school entry varied substantially by context. Dee and Sievertsen (2018) used data from Denmark and found that being a year older at school entry reduces inattention and hyperactivity measured at seven and 11 years old. Applying a similar design, Depew and Eren (2016) found reductions in juvenile crime for black females. In contrast, studies using samples of Australian and Chinese children showed that entering school at an older age is detrimental to cognitive skills (Suziedelyte & Zhu, 2015; Zhang, Zhong, & Zhang, 2017). Cook and Kang (2018) examined data from North Carolina and found that being a year older at school entry raised test scores in both math and reading. Jenkins and Fortner (2019), likewise using North Carolina data, also concluded that older entry age yielded (in this case small) benefits to students' test scores in the spring of 3rd grade, as well as reduced probabilities of being identified as having a disability.

The RD literature improves upon other studies in terms of addressing selection into entering school at an older age. However, the extant RD studies share two key limitations. First, studies that focused on the US used data from only one state (e.g., Cook & Kang, 2018; Depew & Eren, 2016; Jenkins & Fortner, 2019). These studies yielded mixed findings, suggesting that the unique context of each study may limit the generalizability of the results. We know of no study that compared effects across states or across schools within a state. Second, the outcomes in previous RD studies were only measured at one or two points in time, and the achievement measures were not vertically scaled. No research has examined the causal effect on academic growth. This limitation means that we do not know whether any initial causal effects of older entry age on achievement fade as students progress through school. This is an important question, given the cost associated with an additional year of out-of-school childcare.

Studies that examined academic growth using descriptive and instrumental variables approaches presented mixed findings. To our knowledge, only two have used vertically-scaled measures to study the relation between school entry age and academic growth over time. One paper used survey data on 900 children from 10 geographical locations in the US and applied hierarchical linear modeling (HLM) to Woodcock-Johnson test scores (NICHD Early Child Care Research Networks, 2007). The study found that children who entered kindergarten at a younger age had higher initial scores, but children who entered at older ages experienced greater increases overtime, outperforming younger entrants in 3rd grade. Datar (2006) used the nationally representative Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K) survey data and an instrumental variable approach to estimate the relation between kindergarten entry age and gains in math and reading test scores between the fall of kindergarten and the spring of 1st grade. She found that older kindergarten entrants had significantly higher initial test scores, as well as steeper score trajectories during the first two years of school. These results, along with mixed findings from the RD literature, highlight the need to further examine the effects of school entry age on not only achievement but growth over time by applying more rigorous methods to vertically-scaled achievement measures.

Contributions of the Current Study

This study addresses these important gaps in the research. We employ a multilevel, fuzzy RD framework to examine the causal impact of entering kindergarten a year older on academic growth in the first three years of school. We use rich data from NWEA's MAP Growth K-2 assessment, which has been administered in a consistent format longitudinally across multiple states. Our analytic sample includes 30,552 students across three states who were born within 30 days of the school entry cutoff date. We ask whether children who turned five years old around the same time but entered school one year apart had different academic growth trajectories. Our focus is on the extent to which school entry age affects the rate at which students learn in each school year, a question immediately relevant to education policy.

This paper makes four important contributions to the literature. First, our fuzzy RD design leverages knowledge of the rule by which school entry age is assigned to credibly identify the causal effect at the cutoff date. Second, using repeated measures of achievement within year, we estimate students' growth trajectories separately for each grade level, unmasking the differences between grades that earlier studies pooled. Datar (2006) only observed student achievement in the fall of kindergarten and the spring of 1st grade. The NICHD (2007) study tested children at 54 months then at the end of 1st and 3rd grades. As a result, these two studies extrapolated children's growth trajectories using only two or three data points measured 18 or more months apart. Our study greatly improves upon these studies by modeling growth within as well as across school years, providing a much clearer, more detailed picture of how student achievement develops over time and how age of school entry impacts development. Our unique data include academic calendars, allowing us to model instructional time that elapsed between test administrations. Third, in addition to local average treatment effects for the average student, we estimate between-school variations in the effect of older entry age and explore school-level covariates that contribute to the variations. Finally, we expand the observed time window of growth from the first two years to the first three years of school for the main sample and one additional year for a subsample. This is a significant extension, considering the importance of examining growth in this critical period of early learning.

Data

Data Source

The data used in this study came from the Growth Research Database (GRD) at NWEA. The GRD contains longitudinal test score data from students in thousands of public school districts across the country that partner with NWEA for a variety of purposes (to monitor growth throughout the school year, school improvement, teacher/school evaluation, to inform parents of their children's progress, as an indicator for intervention or special programming, etc.). In this study, we focus on three states: two Midwestern and one Southern. We chose these three states because (a) all three shared a birthdate cutoff for kindergarten entry during the years of the study and (b) the GRD contained data on a sizable proportion of the schools serving kindergarten students in each state during the study period (36%, 59%, and 77% of schools serving kindergarten in the three states, respectively). It should be noted that the schools in the NWEA sample are not randomly selected within each state. Districts and schools select into administering the MAP Growth assessment to their students. Most schools that partner with NWEA test the majority of students within a grade (an average of 80% of enrolled students). A comparison of the schools in our sample with the public schools serving kindergarten through 2nd graders in each state is available upon request.

Measures

We examine students' reading and mathematics scores on NWEA's MAP Growth assessment. Each test is aligned to state content standards and takes approximately 40 to 60 minutes depending on the grade and subject area. The MAP Growth assessments are computerized, adaptive tests typically administered in the fall, winter, and spring. In the early grades, MAP Growth includes developmentally-appropriate items, interactive elements, and audio supports to engage and accurately assess early learners. Test scores are reported on the RIT scale, where RIT stands for Rasch Unit and is a linear transformation of the logit scale units of the Rasch item response theory model.

In addition to assessment scores, we also have students' race/ethnicity, gender, school calendars (e.g., start and end dates), and students' birthdate. We use school calendar, test dates, and birthdates to calculate age at school entry and how many months students have been in school prior to testing. Schools set their own testing windows and there is typically a fair amount of variation in how many weeks into the school year that students are first assessed. Furthermore, we use a set of school characteristics reported by the Common Core of Data (CCD) from the National Center for Education Statistics (NCES). The CCD variables used in this study include school percentage of Free or Reduced- Price Lunch (FRPL) receipt, percentage of White students in the school, and percentage of Black students in the school. We use these school-level covariates in the multi-level models employed to estimate growth.

Sample

We focus on students in three birth year cohorts, following students who were born in calendar years 2009 to 2011 and entered kindergarten between the fall of 2014 and the fall of 2017. There are a total of 9 possible observations per test subject per student across the fall, winter, and spring of K, 1st, and 2nd grades. Students born in calendar year 2009 form a subsample and were followed to the spring of 3rd grade. Table 1 provides a visualization of the birth year cohorts. For example, students in the 2009 birth cohort who turned five on or before September 1st, 2014 were eligible to enter kindergarten in 2014-15 whereas students who turned five after September 1, 2014 were not be eligible to enter kindergarten until 2015-16. The comparison of interest is between students within the same birth year who were born just prior to the cut-off (and entered school at around five-years-old) and those born just after the birthday cut-off (and entered school at almost six-years-old). In any given school year, the two groups were in different grades (e.g., in 2015-16, younger entrants in the 2009 birth cohort were in 1st grade, while older entrants in the same birth cohort were in kindergarten).

[INSERT TABLE 1 HERE]

In total, we have data for 181,876 students across the three states who took the MAP Growth reading assessment or math assessment in the fall of kindergarten between academic years 2014-15 and 2017-18. We restrict the sample for analysis as follows. First, we eliminate records of 6,105 students for whom we were unable to match their school to NCES public school records. Then, we drop 214 students who attended schools with fewer than 10 students represented in the MAP growth data for kindergarten to 2nd grade students from 2014-15 to 2017-18. We then drop 118 students with missing demographic data. Finally, we retain in our analytic sample students whose birthdates were within 30 days of the states' school of entry cutoff date, following previous RD studies on school entry age (e.g., Dee & Sievertsen, 2018).

The analytic sample includes data for 30,552 students in 1,305 schools. Table 2 presents demographic information for reading and math test takers separately. The reading sample is 49% female, 44% White, 23% Black, 16% Hispanic, and 17% other race/ethnicity. The math sample is 49% female, 42% White, 24% Black, 18% Hispanic, and 16% other race/ethnicity. A comparison of the schools in our analytic sample with the public schools serving kindergarten through 2nd graders in each state is available upon request.

[INSERT TABLE 2 HERE]

Research Design

We first descriptively compare the math and reading trajectories of the students who were born just before and just after the cutoff by plotting the mean achievement level of each group within the fall, winter, and spring of kindergarten, 1st grade, and 2nd grade. These analyses allow for general understanding of trends prior to specifying more sophisticated statistical models.

Then, we estimate the causal effect of being a year older on test scores at kindergarten entry and on growth rates in the first three years of school. We employ a "fuzzy" RD design, which incorporates a two-stage least squares approach, with students' date of birth as the running variable (Lee & Lemieux, 2010). States set a cutoff date on or before which students must turn five to enter kindergarten. We center the students' date of birth such that $Days_i$ = 0 for students born on the first day after the cutoff. If parents follow their state's cutoff dates for kindergarten entry, children would enter kindergarten when they are five years old. Students born just before the cutoff would have just turned five; students born just after the cutoff would enroll the following year, just before turning six. Thus, we expect to observe a discontinuity, or "jump", in the kindergarten entry age of children born around the cutoff date. The RD design exploits this discontinuity in entry age to estimate the causal effect of entering a year older.

Our reduced-form equations model academic outcomes as a flexible function of the running variable and an indicator for having a birthdate after the cutoff:

$$y_i = \alpha_0 + \alpha_1 \mathbf{1}(Days_i \geq 0) + f(Days_i) + \rho' X_i + \varepsilon_i \quad (1)$$

where y_i represents the outcomes of interest (initial test score in the fall of kindergarten year; growth rates estimated using multi-level modeling, described later); $Days_i$ represents the distance between the student's birthdate and the cutoff date; and X_i is a vector of student-level covariates. α_1 is the parameter of interest, which represents the causal effect of being a year older at kindergarten entry. For the first outcome of interest, initial test score, this parameter represents the gap in academic achievement level at kindergarten entry between five-year-old and six-year-old students. For the second set of outcomes, academic growth, this parameter represents the difference between the growth rates of students who enter kindergarten at five years old and six years old.

Validity of the RD Design

The validity of the design hinges on whether the cutoff induced variation in kindergarten entry age and whether assignment to either side of the cutoff was "as good as random" (Lee & Lemieux, 2010). We report evidence on these two conditions below.

For the first stage in the fuzzy RD design, we test if there is a discontinuity in the age of entry at the cutoff. As shown in Appendix A1, having a birthdate after the state cutoff significantly increased the age at which children entered kindergarten by approximately three quarters of a year. Then, we check that students with birthdates around the cutoff are similar on observable pretreatment characteristics. Using an RD model with linear splines, we test if the densities of students' gender and race/ethnicity are continuous (see Appendix Table A2 Panel A). Unfortunately, we do not observe students' socioeconomic status and were unable to use it to test for balance or as a control. Prior literature shows that parents are very unlikely to precisely manipulate their children's birthdates (Dickert-Conlin & Elder, 2010). We perform density tests and verify that there is no evidence of precise manipulation of the running variable in our data

(McCrary, 2008). Results are reported in Appendix A3. These checks provide some reassurance that close to the cutoff date, assignment of students to their age at kindergarten entry is "as good as random" (Lee & Lemieux, 2010).

Another concern regarding the RD design is the appropriate choice of bandwidth and functional form. We address this by taking a local linear regression approach. For brevity, we describe below our preferred RD model specification, which includes linear splines and uses data within a bandwidth of 30 days. To test the sensitivity of our results, we compare estimates across a variety of bandwidths and estimates from linear and quadratic RD specifications.

Having conducted checks for the validity of the RD design, we use an RD model with linear splines to estimate the impact of being a year older on test scores in the fall of kindergarten. We use two approaches to estimate the impact on initial test score. Our first approach follows the standard RD framework, using the students' kindergarten fall term test score as the outcome. This approach treats the students' test score as their achievement level at the start of kindergarten, regardless of when testing occurred in the fall. The limitation of this approach is that some students in our sample have been exposed to up to two months of instruction prior to their fall test, which confounds estimates of achievement differences at school entry. Our second approach incorporates RD into a multilevel growth model, similar to Umansky (2016). This approach accounts for individual differences in testing dates in both the estimation of (a) initial test scores at school entry and (b) growth rates across the school year. As described in additional detail in the following section, the intercept parameter in the multilevel growth model is an extrapolation of student achievement to the first day of school, allowing for an estimate of the impact of entering school a year older that is uncontaminated by differences in school exposure We present both sets of results for comparison.

Multilevel Growth Model

We use a piecewise multilevel growth model (e.g., Downey, von Hippel, & Broh, 2004; von Hippel, Workman, Downey, 2018). In this set-up, MAP Growth test scores (level 1) observed in each term are nested within students (level 2) and schools (level 3). We include any student who has at least one MAP Growth score, even if he or she did not test in all nine waves.

In these models, school affiliation is considered time-invariant. In the situation where a student switched schools during the study years, students are assigned to the school in which the student entered kindergarten because transferring to another school may be endogenous. We use school characteristics reported for academic year 2013-14, though we note that our examination of the multiple years of school characteristics available from the CCD indicate that school composition tends to be highly stable over time (Chingos, 2020). For schools with missing data for the 2013-14 school year, we use characteristics for an adjacent school year (e.g., 2014-15).

Rather than examining overall growth trajectories across years, our piecewise growth model specification allows us to separately examine differences in academic growth estimates between younger and older students in kindergarten, 1st grade, and 2nd grade as well as the summers after kindergarten and 1st grade. Seasonal patterns of learning, where gains during the school year are followed by flattening or dropping of test scores during the summer, have been observed across a range of datasets (von Hippel & Hamrock, 2019). Additionally, average growth rates have been found to decelerate across school years (Bloom et al., 2008; Thum & Hauser, 2015), which means that estimating a single overall school-year growth rate will mask systematic differences in learning rates across grade levels. For these reasons, researchers interested in modeling growth across multiple timepoints during the school year typically rely on a multilevel piecewise growth models that accounts for variation in testing date within the school year and allows for separate growth terms per school year and summer (e.g., Downey, von Hippel, Broh, 2004; Quinn et al., 2016; von Hippel, Workman, Downey, 2018). By separately specifying growth terms for each school year, we can test whether the potential advantage of entering school a year older is constant across grades or if it begins to fade as students progress through school. This model estimates students' academic growth as a linear function of their "months of exposure" to each school year and summer break. Months of exposure is calculated based on a student's school start and end dates and the test administration dates. For example, a hypothetical student testing at the end of August in 1st grade may have 9.3 months of exposure to kindergarten, 2.7 months exposure to summer break following kindergarten, and one week of exposure to 1st grade (Online Appendix Table OA2 provides more detail about the "months of exposure" calculations).

We first estimate the monthly learning rates during each school year and summer from kindergarten to 2nd grade. At level 1, the growth model is:

$$y_{tij} = \pi_{0ij} + \pi_{1ij}G0_{ij} + \pi_{2ij}S1_{ij} + \pi_{3ij}G1_{ij} + \pi_{4ij}S2_{ij} + \pi_{5ij}G2_{ij} + e_{tij}. \quad (2)$$

We view each test score y_{tij} as a linear function of the months that student *i* in school *j* has been exposed to kindergarten ($G0_{ij}$), 1st grade ($G1_{ij}$), and 2nd grade ($G2_{ij}$); and the number of months that the student has been exposed to the summer after kindergarten ($S1_{ij}$) and 1st grade ($S2_{ij}$). As von Hippel and colleagues (2018) note, this model "implicitly extrapolates beyond the test dates to the scores that would have been achieved on the first and last day of the school year" (p. 335). The intercept (π_{0ij}) therefore is the predicted score for student *i* in school *j* testing on the first day of kindergarten, even if the student tested four weeks into the school year. The slopes ($\pi_{1ij}, \ldots, \pi_{5ij}$) are the monthly learning rates of student *i* during each school year and summer.

At level 2, we include an indicator for having a birthdate after the state cutoff ($A_{ij} = 1$ if students were born after the cut-off, zero otherwise), a measure of the distance between student's birthdate and the state's age cutoff date (Days_{ij}), and the interaction between these terms ($A_{ij} * \text{Days}_{ij}$). The inclusion of these terms allows us to capture differences between younger and older kindergarten entrants with respect to the intercept (estimated test scores at kindergarten entry) as well as the school year/summer learning rates. This level-2 equation is analogous to an RD model with linear splines in the traditional RD framework. As in equation (1), the coefficient for the A_{ij} term is the estimated effect of being a year older at kindergarten entry on learning rates. Additionally, we include dummy variables at level 2 for cohort ($C2_{ij}$ and $C3_{ij}$) to allow for cohort differences in the estimation of the intercept and growth terms (where the 2009 birth year is the reference group). Random effects are included at both the student- and school-level to capture variation in the intercept and slopes across levels of the model. In Equation (3) below, we display the specification of the student-level random intercept term (π_{0ij}). Omitted here for brevity (see Appendix A4 for the full set of student- and school-level equations), the same specification is used for each of the student slope terms ($\pi_{1ij}, \ldots, \pi_{5ij}$). Lastly, state fixed effects are included at level 3 of the model.

Level-2 Model (student (i) within school (j)): (3)

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\text{A}_{ij} + \beta_{02j}\text{Days}_{ij} + \beta_{03j}\left(\text{A}_{ij} * \text{Days}_{ij}\right) + \beta_{04j}\left(C2_{ij}\right) + \beta_{05j}\left(C3_{ij}\right) + r_{0ij}$$
$$\vdots$$
$$\pi_{5ij} = \beta_{50j} + \beta_{51j}\text{A}_{ij} + \beta_{52j}\text{Days}_{ij} + \beta_{53j}\left(\text{A}_{ij} * \text{Days}_{ij}\right) + \beta_{54j}\left(C2_{ij}\right) + r_{5ij}$$

This model described above represents our baseline model (which we will refer to throughout the paper as Model I). We also test whether the estimated effect of entering kindergarten a year older is sensitive to the inclusion of key student and school-level covariates. In a second model (Model II), we add indicators for gender ($Female_{ij}$) and race/ethnicity

($Black_{ij}$, $Hispanic_{ij}$, and $OtherRace_{ij}$), where younger male White students are the reference group. A third model (Model III) extends Model II to include three grand-mean centered school-level covariates: (a) school percentage of Free or Reduced-Price Lunch (FRPL) receipt, (b) percentage of White students, and (c) percentage of Black students.

Effect Heterogeneity

Extant literature suggests that the treatment effects of entering kindergarten a year older may not be homogenous (e.g., Dee & Sievertsen, 2018; Depew & Eren, 2016; Jenkins & Fortner, 2019). To address our second research question ("Does the impact of being a year older at kindergarten entry on academic growth vary by student characteristics or by state?"), we examine the heterogeneity of our treatment estimates by estimating three additional growth models that test for interactions between entering school a year older (A_{ij}) and key covariates. These additional models build upon Model II, which already contains dummy variables for birth year cohort, state, race/ethnicity, and gender. First, we interact having a birthdate after the cutoff with the state dummy variables to see if the effect is constant across the three states in our study (Model IV). Second, we test whether the effect of entering school a year older is different between boys and girls (Model IV). Lastly, we test interactions between entering a year older and race/ethnicity to estimate differential treatment effects between racial/ethnic subgroups (Model VI). Online Appendix OA1 shows the full specification for Models IV through VI.

In the presence of effect heterogeneity, our estimates would only be defined for the "compliers" in the sample, or children who choose the school entry age assigned to them by their state's cutoff date. To further explore the heterogeneity of treatment effects, we follow Dee and Sievertsen (2018) and distinguish compliers from "always-takers" and from "never-takers' in the intent-to-treat design. Students who choose a school entry age that is different from the age assigned by the state's cutoff date may be distinct from students who follow the policy in important ways. For example, if redshirted children come from more socioeconomically-advantaged families than children who enter school on time, the impact of entering school a year older on redshirted children may be different from the impact on other children. We test the generalizability of our findings by interrogating the extent to which the students who followed the state's age of entry policy differ from those who chose to delay or enter early (Dee & Sievertsen, 2018; Bertanha & Imbens, 2014). We make the reasonable assumption that no students were "defiers" who would have chosen to disobey their state's cutoff regardless of the age of entry it assigned (i.e., would have entered at six years old had the cutoff mandated entering at five, but would have entered at five years old had the cutoff mandated entering close to six). Defining treatment as entering kindergarten a year older, we refer to students who, regardless of the policy cutoff, would have entered kindergarten close to five years old as "never-takers" and students who would have entered close to or older than six as "always-takers".

We conduct this analysis graphically by plotting the initial achievement and learning rates of two subsamples. First, we focus on a subsample of students who entered kindergarten after turning five years 11 months old and compare the outcomes of students born on or before the cutoff date (always-takers, N=1,824) and students born after the cutoff date (compliers and always-takers). This comparison provides an indication for whether "redshirters" are distinct from students who enter a year older because of the state's cutoff date. Then, we focus on a subsample of students who entered kindergarten before turning five years one month old and compare the growth rates of students born on or before the cutoff date (compliers and never-takers) and students born after the cutoff date (never-takers, N=448). This comparison provides an indication for whether students who would always choose to enter kindergarten around five years old are fundamentally different from students who enter young because of the state's cutoff date. Significant differences would suggest limited generalizability of the estimated local average treatment effect to students who would always choose to enter school young or old.

Results

K-2 Academic Achievement Trends

[INSERT FIGURE 1 HERE]

Figure 1 presents the average trajectory of math and reading test scores from the fall of kindergarten to the spring of 2nd grade for students with birthdates within 30 days prior to the cut date and students with birthdates within 30 days after the cut date. These results are pooled across cohorts and states (separate state plots available upon request). In this figure, we present group means across time on the RIT scale as well as standardized difference scores within each timepoint. For reference, the average standard deviation (SD) in the fall of kindergarten is 10.06 RIT points in math and 9.42 RIT points in reading.

In the fall of kindergarten, there is a sizable gap in test scores favoring students who enter school a year older over students entering at around five (0.66 SD in math and 0.57 SD in reading). These gaps mostly hold steady in kindergarten (the reading gap even widens) but shrink during 1st and 2nd grade. By the end of 2nd grade, the advantage of being older has almost halved in math (to 0.37 SD) and shrunk considerably in reading (to 0.35 SD), though both gaps do remain sizable. It is also worth noting that the trends presented in Figure 1 demonstrate that both groups of students show learning gains on average during the school year followed by a flattening or drop in test scores during the summer (e.g., summer learning loss). This pattern of gains followed by summer losses, which has also been observed with MAP Growth data by Kuhfeld (2019) and Atteberry and McEachin (2019), among others, necessitates the piecewise growth structure that we specified in our multilevel growth models. In the remainder of the paper, we focus on initial academic achievement and growth during academic years since (a) effects on summer learning rates were imprecisely estimated and (b) schools have more influence over student learning during the year than during the summer. Estimated effects of being older on summer learning rates can be found in Online Appendix OA6.

RQ1. What is the impact of being a year older at kindergarten entry on students' academic growth during the first three years of school?

Using the traditional RD approach, we estimated the impact of being a year older on students' observed test scores in the fall of kindergarten. Appendix Table A2 Panel B presents the estimates. The mean observed score for students who enter at five years old is 139.5 for math and 136.7 for reading. The estimated impact of entering kindergarten a year older is 5.19 RIT for math and 4.05 RIT for reading. In the absence of repeated measures, the observed scores can serve as a proxy for students' achievement level prior to kindergarten entry. However, these scores likely have been affected by days of instruction prior to testing and be higher than scores students would have received had they been tested on the first day of kindergarten. This may bias the estimates on growth. We therefore privilege the results presented in Table 3, which use estimated initial achievement prior to kindergarten instruction as the outcome.

[INSERT TABLE 3 HERE]

Table 3 presents the reduced form estimates from the first set of hierarchical linear models (Models I-III) applied to examine the impact of being a year older on students' initial achievement and academic growth across the first three years of school. For parsimony, only the key parameters of interest pertaining to initial status and growth are included in this table (the full set of coefficients is available upon request). The baseline model (Model I) indicates that as expected, the estimated initial scores prior to instruction, (137.8 for math and 134.8 for reading), are lower than observed scores. Older students enter kindergarten with a significant advantage over younger students (5.47 RIT points on average in math, 4.10 in reading) and show significantly higher monthly learning rates within kindergarten (0.14 RIT points per month in math, 0.24 RIT points per month in reading). However, the advantage seems to flip in 1st grade, with older students showing significantly lower average monthly learning rates in reading and math in both 1st and 2nd grade. For example, in 1st grade, the average student with a birthdate before the cutoff is gaining 2.34 RIT points per month in math, and the average student born after the cut date (who enter older) is learning at a rate that is 0.15 RIT points lower per month than the younger students. These findings hold when student and school background variables are included (Model II and Model III, respectively).

The bottom of Table 3 presents the estimated student- and school-level random effects SDs, which allow us to examine whether there is significant between-school variation in the effect of being older on initial status and growth. In math, there appears to be significant between-school variation in the advantage of being older at school entry (SD=1.25) as well as the monthly learning rates in kindergarten (SD=0.19) and 2nd grade (SD=0.12). Given that overall average advantage of being older in terms of monthly learning rates in kindergarten is only 0.14 RIT points, this degree of between-school variation indicates that there are some schools where being older has a net zero or negative impact on kindergarten growth rates. In reading, between-school variation in the impact of being older is only statistically significant in kindergarten.

[INSERT FIGURE 2 “DISCONTINUITY” HERE]

Figure 2 shows the changes at the birthdate cutoff in initial test score at kindergarten entry and in growth rates over the first three years of school. Data are pooled across states and birth year cohorts (see Online Appendix OA4 for separate figures by state). The horizontal axis represents students' birthdates, centered at the entry cutoff. Values to the left of the black, vertical dash line represent students born on or before the cutoff; values to the right, including 0, represent students born after the cutoff. The vertical axis represents initial performance or growth in RIT points. Each black circle represents the average initial score or average monthly growth for students in the corresponding centered birthdate bin (containing 350 to 700 students). The solid lines illustrate linear fit. The gap between where the two solid fit lines approach the cutoff is the effect of entering kindergarten a year older on test score or growth. The dash lines around the solid line represent the confidence intervals. A jump in RIT at the cutoff, represented by non-overlapping confidence intervals on the left and right sides, indicates a significant effect. Visually, we can detect a difference of a little over five RIT points, and the confidence intervals do not overlap. These correspond to the statistically significant effect of 5.47 RIT points reported in Table 3. Also consistent with the results in Table 3, in both math and reading, the monthly learning rates for older students were significantly higher than younger students during the kindergarten year but significantly lower during the 1st and 2nd grade school years.

Growth in 3rd Grade (Cohort 2009)

We followed the 2009 birth year cohort for one additional school year. In both reading and math, the older students showed significantly lower growth rate in 3rd grade, indicating the initial advantage of being older continues to shrink as students move through school (see Appendix Table A5).

RQ2. Does the impact of being a year older at kindergarten entry on academic growth vary by student characteristics (e.g., gender, race/ethnicity), or by state?

[INSERT TABLE 4 HERE]

Table 4 presents the findings examining whether the impact of being a year older at kindergarten entry on academic growth varies by student characteristics or state. We do not see strong evidence of state-level differences in the impact of being older on initial status or growth in math, though there are some significant interactions by state in reading. For example, State 2 shows a smaller advantage (-1.37 RIT points) of being older in students' initial reading scores than State 3 (the reference state). In addition, the impact of entering school older on students' academic growth across the three school years observed in State 3 is absent in State 2.

The results from Model V (Table 4) indicate that while the effect of being older at school entry on initial test scores is slightly larger for girls than boys (0.59 RIT points in math, 0.93 RIT points in reading), the impact of being older on growth rates across kindergarten through 2nd grade does not significantly differ by gender. Lastly, we examined interactions between being older and students' race/ethnicity (Model VI). Results indicate that Hispanic students experienced a significantly larger effect of being older on initial achievement (0.96 RIT points in math, 1.23 RIT points in reading) but similar effects of being older on growth rates as White students. The only significant interaction for Black students is related to the impact of being older on kindergarten growth rates, where older Black students had larger advantages (0.12 RIT points in math, 0.15 RIT points in reading) compared to White students.

The fuzzy RD design produces results that are defined for compliers, or students who entered kindergarten at an age assigned to them by the state policy. The findings would not be generalizable to students who would have chosen to enter kindergarten at an older (or younger) age regardless of their policy-assigned age of entry. We interrogate the generalizability of our findings by looking at whether always-takers and never-takers of the treatment (being a year older at entry) are distinct. Online Appendix OA5 shows the comparison between (a) always-takers on the left side of the vertical dash line and (b) compliers and always takers (on the right side of the dash line. Always-takers (redshirted children) have higher achievement at kindergarten entry and somewhat higher growth during kindergarten than compliers and always-takers born after the cutoff, but their growth rates during 1st and 2nd grade were indistinguishable. Never-takers (children who would have always entered kindergarten young) were indistinguishable from compliers and never-takers born before the cutoff in all measures. We interpret these results as suggestive evidence that the negative estimated effects on growth rates in 1st and 2nd grade are likely generalizable to students who would always choose to enter school older or younger than their policy-mandated entry age.

Robustness Checks

In addition to checking that our results are robust to the inclusion of student- and school-level covariates (Models II and III), we perform two other sensitivity tests. First, we vary the RD bandwidth from 5 to 45 days around the birthdate cutoff. Results are very similar across the bandwidths (see Online Appendix OA3). Second, we include quadratic splines in our model and find that the addition of quadratic terms, which had insignificant estimates, does not improve model fit (results available upon request). These analyses provide reassurance that the findings from our preferred linear model and bandwidth of 30 days are robust.

Discussion

This study demonstrates a novel approach to estimating the causal impact of kindergarten entry age on academic growth by integrating the RD design with multilevel growth modeling. We report three main findings. First, being a year older at kindergarten entry has significant positive effects on initial math and reading achievement, as well as monthly learning rates during the kindergarten school year. Second, the effects of being older at kindergarten entry on 1st and

2nd grade math and reading growth are negative and significant. In other words, the initial gaps between older and younger students start to close as students progress through the early grades because after kindergarten, older students grow at slower rates compared to younger students. Third, while there are some heterogeneous effects on initial test score, the impact of being older on growth, especially growth after kindergarten, generally does not differ by gender or ethnicity.

The gap we observe in initial test score between older and younger kindergarten students is consistent with previous literature. For example, Datar (2006) and Stipek (2001) similarly found that children who entered kindergarten at an older age had higher initial test scores upon entry. Relatedly, being older has been shown to raise achievement measured at a fixed point relatively early in a student's academic career. Employing a fuzzy RD framework with a slightly weaker first stage than ours (0.72 vs. 0.78), Cook and Kang (2018) found that an extra year of age has significant causal effects on end-of-3rd-grade test scores: .36 SD in reading and .30 SD in math. Our descriptive findings (Figure 1) are consistent with these estimates, showing that a gap of similar magnitude (.37 SD for math, .35 SD for reading) remains between older and younger students in achievement level by the end of 2nd grade.

However, our results on early growth trajectories contrast with earlier studies and add novel evidence to this line of inquiry. Datar (2006) and the NICHD Early Child Care Research Network (2007) also used vertically-scaled measures and found that the achievement of older kindergarten entrants grows faster during the early grades. In contrast, we found that older students only learned faster than younger students during kindergarten; in fact, they grew significantly more slowly during 1st through 3rd grade. The previous papers were not able to distinguish kindergarten year growth from other grade levels because they only observed achievement data before or immediately after kindergarten entry and then again at the end of 1st grade. In modeling growth within and across grades, our findings illuminate a richer trajectory and make important distinctions between grades that allow policymakers and educators to design programs and policies targeted at the appropriate grade levels.

More generally, our findings suggest that the extant evidence on the advantages of older school entry age ought to be interrogated. The decline we see in 1st and 2nd grade in the older-age premium is consistent with the findings of other studies that examined outcomes measured in later grades. For instance, Bedard and Dhuey (2006) found that the 8th grade achievement gap between students who entered school older and younger was much smaller than the gap in 4th grade. Fletcher and Kim (2016) also found that the impact on math and reading achievement of increasing entry age was large in 4th grade, much smaller in 8th grade, and negligible in 12th grade. The gap reduction that begins in 1st grade may continue into later grades, but more rigorous research is required to clarify effects on longer term outcomes.

The literature has reported an array of outcomes, from achievement in early grades to juvenile incarceration to adult wages. But almost all studies on school entry age, like Bedard and Dhuey (2006) and Fletcher and Kim (2016), only observed outcomes that were not vertically-scaled and, if repeated at all, measured at least a few years apart. Thus, the current knowledge is built on inconsistent measures taken across disparate student ages, grade levels, and geographical regions, and consumers of the research are having to base their practices on forced connections between distinct findings that may not be comparable or generalizable. Policymakers and educators need more actionable evidence drawn from repeated, vertically-scaled measures of student achievement over time. A viable first step is to examine the growth trajectories of students from 4th grade to high school, bridging the gaps in our understanding of academic development in the middle grades. Research is also needed to improve measurements of academic engagement and socioemotional development. For instance, Dee and Sievertsen (2018) showed using data from Denmark that entering school a year older resulted in lower levels of inattention and hyperactivity at ages 7 and 11. Establishing more reliable and comparable measures will allow the tracking and modeling of these important "non-academic outcomes" across time to further understand the effects of policy and practices.

There are multiple possible explanations for the initial advantage of being older observed in our study. As Jenkins and Fortner (2019) elaborate, students who just miss the birthdate cutoff have an extra year in which they may experience high-quality early education that better prepare them to enter school. Additionally, the extra year of development may result in improved executive functioning and behavior that translate to a higher ability to focus on the MAP Growth assessments in the early grades. However, it is less clear what might explain the fade out of the age advantage in 1st through 3rd grade. It may be that teachers in these grades focus more of their attention on helping to catch up the younger students. Alternatively, the maturational advantage in kindergarten mat begin to narrow in later grades independent of teacher attention (e.g., developmental differences are larger between five and six-year-olds than between seven and eight-year-olds). It is also possible that younger students benefit from having older peers who are higher-achieving and better-behaved. Research on classroom peer effects found that holding the child's own age constant, being relatively young in the class led to higher test scores in kindergarten and 8th grade (Cascio & Schanzenbach, 2016). Relative age may be influential in reducing the gap over time. Our findings beg future research to explore the mechanisms behind the causal link between age and growth trajectories.

Limitations

This study is not without limitations. Since our data contained only a few student covariates, we were unable to examine how the impact of kindergarten entry age on growth differs by socioeconomic status or other subgroup membership, such as English Learners. Future research should examine whether there are additional benefits of being older for these groups. Further research also is needed to disentangle the factors that drove the initial achievement gap between older and younger students. Without data on childcare participation, we were unable to distinguish between the effects of relative maturity and hold-out year experience on initial achievement. Our sample of schools within each state is not a random subset of schools serving kindergarteners and may differ from all public schools in the state. Second, we cannot yet track student cohorts beyond 3rd grade. It is important to know whether the gap between younger and older students continues to narrow after 3rd grade. Finally, we are only able to follow students who enrolled in public schools and took the MAP Growth assessments. The results may be susceptible to bias if the probability of enrolling in a school that administers MAP Growth jumps at the birthdate cutoff, though we are not aware of a theoretical basis for this concern.

Concluding Remarks

The age of school entry matters to schools and families because of its implications on students' academic readiness and subsequent experiences and on the private and public costs associated with delaying schooling. This study contributes novel causal evidence to the literature by showing that initial advantages to entering kindergarten a year older fade as older students grow at significantly lower rates during the second and third years of school. Based on these results, we recommend that policymakers and families consider these fade out effects before adopting policies and practices that raise children's school entry age.

Disaggregating growth trajectories by grade level can be especially critical to addressing equity among subgroups of students. A few studies have investigated effect heterogeneity for school entry age by gender and race, but due to the lack of comparable repeated measures, none has been able to identify the grade level(s) in which the differential effects manifest. For instance, Jenkins and Fortner (2019) examined test scores in 3rd grade and found that practices that increased students' age at school entry differentially benefitted low-income students but disadvantaged non-White students. What remained unclear is how different subgroups grow over each of the early grades and how age contributes to the growth patterns. One might conclude from previous research, for example, that low-income students should be redshirted at higher rates since, judging from 3rd grade test scores, they benefit more from older entry age than higher income students (Jenkins & Fortner, 2019). However, it is possible that older age has a positive differential effect on low income students' growth during kindergarten but a negative differential effect in the years that follow, such that in the long run the benefits disappear. Without observing the subgroups' growth trajectories within and across years, heterogeneous effects on achievement measured in one point in time should be interpreted with caution.

We show that initial achievement gaps between older and younger kindergarten entrants exist across subgroups but growth rates over the first few years of school do not vary significantly by subgroup status. This is an important insight unfound in previous research. To design policies aimed at providing equitable opportunities to all student subgroups, knowledge of both achievement levels and development over time is required. This study demonstrates that research designs combining the strengths of causal inference and multilevel growth modeling offer great potential for identifying the impact of policies on student achievement and growth.

References

Bassok, D. & Reardon, S. F. (2013). Academic "Redshirting" in Kindergarten: Prevalence, Patterns, and Implications. *Educational Evaluation and Policy Analysis*. doi:10.3102/0162373713482764

Bedard, K. & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics, 121*(4), 1437-1472. doi:10.1093/qje/121.4.1437

Bedard, K. & Dhuey, E. (2012). School-entry policies and skill accumulation across directly and indirectly affected individuals. *Journal of Human Resources, 47*(3), 643–683.

Bertanha, M., & Imbens, G.W. (2014). External validity in fuzzy regression discontinuity designs. (*Technical report*). Cambridge, MA: National Bureau of Economic Research.

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289-328.

Buckles, K. S., & Hungerman, D. M. (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics, 95*(3), 711-724.

Cascio, E. U., & Schanzenbach, D. W. (2016). First in the class? Age and the education production function. *Education Finance and Policy, 11*(3), 225-250.

Chingos, M. (2020, January 29). *Features: Explore your school's changing demographics*. Urban Institute. Retrieved from https://www.urban.org/features/explore-your-schools-changing-demographics

Cook, P. J., & Kang, S. (2016). Birthdays, Schooling, and Crime: Regression-Discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation. *American Economic Journal: Applied Economics, 8*(1), 33-57. doi: 10.1257/app.20140323

Cook, P. J., & Kang, S. (2018). The School-Entry-Age Rule Affects Redshirting Patterns and Resulting Disparities in Achievement. *National Bureau of Economic Research Working Paper Series, No. 24492*. doi:10.3386/w24492

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review, 25*(1), 43-62. doi:10.1016/j.econedurev.2004.10.004

Dee, T. S., & Sievertsen, H. H. (2018). The gift of time? School starting age and mental health. *Health Economics, 27*(5), 781-802. doi:10.1002/hec.3638

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

Depew, B. & Eren, O. (2016). Born on the wrong day? School entry age and juvenile crime. *Journal of Urban Economics 96,* 73–90.

Dhuey, E., Figlio, D., Karbownik, K., & Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management, 38*(3), 538-578. doi:10.1002/pam.22135

Dickert-Conlin, S., & Elder, T. (2010). Suburban legend: School cutoff dates and the timing of births. *Economics of Education Review, 29*, 826-841.

Downey, D. B., von Hippel, P. T., & Broh, B. A. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review, 69,* 613–635.

Education Commission of the States. (2018). Kindergarten entrance age. Retrieved from http://ecs.force.com/mbdata/MBQuest2RTanw?rep=KK3Q1802

Elder, T. E., & Lubotsky, D. H. (2009). Kindergarten entrance age and children's achievement. *Journal of Human Resources, 44*(3), 641-683. Retrieved from http://jhr.uwpress.org/content/44/3/641.abstract

Fletcher, J. & Kim, T. (2016). The effects of changes in kindergarten entry age policies on educational achievement. *Economics of Education Review, 50,* 45-62.

Fortner, C. K., & Jenkins, J. M. (2017). Kindergarten redshirting: Motivations and spillovers using census-level data. *Early Childhood Research Quarterly, 38*, 44-56.

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

Herbst, M. & Strawiński, P. (2016). Early effects of an early start: Evidence from lowering the school starting age in Poland. *Journal of Policy Modeling, 38,* 256-271.

Huang, F. L. (2015). Investigating the prevalence of academic redshirting using population-level data. *AERA Open, 1*(2), 1-11.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *The Journal of Economic Literature*, *48*(2), 281–355.

Lenard, M. A. & Peña, P. A. (2018). Maturity and minorities: The impact of redshirting on achievement gaps. *Education Economics, 26*(6): 593-609. DOI: 10.1080/09645292.2018.1468873

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, 142*(2), 698-714. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304407607001133

Meisels, S. J. (1999). Assessing readiness. In R. Pianta & M. J. Cox (Eds.), *The transition to kindergarten* (pp. 39-66). Baltimore: Paul H. Brookes.

Morrison, F. J., Alberts, D. M., & Griffith, E. M. (1997). Nature-nurture in the classroom: Entrance age, school readiness, and learning in children. *Developmental Psychology, 33*(2), 254-262. doi:10.1037/0012-1649.33.2.254

NICHD Early Child Care Research Network. (2007). Age of entry to kindergarten and children's academic achievement and socioemotional development, *Early Education and Development, 18*(2), 337-368, DOI: 10.1080/10409280701283460

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher*, *45*, 443–453.

Schanzenbach, D. W., & Larson, S. H. (2017). Is Your Child Ready for Kindergarten? *Education Next,* 17(3), 18-24.

Stipek, D. (2002). At what age should children enter kindergarten?: A question for policy makers and parents. *SRCD Social Policy Report, 16*(2). Retrieved from http://www.srcd.org/Documents/Publications/SPR/sprv16n2.pdf

Stipek, D. & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Applied Developmental Psychology, 22,* 175-189.

Suziedelyte, A. & Zhu, A. (2015). Does early schooling narrow outcome gaps for advantaged and disadvantaged children? *Economics of Education Review, 45,* 76-88.

Thum Y. M., & Hauser, C. H. (2015). NWEA 2015 MAP Norms for Student and School Achievement Status and Growth. NWEA Research Report. Portland, OR: NWEA.

Umansky, I. M. (2016). To be or not to be EL: An examination of the impact of classifying students as English Learners. *Educational Evaluation and Policy Analysis, 38*(4), 714-737.

von Hippel, P. T., & Hamrock, C. (2019). Do test score gaps grow before, during, or between the school years? Measurement artifacts and what we can know in spite of them. *Sociological Science, 6,* 43–80.

von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in reading and math skills forms mainly before kindergarten: A replication, and partial correction, of "Are schools the great equalizer?" *Sociology of Education, 91*(4), 323-357.

Zhang, S., Zhong, R., & Zhang, J. (2017). School starting age and academic achievement: Evidence from China's junior high schools. *China Economic Review, 44,* 343-354.

Table 1

Visualization of the Birth-Year Cohorts and the School Years in which Students were Assessed

		Birthday Before Cut Date									Birthdate after Cut Date			
Cohort	Year of Birth	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	**Cut Date**	Sep.	Oct.	Nov.	Dec.
1	2009				K: 2014-15 1st: 2015-16 2nd: 2016-17 3rd: 2017-18					**9/1/2014**		K: 2015-16 1st: 2016-17 2nd: 2017-18 3rd: 2018-19		
2	2010				K: 2015-16 1st: 2016-17 2nd: 2017-18					**9/1/2015**		K: 2016-17 1st: 2017-18 2nd: 2018-19		
3	2011				K: 2016-17 1st: 2017-18 2nd: 2018-19					**9/1/2016**		K: 2017-18 1st: 2018-19		

Note. Cohort in our study is defined by students' birth year (2009 to 2011). In this set-up, the students with birthdates right after the Cut Date are the "treated" group, as they enter school almost a year older than the students whose birthdates fall in the month right before the Cut Date. Due to the unavailability of treated group outcome data for Cohort 3, we omit the Cohort 3 dummy coefficients from Table 3, "Growth in 2nd Grade" results.

Table 2

Demographic Characteristics of the Sample for Students Within 30 Days Bandwidth

	All 3 States	State 1	State 2	State 3
	Reading (BW = 30)			
Female	0.49	0.49	0.49	0.49
White	0.44	0.4	0.47	0.51
Black	0.23	0.19	0.13	0.33
Hispanic	0.16	0.21	0.11	0.09
Other race	0.17	0.20	0.28	0.08
Born after entry cutoff	0.55	0.53	0.56	0.58
Cohort 1 - 5th birthday in 2014	0.29	0.29	0.32	0.27
Cohort 2 - 5th birthday in 2015	0.35	0.34	0.35	0.36
Cohort 3 - 5th birthday in 2016	0.36	0.37	0.33	0.37
N. Students	26,172	14,714	3,008	8,450
N. Schools	1,155	706	146	303
	Math (BW = 30)			
Female	0.49	0.49	0.49	0.49
White	0.42	0.36	0.47	0.51
Black	0.24	0.21	0.14	0.33
Hispanic	0.18	0.24	0.11	0.09
Other race	0.16	0.18	0.28	0.08
Born after entry cutoff	0.55	0.53	0.56	0.58
Cohort 1 - 5th birthday in 2014	0.29	0.31	0.32	0.26
Cohort 2 - 5th birthday in 2015	0.35	0.34	0.35	0.37
Cohort 3 - 5th birthday in 2016	0.36	0.35	0.33	0.37
N. Students	30,128	17,166	3,118	9,844
N. Schools	1,298	809	149	340

Note. There is a slightly higher percentage of students in the "Born after entry cut-off" group because this group contains an extra day ($0 \leq Days_{ij} \leq 30$) compared to the "Born before or on entry cutoff" group ($-30 \leq Days_{ij} < 0$).

Table 3
HLM results for the first three model specifications

	Math			Reading		
Variable	(I)	(II)	(III)	(I)	(II)	(III)
			Starting RIT			
A	**5.47 (0.23)**	**5.51 (0.22)**	**5.50 (0.22)**	**4.10 (0.23)**	**4.13 (0.23)**	**4.12 (0.23)**
Control Group Fall K Score	137.8 (0.30)	140.2 (0.30)	139.4 (0.30)	134.8 (0.27)	135.9 (0.28)	135.5 (0.29)
State 1	0.61 (0.30)	0.92 (0.25)	1.78 (0.27)	0.98 (0.26)	1.14 (0.23)	1.74 (0.27)
State 2	2.65 (0.58)	1.88 (0.48)	1.28 (0.41)	1.76 (0.49)	1.35 (0.43)	0.91 (0.37)
Cohort 2	-1.32 (0.15)	-1.32 (0.15)	-1.33 (0.15)	-0.95 (0.16)	-0.97 (0.16)	-0.98 (0.16)
Cohort 3	-1.85 (0.16)	-1.86 (0.16)	-1.87 (0.16)	-1.14 (0.17)	-1.17 (0.17)	-1.17 (0.17)
Female		0.75 (0.11)	0.74 (0.11)		1.45 (0.11)	1.45 (0.11)
Black		-5.73 (0.19)	-5.01 (0.22)		-3.69 (0.20)	-3.19 (0.22)
Hispanic		-6.20 (0.23)	-5.62 (0.25)		-4.87 (0.22)	-4.44 (0.24)
Other race		-2.02 (0.24)	-1.76 (0.24)		-1.35 (0.23)	-1.15 (0.23)
% FRPL			-6.88 (1.37)			-5.51 (1.43)
% White			4.73 (0.57)			4.00 (0.60)
% Black			1.08 (0.56)			*1.04 (0.64)*
			Growth in Kindergarten			
A	**0.14 (0.03)**	**0.14 (0.03)**	**0.14 (0.03)**	**0.24 (0.03)**	**0.24 (0.03)**	**0.24 (0.03)**
Control Group K Growth	2.28 (0.03)	2.41 (0.04)	2.42 (0.04)	2.15 (0.04)	2.26 (0.04)	2.26 (0.04)
State 1	-0.06 (0.03)	-0.08 (0.03)	-0.08 (0.03)	-0.12 (0.03)	-0.13 (0.03)	-0.12 (0.04)
State 2	-0.12 (0.05)	-0.15 (0.05)	-0.14 (0.05)	-0.15 (0.05)	-0.19 (0.05)	-0.20 (0.05)
Cohort 2	0.10 (0.02)	0.10 (0.02)	0.10 (0.02)	0.07 (0.02)	0.07 (0.02)	0.07 (0.02)
Cohort 3	0.18 (0.02)	0.19 (0.02)	0.19 (0.02)	0.06 (0.02)	0.06 (0.02)	0.05 (0.02)
Female		-0.11 (0.01)	-0.11 (0.01)		0.02 (0.01)	*0.02 (0.01)*
Black		-0.20 (0.02)	-0.23 (0.02)		-0.27 (0.03)	-0.28 (0.03)
Hispanic		-0.04 (0.02)	-0.06 (0.03)		-0.18 (0.03)	-0.19 (0.03)
Other race		*-0.02 (0.02)*	-0.03 (0.02)		*-0.05 (0.03)*	-0.06 (0.03)
% FRPL			-0.06 (0.12)			-0.29 (0.14)
% White			-0.15 (0.06)			*0.00 (0.07)*
% Black			*0.05 (0.07)*			*0.02 (0.09)*
			Growth in 1st Grade			
A	**-0.15 (0.03)**	**-0.15 (0.03)**	**-0.15 (0.03)**	**-0.14 (0.03)**	**-0.14 (0.03)**	**-0.14 (0.03)**
Control Group 1st Grade Growth	2.34 (0.03)	2.46 (0.03)	2.46 (0.03)	2.18 (0.03)	2.26 (0.03)	2.24 (0.04)
State 1	-0.04 (0.03)	-0.05 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.05 (0.03)	-0.04 (0.03)
State 2	0.00 (0.04)	-0.03 (0.04)	-0.05 (0.04)	0.07 (0.04)	0.04 (0.04)	0.03 (0.04)
Cohort 2	0.03 (0.02)	0.04 (0.02)	0.04 (0.02)	-0.06 (0.02)	-0.06 (0.02)	-0.06 (0.02)
Cohort 3	0.09 (0.02)	0.09 (0.02)	0.09 (0.02)	-0.07 (0.02)	-0.07 (0.02)	-0.07 (0.02)
Female		-0.14 (0.01)	-0.14 (0.01)		*-0.01 (0.01)*	*-0.01 (0.01)*

Black		-0.15 (0.02)	-0.15 (0.02)		-0.18 (0.02)	-0.15 (0.03)
Hispanic		-0.06 (0.02)	-0.06 (0.02)		-0.11 (0.03)	-0.08 (0.03)
Other race		*-0.01 (0.02)*	-0.02 (0.02)		*-0.03 (0.03)*	*-0.02 (0.03)*
% FRPL			-0.24 (0.11)			*-0.02 (0.11)*
% White			*-0.02 (0.05)*			0.16 (0.06)
% Black			*-0.03 (0.06)*			*-0.01 (0.08)*
			Growth in 2nd Grade			
A	**-0.11 (0.03)**	**-0.11 (0.03)**	**-0.11 (0.03)**	**-0.10 (0.04)**	**-0.10 (0.04)**	**-0.10 (0.04)**
Control Group 2nd Grade Growth	1.68 (0.03)	1.74 (0.03)	1.73 (0.03)	1.77 (0.04)	1.83 (0.04)	1.81 (0.04)
State 1	*-0.01 (0.02)*	-0.01 (0.02)	0.02 (0.03)	-0.10 (0.03)	-0.10 (0.03)	-0.08 (0.03)
State 2	*0.06 (0.04)*	0.06 (0.04)	0.07 (0.04)	*0.04 (0.05)*	0.04 (0.05)	0.05 (0.05)
Cohort 2	*0.01 (0.02)*	0.01 (0.02)	0.01 (0.02)	*0.00 (0.02)*	0.00 (0.02)	0.00 (0.02)
Female		-0.06 (0.01)	-0.06 (0.01)		0.01 (0.02)	0.01 (0.02)
Black		-0.06 (0.02)	-0.08 (0.02)		-0.13 (0.03)	-0.12 (0.03)
Hispanic		*-0.02 (0.02)*	0.00 (0.02)		-0.08 (0.03)	-0.05 (0.03)
Other race		-0.06 (0.02)	-0.06 (0.02)		-0.12 (0.03)	-0.11 (0.03)
% FRPL			-0.15 (0.11)			*0.07 (0.13)*
% White			0.11 (0.05)			*0.13 (0.06)*
% Black			0.15 (0.05)			*0.10 (0.07)*
			Student-level Random Effect SD			
Starting RIT	8.26	8.06	8.06	7.13	6.96	6.96
Growth in K	0.64	0.63	0.63	0.63	0.63	0.63
Growth in 1st	0.51	0.51	0.51	0.44	0.44	0.44
Growth in 2nd	0.21	0.21	0.21	0.57	0.57	0.57
			School-level Random Effect SD			
Starting RIT	4.72	3.74	3.34	3.92	3.33	3.04
Starting RIT - A	1.25	1.26	1.25	1.15	1.20	1.20
Growth in K	0.43	0.42	0.42	0.42	0.41	0.40
Growth in K - A	0.19	0.18	0.18	0.19	0.19	0.19
Growth in 1st	0.34	0.34	0.34	0.32	0.31	0.30
Growth in 1st - A	0.14	0.14	0.14	*0.14*	*0.14*	*0.14*
Growth in 2nd	0.27	0.27	0.26	0.25	0.25	0.25
Growth in 2nd - A	0.12	0.12	0.12	*0.12*	*0.12*	*0.12*

Note. Robust Standard Errors in parentheses. Italicized parameters are not statistically significant. Each panel-column represents a separate regression with the panel title as the dependent variable. Model specifications are detailed in Appendix A4. Control group (students who entered kindergarten at 5 years old) estimates are presented for context. State 1 and State 2 are dummy variables. State 3 is the omitted category. Cohort 2 and Cohort 3 refer to students born in calendar years 2010 and 2011. Students born in 2009 are the omitted category. Female, Black, Hispanic, and Other race are student-level controls. % FRPL (free or reduced-price lunch eligibility), % White, and % Black are school-level controls. For brevity, summer loss estimates are excluded from this table but included in Online Appendix OA6. Some predictors of secondary interest (including C and A*C) were included in each model but excluded here.

Table 4

Results for the Growth Models Examining the Potential Interactions between Being Older and a set of Student Characteristics

Variable	Math (IV)	Math (V)	Math (VI)	Reading (IV)	Reading (V)	Reading (VI)
			Starting RIT			
A	5.72 (0.28)	5.22 (0.25)	5.30 (0.25)	4.32 (0.26)	3.67 (0.25)	3.90 (0.25)
A*State 1	*0.06 (0.26)*			*-0.06 (0.26)*		
A*State 2	-2.34 (0.43)			-1.37 (0.42)		
A*Female		0.59 (0.21)			0.93 (0.22)	
A*Black			*0.15 (0.24)*			*0.14 (0.26)*
A*Hispanic			0.96 (0.30)			1.23 (0.30)
			Growth in Kindergarten			
A	0.15 (0.03)	0.15 (0.03)	0.09 (0.03)	0.34 (0.04)	0.23 (0.03)	0.21 (0.03)
A*State 1	*-0.01 (0.03)*			-0.10 (0.04)		
A*State 2	*-0.07 (0.05)*			-0.31 (0.06)		
A*Female		*-0.02 (0.02)*			*0.02 (0.03)*	
A*Black			0.12 (0.03)			*0.15 (0.04)*
A*Hispanic			0.11 (0.03)			*0.03 (0.04)*
			Growth in 1st Grade			
A	-0.18 (0.03)	-0.14 (0.03)	-0.14 (0.03)	-0.21 (0.04)	-0.13 (0.03)	-0.13 (0.03)
A*State 1	*0.04 (0.03)*			0.09 (0.03)		
A*State 2	*0.10 (0.05)*			0.18 (0.06)		
A*Female		*-0.02 (0.03)*			*-0.01 (0.03)*	
A*Black			*-0.05 (0.03)*			*-0.01 (0.04)*
A*Hispanic			*0.00 (0.04)*			*-0.04 (0.04)*
			Growth in 2nd Grade			
A	-0.07 (0.03)	-0.11 (0.03)	-0.13 (0.03)	-0.11 (0.05)	*-0.08 (0.04)*	-0.12 (0.04)
A*State 1	-0.07 (0.03)			*0.00 (0.04)*		
A*State 2	*0.05 (0.05)*			*0.08 (0.06)*		
A*Female		*-0.01 (0.03)*			*-0.04 (0.04)*	
A*Black			*0.07 (0.03)*			*0.03 (0.04)*
A*Hispanic			*0.00 (0.04)*			*0.06 (0.05)*

Note. Robust Standard Errors in parentheses. Italicized parameters are not statistically significant. Each panel-column represents key parameter estimates from a separate regression with the panel title as the dependent variable. Model specifications are detailed in Online Appendix OA1. State 1 and State 2 are dummy variables. State 3 is the omitted category. Female, Black, and Hispanic are student-level controls. For brevity, summer loss estimates are excluded from this table but included in Online Appendix OA6. Estimates of secondary interest (main effects) are suppressed.

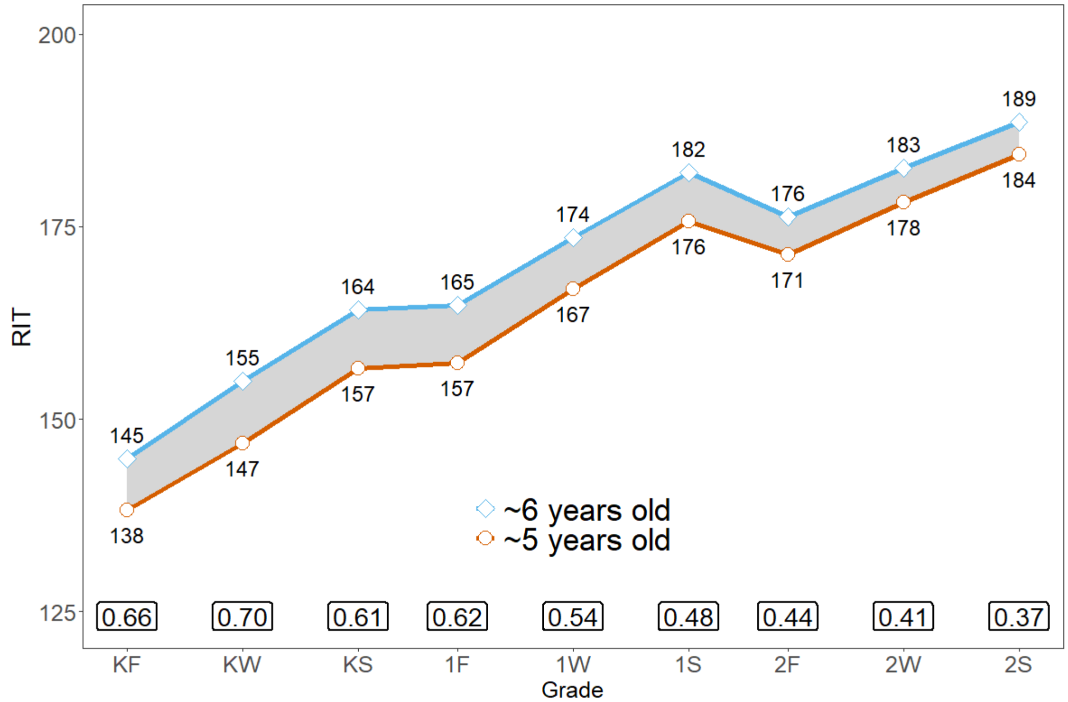
Figure Captions

Figure 1. K-2 Growth Trajectories
Notes: Average trajectories (pooled across states and cohorts) of the students with birthdates within 30 days prior to cut date (circles) and students with birthdates within 30 days after the cut date (diamonds). Group means (rounded) are presented next to the lines, and standardized mean differences between the groups in each term (standardized by the pooled standard deviation calculated within each grade/term pair) are reported at the bottom of each figure.

Figure 2. Estimated Effects on Test Scores

Figure 1

(a) Math K-2 Trajectories



(b) Reading K-2 Trajectories

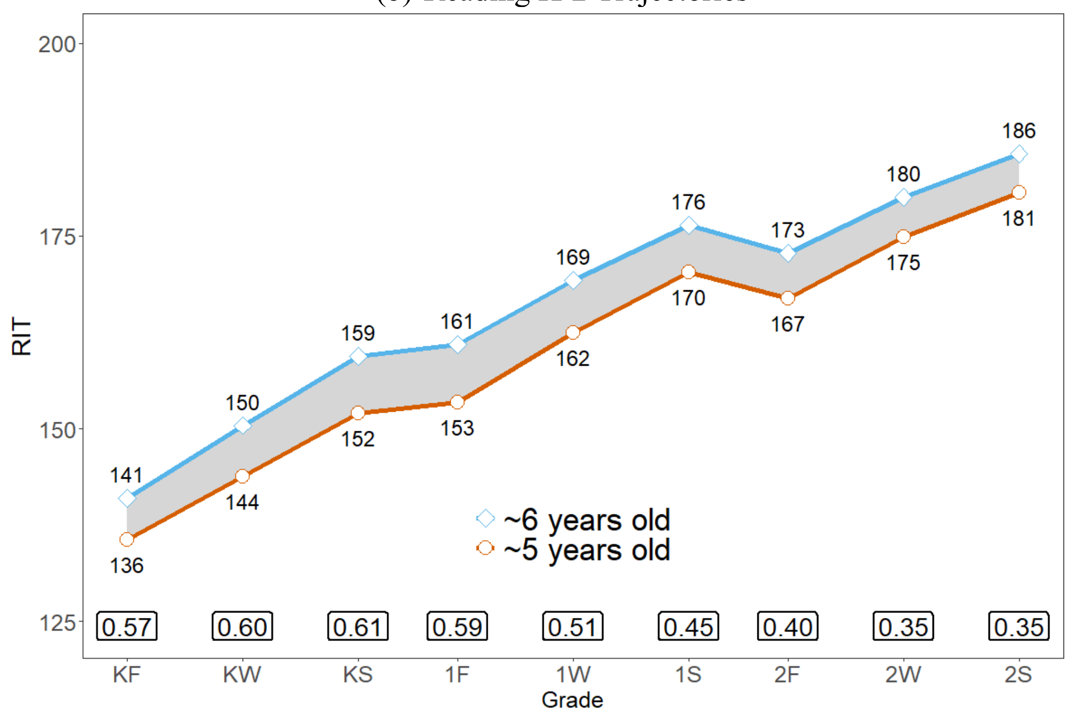


Figure 2

Math Test Score, All 3 States

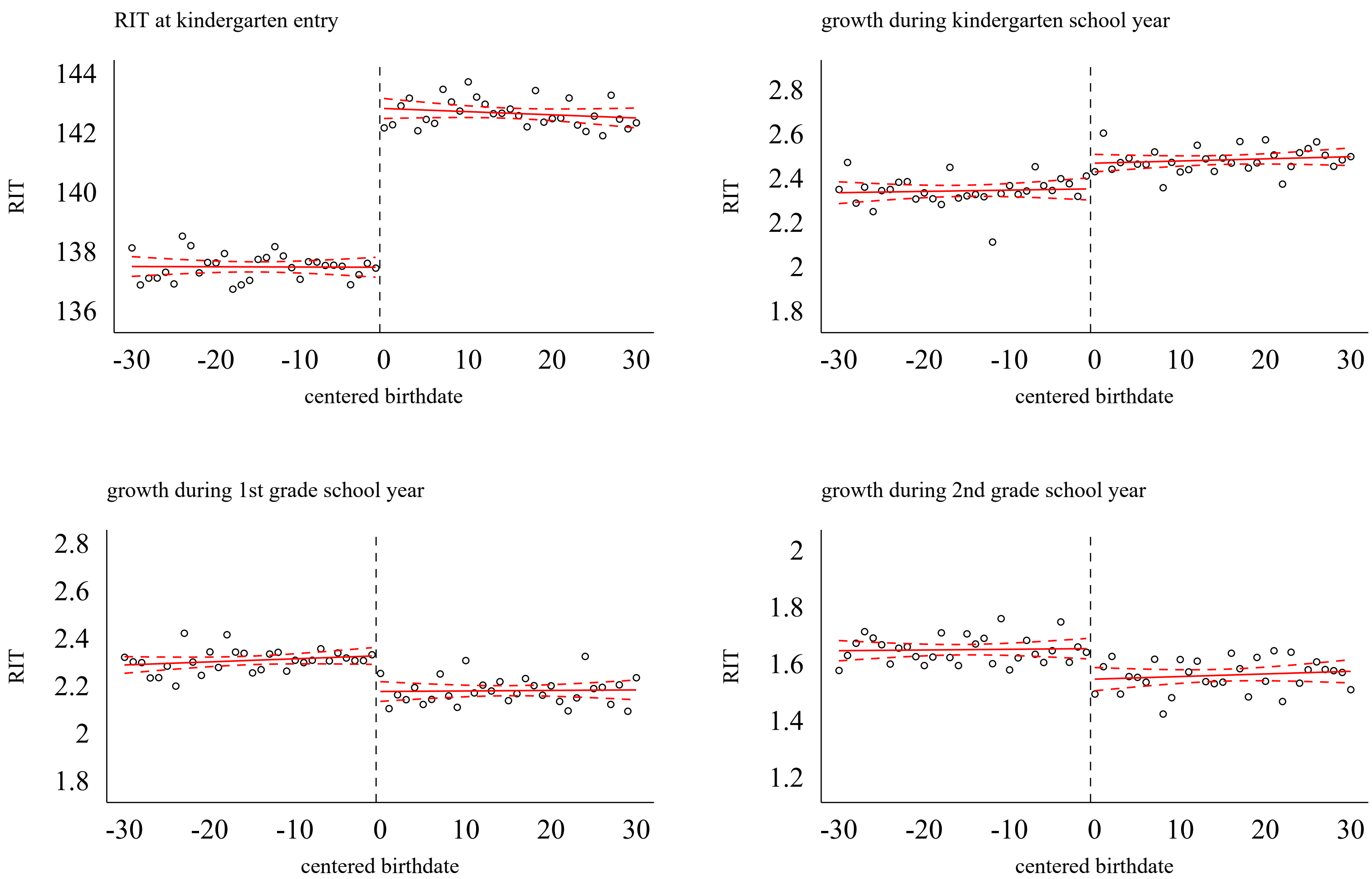
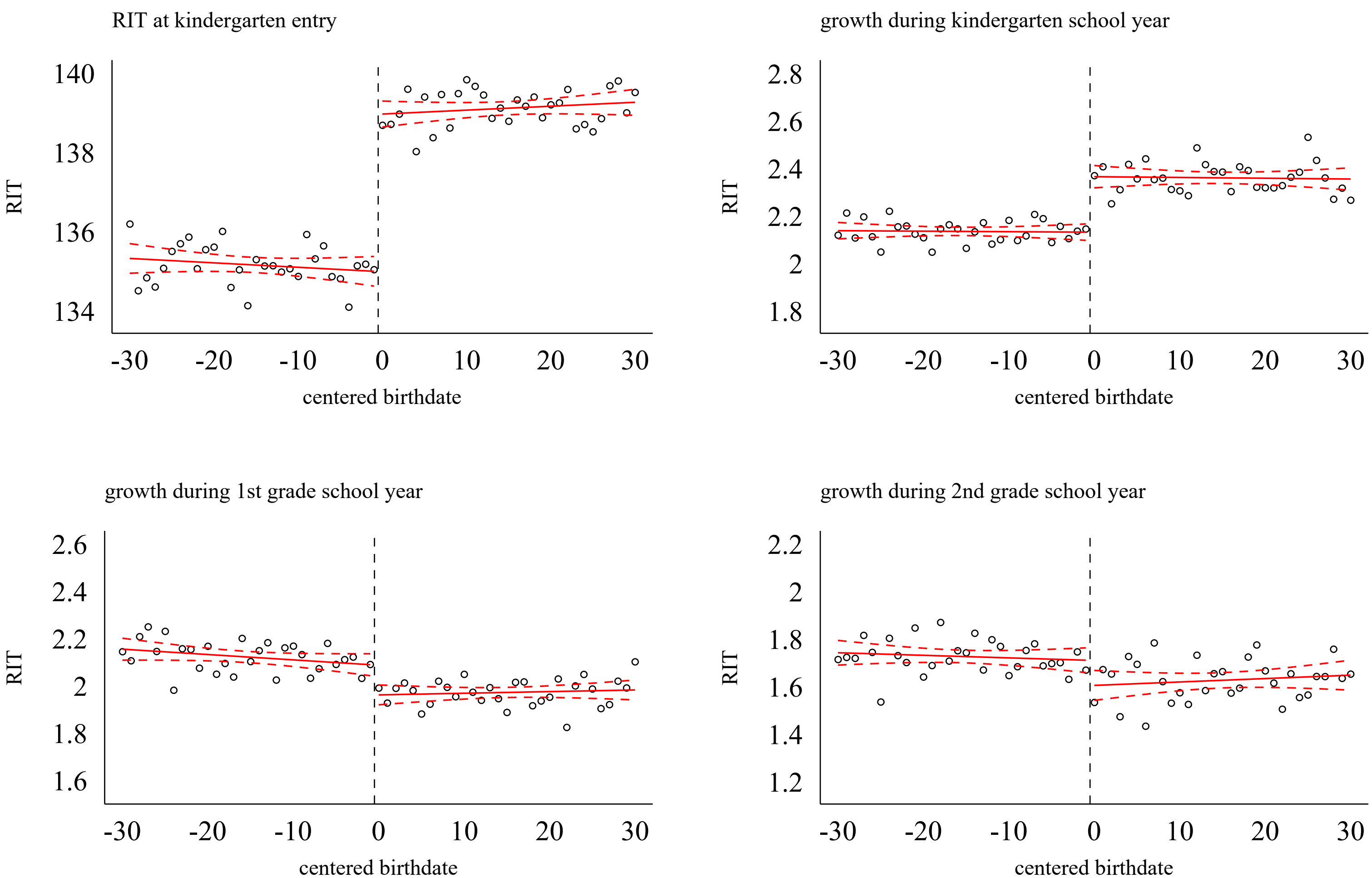


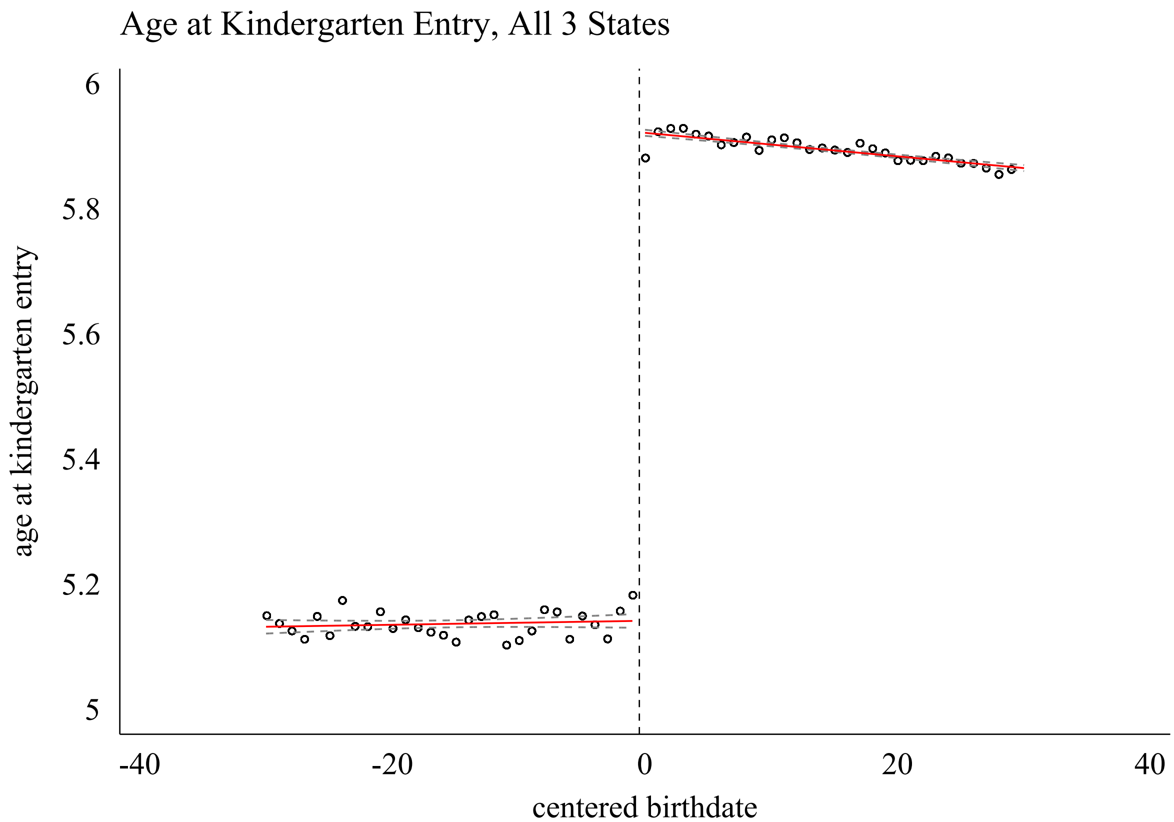
Figure 2 (continued)

Reading Test Score, All 3 States



Appendix

A1. First Stage Figure and Table



Predictor	(1) bandwidth 5	(2) bandwidth 10	(3) bandwidth 20	(4) bandwidth 30
Days ≥ 0	0.729***	0.751***	0.767***	0.780***
	(0.019)	(0.013)	(0.009)	(0.007)
Constant	5.251***	5.235***	5.212***	5.201***
	(0.020)	(0.013)	(0.009)	(0.007)
Observations	5,176	10,356	20,440	30,552
R^2	0.642	0.658	0.684	0.693
F-stat	1431	3561	7930	13303
F-stat p-value	0	0	0	0

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Each column represents a regression with bandwidth indicated in the column title. Days ≥ 0 is a binary variable indicating student was born after the date of school entry cutoff. Model also includes distance from the cutoff date, the interaction between the distance for the cutoff date and the indicator for being born after the cutoff, and student-level demographic variables.

A2. Student Characteristics and Achievement in Fall of Kindergarten

Panel A: Student and School Demographics

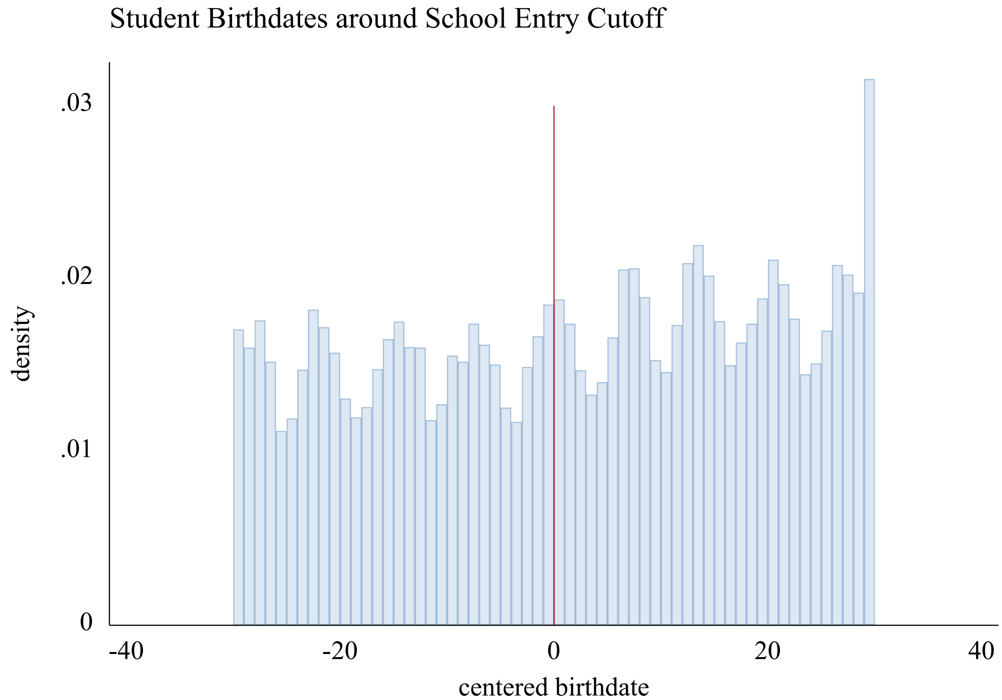
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Characteristics	Female	Black	Hispanic	Other Race	School % Black	School % White	School % FRPL
Days ≥ 0	-0.004	0.005	-0.003	-0.003	0.005	0.008	-0.005*
	(0.012)	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.003)
Observations	30,552	30,552	30,552	30,552	30,552	30,552	30,552
R^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000

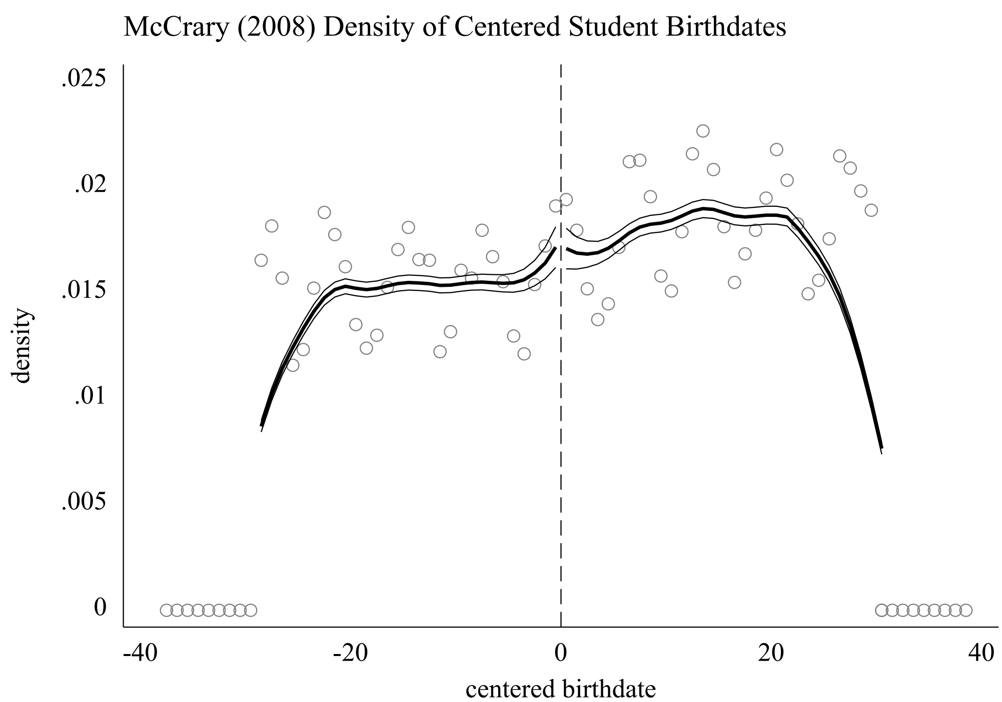
Panel B: Student Achievement in the Fall of Kindergarten

	(1)	(2)
	Math	Reading
Days ≥ 0	5.187***	4.050***
	(0.242)	(0.243)
Control Group Mean	139.52	136.73
State Dummies	yes	yes
Cohort Dummies	yes	yes
Student Covariates	no	no
School Covariates	no	no
Observations	30,128	26,172
R^2	0.061	0.044

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Days ≥0 is a binary variable indicating student was born after the date of school entry cutoff. In Panel A, each column represents a regression discontinuity model with the column title as the dependent variable, linear splines, and no additional controls. Panel A, Column (1) - (4) are student-level demographic variables with binary outcomes. Panel A, Columns (5) - (7) are characteristics of students' first school from NCES Common Core of Data in 2013-2014. In Panel B, the outcome is the observed test scores in RIT points. Consistent with Model I, in Panel B, the RD model includes linear splines, cohort dummies, and state dummies (coefficients suppressed) but no student or school covariates.

A3. Density Check Figures





Discontinuity Estimate = -0.017 (0.045)

A4. Model specification

Model 1:

Our first model estimates the monthly learning rates during each school year and summer from kindergarten to second grade. At level 1, the growth model is:

$$\text{y}_{tij} = \pi_{0ij} + \pi_{1ij}G0_{ij} + \pi_{2ij}S1_{ij} + \pi_{3ij}G1_{ij} + \pi_{4ij}S2_{ij} + \pi_{5ij}G2_{ij} + e_{tij}.$$

For details on how each of the level-1 predictors ($G0_{ij}$ through $G2_{ij}$) were calculated, see Online Appendix OA2. Each level-1 term is regressed on a set of student-level characteristics, including an indicator for having a birthdate after the state cutoff (A_{ij}), a measure of the distance between student's birthdate and the state's age cutoff date (Days_{ij}), and the interaction between these terms ($A_{ij} * \text{Days}_{ij}$), and dummy variables at level 2 for cohort ($C1_{ij}$ and $C2_{ij}$). Additionally, student-level random effects were included to allow for random intercepts and slope terms within each grade/summer.

Level-2 Model (student (i) within school (j)):

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\text{A}_{ij} + \beta_{02j}\text{Days}_{ij} + \beta_{03j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{04j}(C3_{ij}) + \beta_{05j}(C3_{ij}) + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + \beta_{11j}\text{A}_{ij} + \beta_{12j}\text{Days}_{ij} + \beta_{13j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{14j}(C3_{ij}) + \beta_{15j}(C3_{ij}) + r_{1ij}$$
$$\pi_{2ij} = \beta_{20j} + \beta_{21j}\text{A}_{ij} + \beta_{22j}\text{Days}_{ij} + \beta_{23j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{24j}(C3_{ij}) + \beta_{25j}(C3_{ij}) + r_{2ij}$$
$$\pi_{3ij} = \beta_{30j} + \beta_{31j}\text{A}_{ij} + \beta_{32j}\text{Days}_{ij} + \beta_{33j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{34j}(C3_{ij}) + \beta_{35j}(C3_{ij}) + r_{3ij}$$
$$\pi_{4ij} = \beta_{40j} + \beta_{41j}\text{A}_{ij} + \beta_{42j}\text{Days}_{ij} + \beta_{43j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{44j}(C3_{ij}) + \beta_{45j}(C3_{ij}) + r_{4ij}$$
$$\pi_{5ij} = \beta_{50j} + \beta_{51j}\text{A}_{ij} + \beta_{52j}\text{Days}_{ij} + \beta_{53j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{54j}(C3_{ij}) + \beta_{55j}(C3_{ij}) + r_{5ij}$$

At the school-level (level 3), we include state dummy variables and random effects for the intercept and growth terms. Additionally, random effects were included to allow the effect of being older A_{ij} to vary randomly between schools (e.g., random effects u_{01j} to u_{51j}). All other school-level covariates were treated as fixed.

Level-3 Model (school (j)):

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(State1_j) + \gamma_{002}(State2_j) + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + \gamma_{101}(State1_j) + \gamma_{102}(State2_j) + u_{10j}$$
$$\beta_{20j} = \gamma_{200} + \gamma_{201}(State1_j) + \gamma_{202}(State2_j) + u_{20j}$$
$$\beta_{30j} = \gamma_{300} + \gamma_{301}(State1_j) + \gamma_{302}(State2_j) + u_{30j}$$
$$\beta_{40j} = \gamma_{400} + \gamma_{401}(State1_j) + \gamma_{402}(State2_j) + u_{40j}$$
$$\beta_{50j} = \gamma_{500} + \gamma_{501}(State1_j) + \gamma_{502}(State2_j) + u_{50j}$$
$$\beta_{01j} = \gamma_{010} + u_{01j}$$
$$\beta_{11j} = \gamma_{110} + u_{11j}$$
$$\beta_{21j} = \gamma_{210} + u_{21j}$$
$$\beta_{31j} = \gamma_{310} + u_{31j}$$
$$\beta_{41j} = \gamma_{410} + u_{41j}$$
$$\beta_{51j} = \gamma_{510} + u_{51j}$$
$$\beta_{11j} = \gamma_{110}$$
$$\vdots$$
$$\beta_{55j} = \gamma_{550}$$

Model 2:

Model 2 expands upon Model 1 by adding student-level gender and race/ethnicity indicators.

Level-2 Model (student (i) within school (j)):

$$y_{tij} = \pi_{0ij} + \pi_{1ij}G0_{ij} + \pi_{2ij}S1_{ij} + \pi_{3ij}G1_{ij} + \pi_{4ij}S2_{ij} + \pi_{5ij}G2_{ij} + e_{tij}.$$

Level-2 Model (student (i) within school (j)):

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\mathrm{A}_{ij} + \beta_{02j}\mathrm{Days}_{ij} + \beta_{03j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{04j}(C2_{ij}) + \beta_{05j}(C3_{ij}) + \beta_{06j}(Female_{ij}) + \beta_{07}(Black_{ij}) + \beta_{08j}(Hispanic_{ij}) + \beta_{09j}(OtherRace_{ij}) + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + \beta_{11j}\mathrm{A}_{ij} + \beta_{12j}\mathrm{Days}_{ij} + \beta_{13j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{14j}(C2_{ij}) + \beta_{15j}(C3_{ij}) + \beta_{16j}(Female_{ij}) + \beta_{17}(Black_{ij}) + \beta_{18j}(Hispanic_{ij}) + \beta_{19j}(OtherRace_{ij}) + r_{1ij}$$
$$\pi_{2ij} = \beta_{20j} + \beta_{21j}\mathrm{A}_{ij} + \beta_{22j}\mathrm{Days}_{ij} + \beta_{23j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{24j}(C2_{ij}) + \beta_{25j}(C3_{ij}) + \beta_{26j}(Female_{ij}) + \beta_{27}(Black_{ij}) + \beta_{28j}(Hispanic_{ij}) + \beta_{29j}(OtherRace_{ij}) + r_{2ij}$$
$$\pi_{3ij} = \beta_{30j} + \beta_{31j}\mathrm{A}_{ij} + \beta_{32j}\mathrm{Days}_{ij} + \beta_{33j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{34j}(C2_{ij}) + \beta_{35j}(C3_{ij}) + \beta_{36j}(Female_{ij}) + \beta_{37}(Black_{ij}) + \beta_{38j}(Hispanic_{ij}) + \beta_{39j}(OtherRace_{ij}) + r_{3ij}$$
$$\pi_{4ij} = \beta_{40j} + \beta_{41j}\mathrm{A}_{ij} + \beta_{42j}\mathrm{Days}_{ij} + \beta_{43j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{44j}(C2_{ij}) + \beta_{45j}(C3_{ij}) + \beta_{46j}(Female_{ij}) + \beta_{47}(Black_{ij}) + \beta_{48j}(Hispanic_{ij}) + \beta_{49j}(OtherRace_{ij}) + r_{4ij}$$
$$\pi_{5ij} = \beta_{50j} + \beta_{51j}\mathrm{A}_{ij} + \beta_{52j}\mathrm{Days}_{ij} + \beta_{53j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{54j}(C2_{ij}) + \beta_{55j}(C3_{ij}) + \beta_{56j}(Female_{ij}) + \beta_{57}(Black_{ij}) + \beta_{58j}(Hispanic_{ij}) + \beta_{59j}(OtherRace_{ij}) + r_{5ij}$$

Level-3 Model (school (j)):

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(State1_j) + \gamma_{002}(State2_j) + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + \gamma_{101}(State1_j) + \gamma_{102}(State2_j) + u_{10j}$$
$$\beta_{20j} = \gamma_{200} + \gamma_{201}(State1_j) + \gamma_{202}(State2_j) + u_{20j}$$
$$\beta_{30j} = \gamma_{300} + \gamma_{301}(State1_j) + \gamma_{302}(State2_j) + u_{30j}$$
$$\beta_{40j} = \gamma_{400} + \gamma_{401}(State1_j) + \gamma_{402}(State2_j) + u_{40j}$$
$$\beta_{50j} = \gamma_{500} + \gamma_{501}(State1_j) + \gamma_{502}(State2_j) + u_{50j}$$
$$\beta_{01j} = \gamma_{010} + u_{01j}$$
$$\beta_{11j} = \gamma_{110} + u_{11j}$$
$$\beta_{21j} = \gamma_{210} + u_{21j}$$
$$\beta_{31j} = \gamma_{310} + u_{31j}$$
$$\beta_{41j} = \gamma_{410} + u_{41j}$$
$$\beta_{51j} = \gamma_{510} + u_{51j}$$
$$\beta_{11j} = \gamma_{110}$$
$$\vdots$$
$$\beta_{59j} = \gamma_{590}$$

Model 3:

Model 3 expands upon Model 2 by adding a set of school characteristics at level 3.

Level-2 Model (student (i) within school (j)):

$$y_{tij} = \pi_{0ij} + \pi_{1ij}G0_{ij} + \pi_{2ij}S1_{ij} + \pi_{3ij}G1_{ij} + \pi_{4ij}S2_{ij} + \pi_{5ij}G2_{ij} + e_{tij}.$$

Level-2 Model (student (i) within school (j)):

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\text{A}_{ij} + \beta_{02j}\text{Days}_{ij} + \beta_{03j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{04j}(C2_{ij}) + \beta_{05j}(C3_{ij}) + \beta_{06j}(Female_{ij}) + \beta_{07}(Black_{ij}) + \beta_{08j}(Hispanic_{ij}) + \beta_{09j}(OtherRace_{ij}) + r_{0ij}$$
$$\pi_{1ij} = \beta_{10j} + \beta_{11j}\text{A}_{ij} + \beta_{12j}\text{Days}_{ij} + \beta_{13j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{14j}(C2_{ij}) + \beta_{15j}(C3_{ij}) + \beta_{16j}(Female_{ij}) + \beta_{17}(Black_{ij}) + \beta_{18j}(Hispanic_{ij}) + \beta_{19j}(OtherRace_{ij}) + r_{1ij}$$
$$\pi_{2ij} = \beta_{20j} + \beta_{21j}\text{A}_{ij} + \beta_{22j}\text{Days}_{ij} + \beta_{23j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{24j}(C2_{ij}) + \beta_{25j}(C3_{ij}) + \beta_{26j}(Female_{ij}) + \beta_{27}(Black_{ij}) + \beta_{28j}(Hispanic_{ij}) + \beta_{29j}(OtherRace_{ij}) + r_{2ij}$$
$$\pi_{3ij} = \beta_{30j} + \beta_{31j}\text{A}_{ij} + \beta_{32j}\text{Days}_{ij} + \beta_{33j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{34j}(C2_{ij}) + \beta_{35j}(C3_{ij}) + \beta_{36j}(Female_{ij}) + \beta_{37}(Black_{ij}) + \beta_{38j}(Hispanic_{ij}) + \beta_{39j}(OtherRace_{ij}) + r_{3ij}$$
$$\pi_{4ij} = \beta_{40j} + \beta_{41j}\text{A}_{ij} + \beta_{42j}\text{Days}_{ij} + \beta_{43j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{44j}(C2_{ij}) + \beta_{45j}(C3_{ij}) + \beta_{46j}(Female_{ij}) + \beta_{47}(Black_{ij}) + \beta_{48j}(Hispanic_{ij}) + \beta_{49j}(OtherRace_{ij}) + r_{4ij}$$
$$\pi_{5ij} = \beta_{50j} + \beta_{51j}\text{A}_{ij} + \beta_{52j}\text{Days}_{ij} + \beta_{53j}(\text{A}_{ij} * \text{Days}_{ij}) + \beta_{54j}(C2_{ij}) + \beta_{55j}(C3_{ij}) + \beta_{56j}(Female_{ij}) + \beta_{57}(Black_{ij}) + \beta_{58j}(Hispanic_{ij}) + \beta_{59j}(OtherRace_{ij}) + r_{5ij}$$

Level-3 Model (school (j)):

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(State1_j) + \gamma_{002}(State2_j) + \gamma_{003}(\%FRPL_j) + \gamma_{004}(\%White_j) + \gamma_{005}(\%Black_j) + u_{00j}$$
$$\beta_{10j} = \gamma_{100} + \gamma_{101}(State1_j) + \gamma_{102}(State2_j) + \gamma_{103}(\%FRPL_j) + \gamma_{104}(\%White_j) + \gamma_{105}(\%Black_j) + u_{10j}$$
$$\beta_{20j} = \gamma_{200} + \gamma_{201}(State1_j) + \gamma_{202}(State2_j) + \gamma_{203}(\%FRPL_j) + \gamma_{204}(\%White_j) + \gamma_{205}(\%Black_j) + u_{20j}$$
$$\beta_{30j} = \gamma_{300} + \gamma_{301}(State1_j) + \gamma_{302}(State2_j) + \gamma_{303}(\%FRPL_j) + \gamma_{304}(\%White_j) + \gamma_{305}(\%Black_j) + u_{30j}$$
$$\beta_{40j} = \gamma_{400} + \gamma_{401}(State1_j) + \gamma_{402}(State2_j) + \gamma_{403}(\%FRPL_j) + \gamma_{404}(\%White_j) + \gamma_{405}(\%Black_j) + u_{40j}$$
$$\beta_{50j} = \gamma_{500} + \gamma_{501}(State1_j) + \gamma_{502}(State2_j) + \gamma_{503}(\%FRPL_j) + \gamma_{504}(\%White_j) + \gamma_{505}(\%Black_j) + u_{50j}$$
$$\beta_{01j} = \gamma_{010} + u_{01j}$$
$$\beta_{11j} = \gamma_{110} + u_{11j}$$
$$\beta_{21j} = \gamma_{210} + u_{21j}$$
$$\beta_{31j} = \gamma_{310} + u_{31j}$$
$$\beta_{41j} = \gamma_{410} + u_{41j}$$
$$\beta_{51j} = \gamma_{510} + u_{51j}$$
$$\beta_{11j} = \gamma_{110}$$
$$\vdots$$
$$\beta_{59j} = \gamma_{590}$$

A5. Third Grade Results for Cohort 2009

	Math	Reading
Control Group Fall K Score	139.34 (0.40)	136.48 (0.43)
A	4.25 (0.45)	3.20 (0.48)
Control Group K Growth	2.24 (0.04)	2.12 (0.05)
A	0.21 (0.05)	0.24 (0.06)
Control Group Summer After K Growth	-1.21 (0.12)	-0.94 (0.13)
A	-0.05 (0.14)	0.00 (0.17)
Control Group 1st Grade Growth	2.28 (0.04)	2.18 (0.04)
A	-0.14 (0.05)	-0.17 (0.06)
Control Group Summer After 1st Growth	-2.78 (0.14)	-2.19 (0.16)
A	-0.53 (0.16)	-0.27 (0.20)
Control Group 2nd Grade Growth	1.66 (0.04)	1.60 (0.06)
A	-0.11 (0.04)	0.02 (0.07)
Control Group Summer After 2nd Growth	-2.12 (0.12)	-1.36 (0.15)
A	0.09 (0.14)	-0.03 (0.18)
Control Group 3rd Grade Growth	1.58 (0.03)	1.42 (0.04)
A	-0.16 (0.04)	-0.21 (0.06)

Notes: Robust standard errors in parentheses. Estimation uses Model I, which includes state dummies. Each column is a separate regression. Coefficient of secondary interest are suppressed for brevity. Sample is restricted to students born in calendar year 2009. Control group represents students who entered kindergarten close to five years old. A is the estimate for the impact of being a year older.

Online Appendix

OA1. Model Specification

Models 4-6 expands upon Model 2 by adding a set of interaction terms at level 2. We illustrate Model 5 (gender interactions) below but the logic holds for Models 4 & 6 as well.

Level-2 Model (student (i) within school (j)):

$$\mathrm{y}_{tij} = \pi_{0ij} + \pi_{1ij}G0_{ij} + \pi_{2ij}S1_{ij} + \pi_{3ij}G1_{ij} + \pi_{4ij}S2_{ij} + \pi_{5ij}G2_{ij} + e_{tij}.$$

Level-2 Model (student (i) within school (j)):

$$\begin{aligned}
\pi_{0ij} &= \beta_{00j} + \beta_{01j}\mathrm{A}_{ij} + \beta_{02j}\mathrm{Days}_{ij} + \beta_{03j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{04j}(C2_{ij}) + \beta_{05j}(C3_{ij}) + \beta_{06j}(Female_{ij}) \\
&\quad + \beta_{07}(Black_{ij}) + \beta_{08j}(Hispanic_{ij}) + \beta_{09j}(OtherRace_{ij}) + \beta_{010j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{0ij} \\
\pi_{1ij} &= \beta_{10j} + \beta_{11j}\mathrm{A}_{ij} + \beta_{12j}\mathrm{Days}_{ij} + \beta_{13j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{14j}(C2_{ij}) + \beta_{15j}(C3_{ij}) + \beta_{16j}(Female_{ij}) \\
&\quad + \beta_{17}(Black_{ij}) + \beta_{18j}(Hispanic_{ij}) + \beta_{19j}(OtherRace_{ij}) + \beta_{110j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{1ij} \\
\pi_{2ij} &= \beta_{20j} + \beta_{21j}\mathrm{A}_{ij} + \beta_{22j}\mathrm{Days}_{ij} + \beta_{23j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{24j}(C2_{ij}) + \beta_{25j}(C3_{ij}) + \beta_{26j}(Female_{ij}) + \\
&\quad \beta_{27}(Black_{ij}) + \beta_{28j}(Hispanic_{ij}) + \beta_{29j}(OtherRace_{ij}) + \beta_{210j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{2ij} \\
\pi_{3ij} &= \beta_{30j} + \beta_{31j}\mathrm{A}_{ij} + \beta_{32j}\mathrm{Days}_{ij} + \beta_{33j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{34j}(C2_{ij}) + \beta_{35j}(C3_{ij}) + \beta_{36j}(Female_{ij}) \\
&\quad + \beta_{37}(Black_{ij}) + \beta_{38j}(Hispanic_{ij}) + \beta_{39j}(OtherRace_{ij}) + \beta_{310j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{3ij} \\
\pi_{4ij} &= \beta_{40j} + \beta_{41j}\mathrm{A}_{ij} + \beta_{42j}\mathrm{Days}_{ij} + \beta_{43j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{44j}(C2_{ij}) + \beta_{45j}(C3_{ij}) + \beta_{46j}(Female_{ij}) \\
&\quad + \beta_{47}(Black_{ij}) + \beta_{48j}(Hispanic_{ij}) + \beta_{49j}(OtherRace_{ij}) + \beta_{410j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{4ij} \\
\pi_{5ij} &= \beta_{50j} + \beta_{51j}\mathrm{A}_{ij} + \beta_{52j}\mathrm{Days}_{ij} + \beta_{53j}(\mathrm{A}_{ij} * \mathrm{Days}_{ij}) + \beta_{54j}(C2_{ij}) + \beta_{55j}(C3_{ij}) + \beta_{56j}(Female_{ij}) \\
&\quad + \beta_{57}(Black_{ij}) + \beta_{58j}(Hispanic_{ij}) + \beta_{59j}(OtherRace_{ij}) + \beta_{510j}(\mathrm{Days}_{ij} * Female_{ij}) + r_{5ij}
\end{aligned}$$

Level-3 Model (school (j)):

$$\begin{aligned}
\beta_{00j} &= \gamma_{000} + \gamma_{001}(State1_j) + \gamma_{002}(State2_j) + \gamma_{003}(\%FRPL_j) + \gamma_{004}(\%White_j) + \gamma_{005}(\%Black_j) + u_{00j} \\
\beta_{10j} &= \gamma_{100} + \gamma_{101}(State1_j) + \gamma_{102}(State2_j) + \gamma_{103}(\%FRPL_j) + \gamma_{104}(\%White_j) + \gamma_{105}(\%Black_j) + u_{10j} \\
\beta_{20j} &= \gamma_{200} + \gamma_{201}(State1_j) + \gamma_{202}(State2_j) + \gamma_{203}(\%FRPL_j) + \gamma_{204}(\%White_j) + \gamma_{205}(\%Black_j) + u_{20j} \\
\beta_{30j} &= \gamma_{300} + \gamma_{301}(State1_j) + \gamma_{302}(State2_j) + \gamma_{303}(\%FRPL_j) + \gamma_{304}(\%White_j) + \gamma_{305}(\%Black_j) + u_{30j} \\
\beta_{40j} &= \gamma_{400} + \gamma_{401}(State1_j) + \gamma_{402}(State2_j) + \gamma_{403}(\%FRPL_j) + \gamma_{404}(\%White_j) + \gamma_{405}(\%Black_j) + u_{40j} \\
\beta_{50j} &= \gamma_{500} + \gamma_{501}(State1_j) + \gamma_{502}(State2_j) + \gamma_{503}(\%FRPL_j) + \gamma_{504}(\%White_j) + \gamma_{505}(\%Black_j) + u_{50j}
\end{aligned}$$

$$\begin{aligned}
\beta_{01j} &= \gamma_{010} + u_{01j} \\
\beta_{11j} &= \gamma_{110} + u_{11j} \\
\beta_{21j} &= \gamma_{210} + u_{21j} \\
\beta_{31j} &= \gamma_{310} + u_{31j} \\
\beta_{41j} &= \gamma_{410} + u_{41j} \\
\beta_{51j} &= \gamma_{510} + u_{51j} \\
\beta_{11j} &= \gamma_{110} \\
&\vdots \\
\beta_{59j} &= \gamma_{590}
\end{aligned}$$

OA2. Calculating months of exposure to school

To set up the design matrix for this seasonal learning model, we needed to calculate three sets of time variables: (a) number of months in school prior to testing, (b) total number of months spent in school across the whole school year, and (c) months of summer vacation. Time before testing was calculated as the difference between the school start date and test administration date for each student. The number of weeks before testing ranged from zero to 17 weeks in the fall and 18 to 36 weeks in the spring, with the average student testing in week four in the fall and week 33 in the spring. The total number of months in school is calculated as the end date subtracted by the school start date. The months of summer vacation is the fall school start date subtracted by the prior year spring end date. These three sets of values are used to fill in the design matrix (as shown in Table A10). For example, if a student tests in the fall of first grade, he or she has been exposed to all of kindergarten (typically 9.5 months), a couple months of summer vacation after kindergarten, and one or two months of first grade. Since he or she has not been exposed to another summer vacation or 2nd grade, the values for those predictors are set to zero

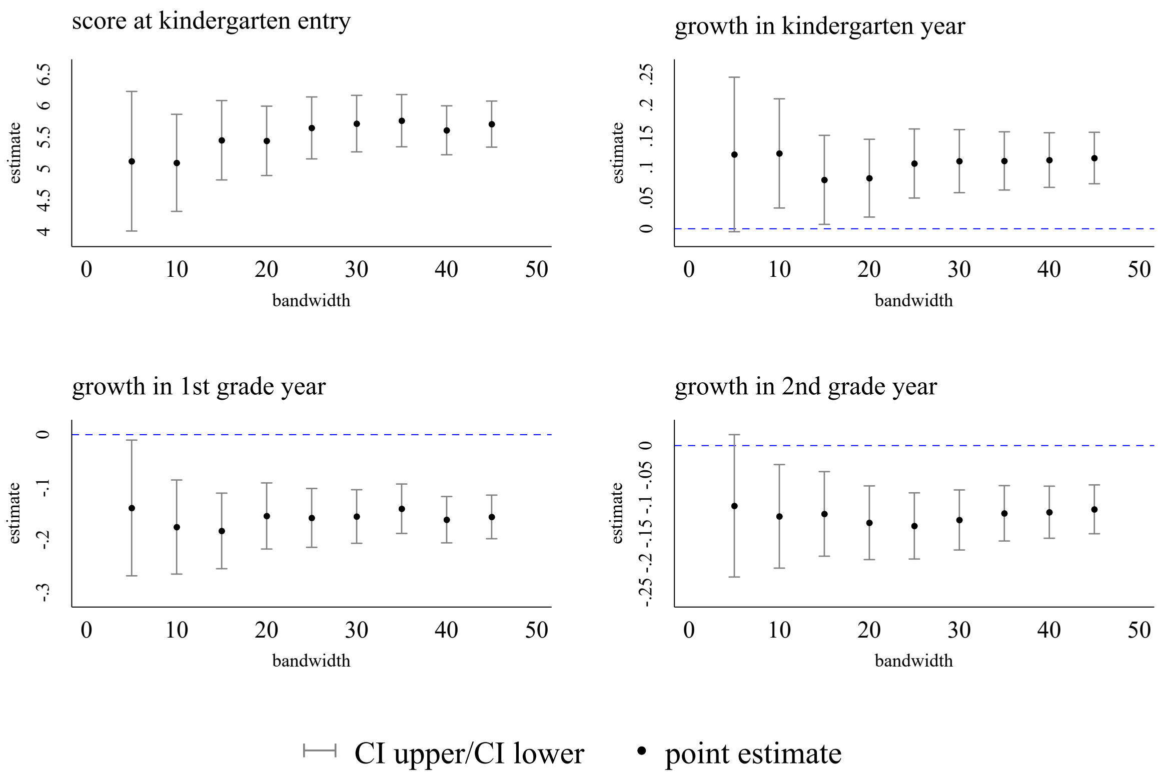
Table OA2

Coding of Monthly Exposure Rates for an Average Student Testing Between Kindergarten and 2nd Grade

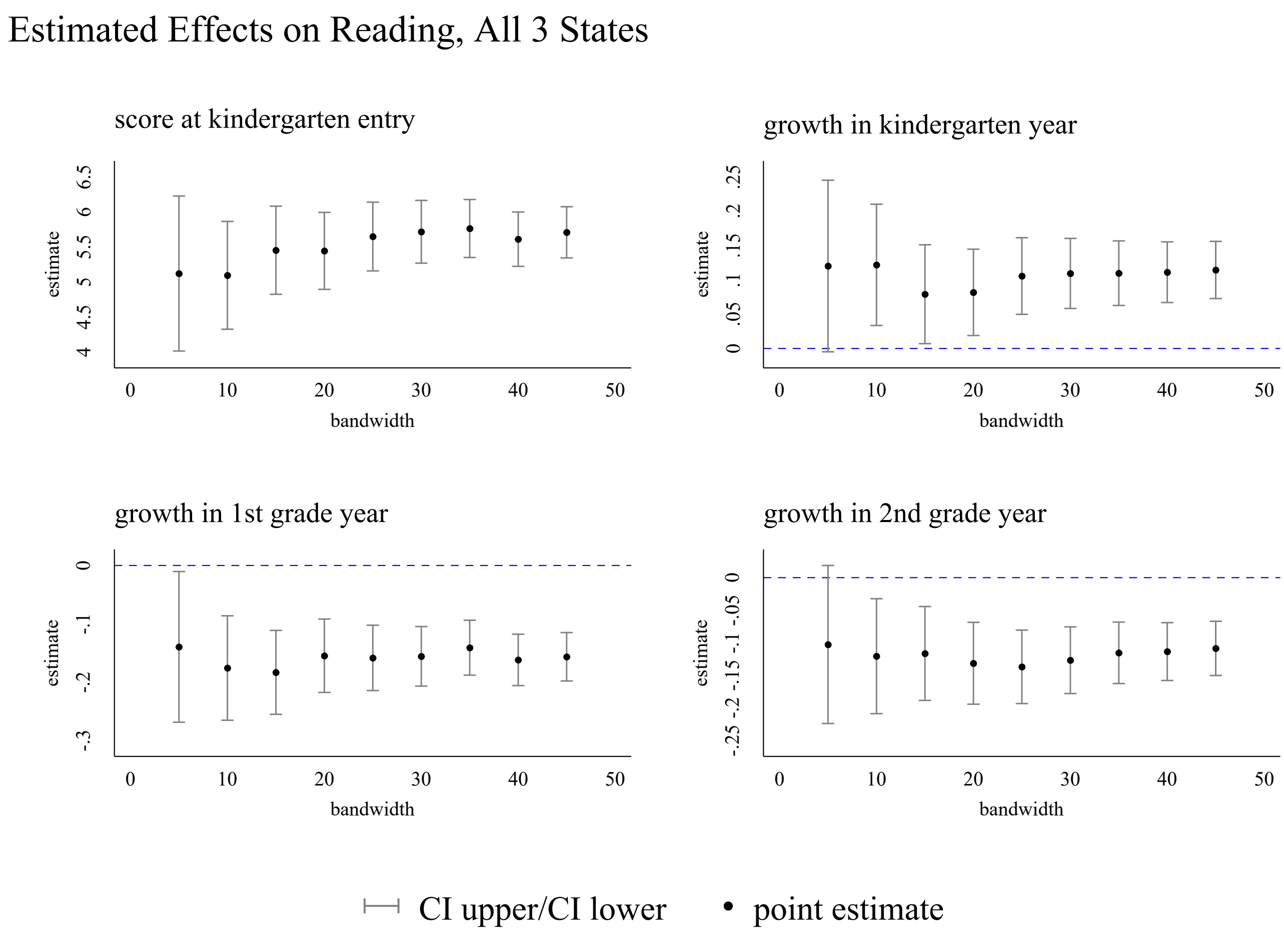
Grade/Term	School Start Date	School End Date	Test date	Exposure to Summer	Exposure to School Year Prior to Testing	Exposure to Full School Year	Monthly Exposure Design Matrix: Int.	G0	Sum1	G1	Sum2	G2
Fall K	8/20/2015	6/12/2016	9/1/2015	NA	0.39	9.58	1.00	0.39	0.00	0.00	0.00	0.00
Spring K	8/20/2015	6/12/2016	5/1/2016	NA	8.23	9.58	1.00	8.23	0.00	0.00	0.00	0.00
Fall 1st	8/8/2016	6/4/2017	10/1/2016	1.84	1.74	9.68	1.00	9.58	1.84	1.74	0.00	0.00
Spring 1st	8/8/2016	6/4/2017	4/1/2017	1.84	7.61	9.68	1.00	9.58	1.84	7.61	0.00	0.00
Fall 2nd	8/14/2017	6/5/2018	9/5/2017	2.29	0.71	9.33	1.00	9.58	1.84	9.68	2.29	0.71
Spring 2nd	8/14/2017	6/5/2018	5/15/2018	2.29	8.84	9.33	1.00	9.58	1.84	9.68	2.29	8.84

OA3. Robustness across Bandwidths

Estimated Effects on Math, All 3 States

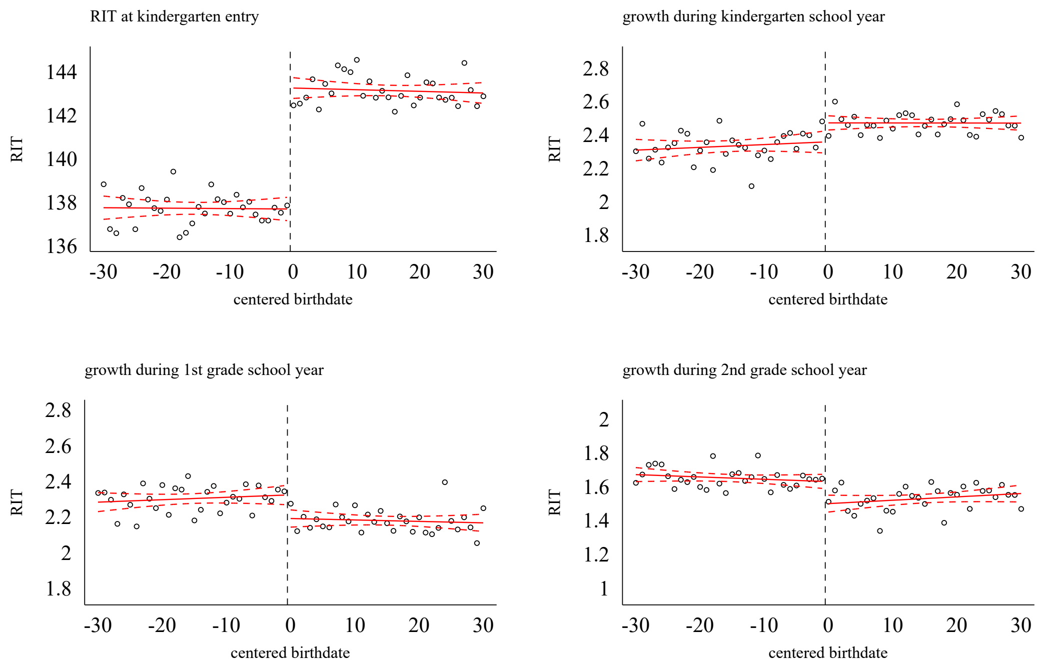


OA3. Robustness across Bandwidths (continued)

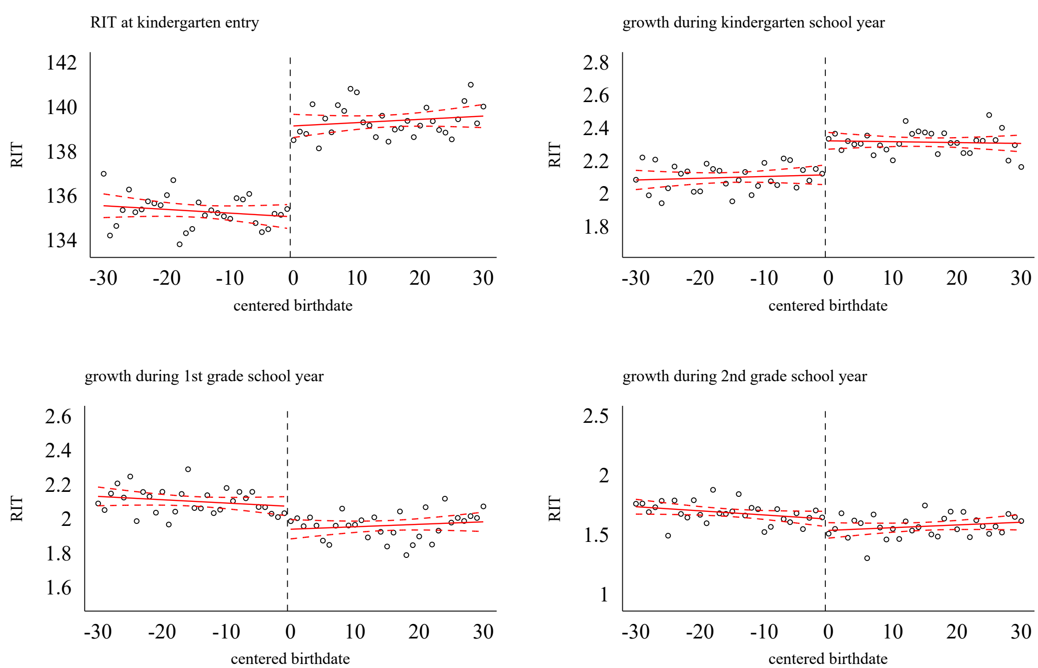


OA4. Discontinuity Figures by State

Math Test Score, State 1

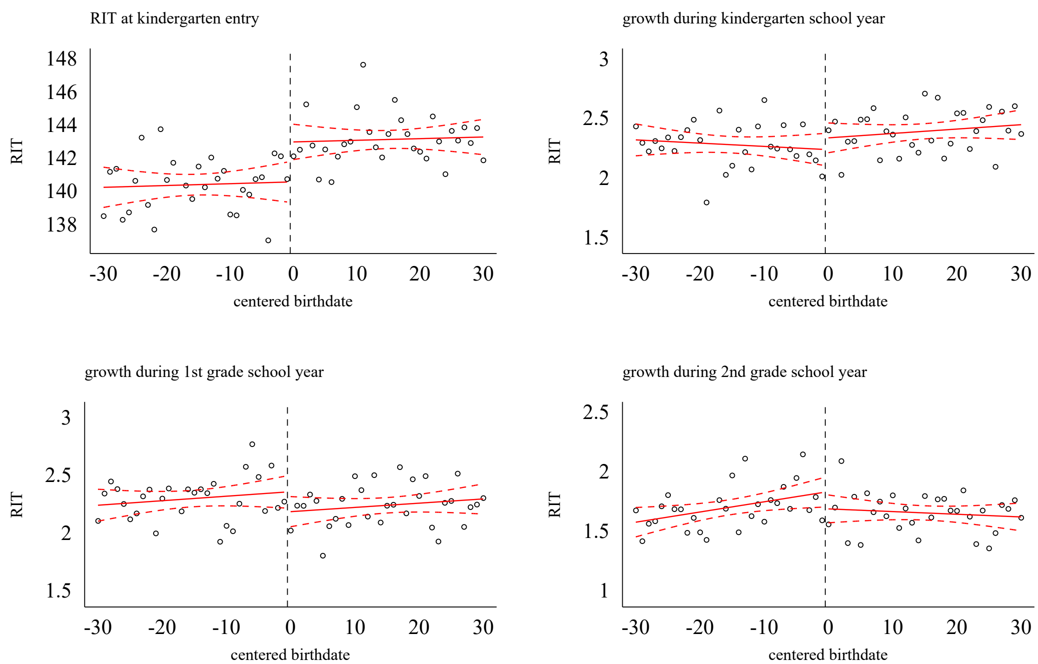


Reading Test Score, State 1

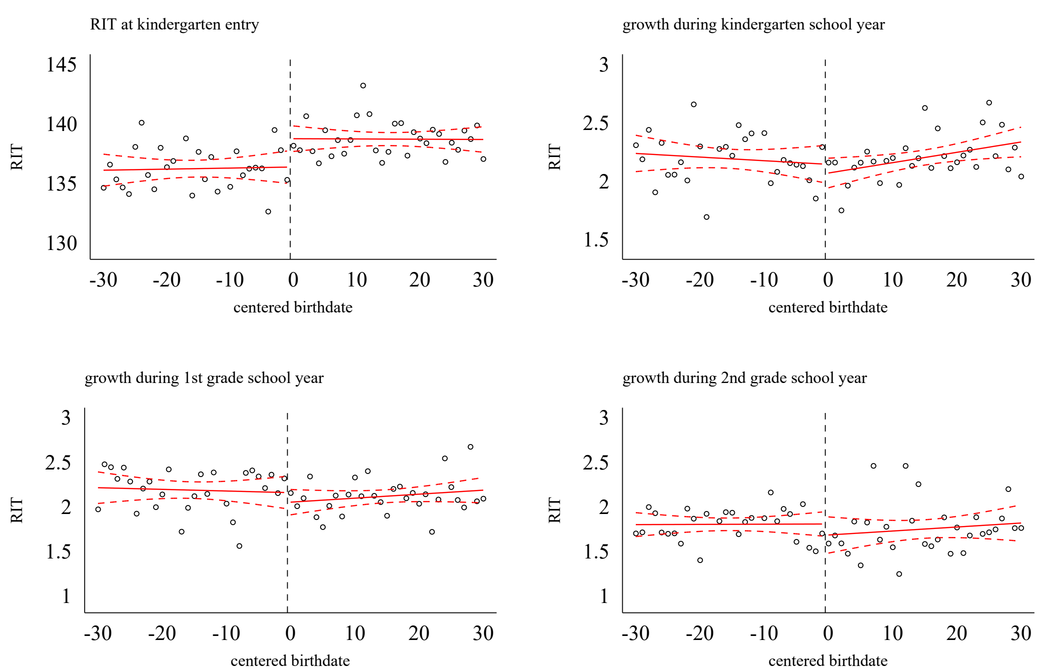


OA4. Discontinuity Figures by State (continued)

Math Test Score, State 2

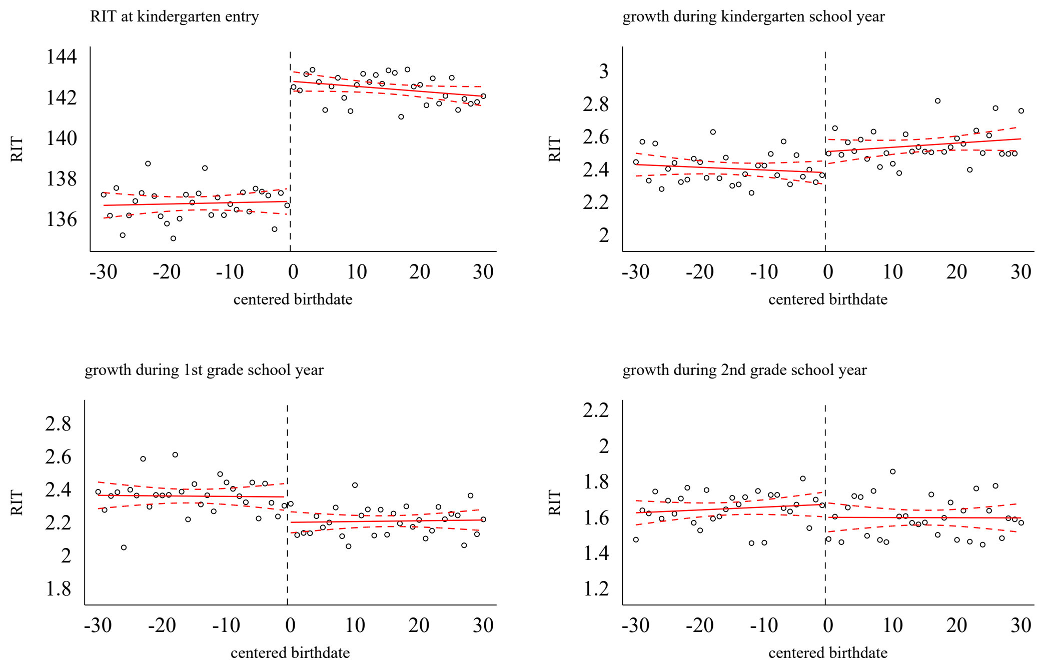


Reading Test Score, State 2

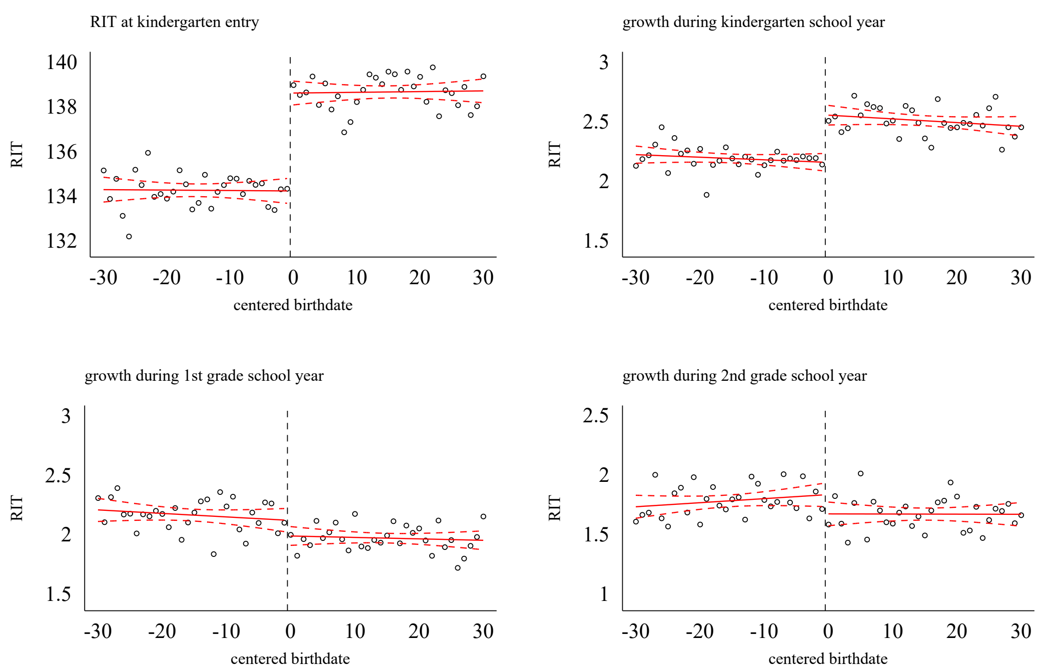


OA4. Discontinuity Figures by State (continued)

Math Test Score, State 3

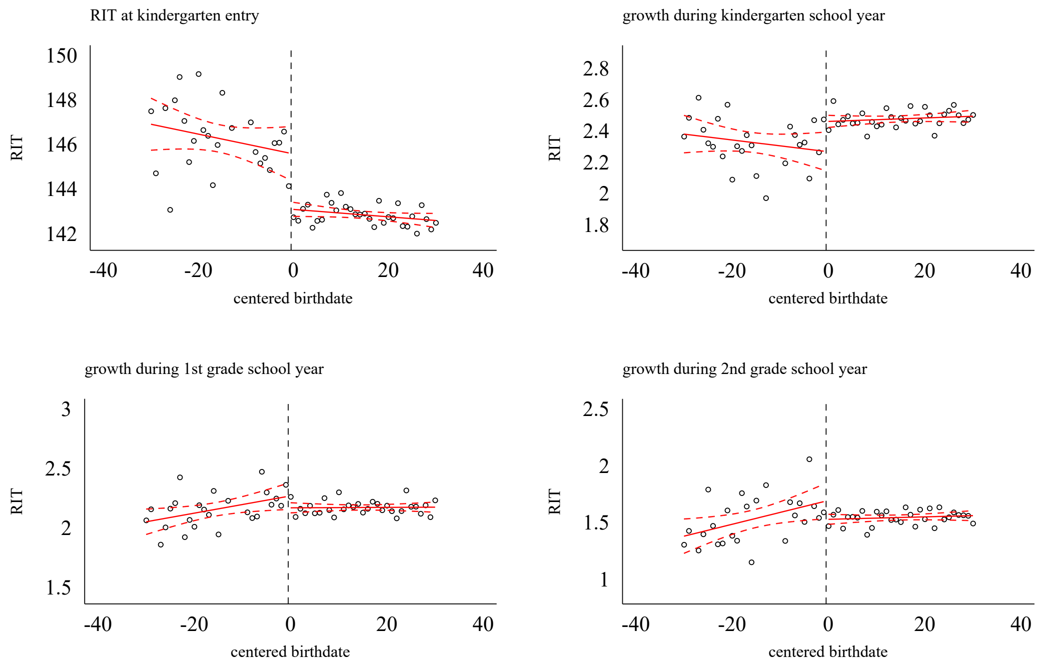


Reading Test Score, State 3

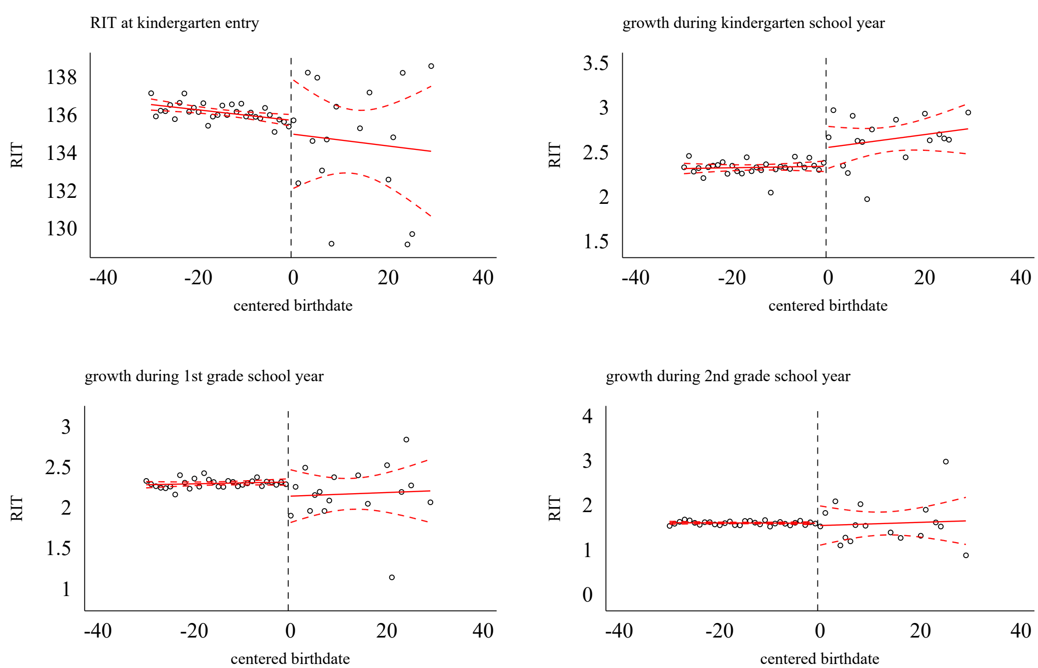


OA5. Effects Heterogeneity

Compliers and Always-takers vs. Always-takers, Math

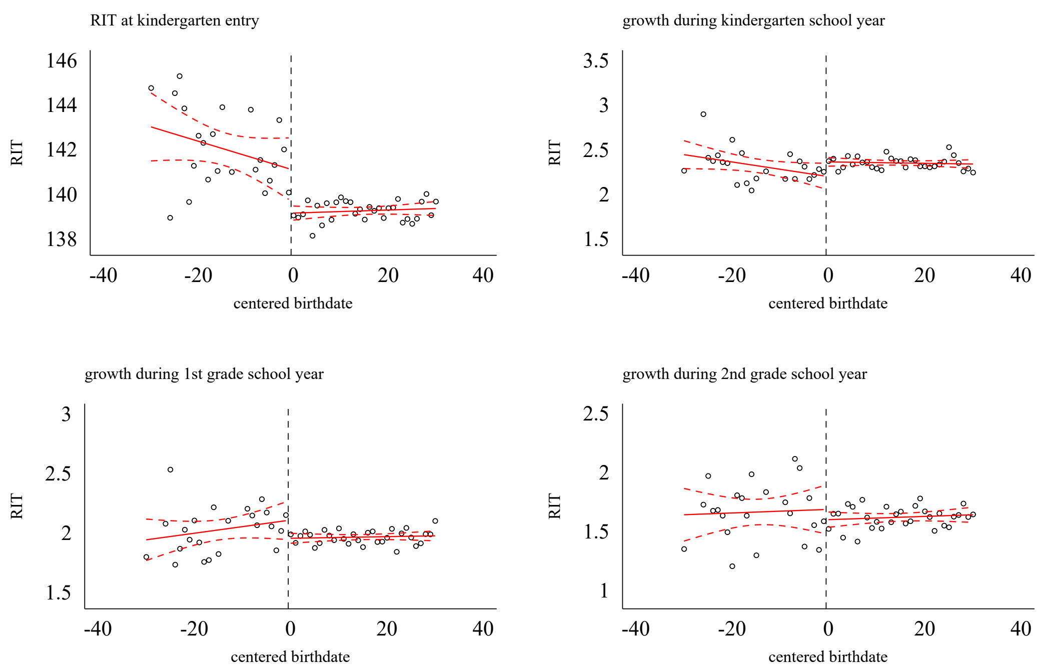


Compliers and Never-takers vs. Never-takers, Math

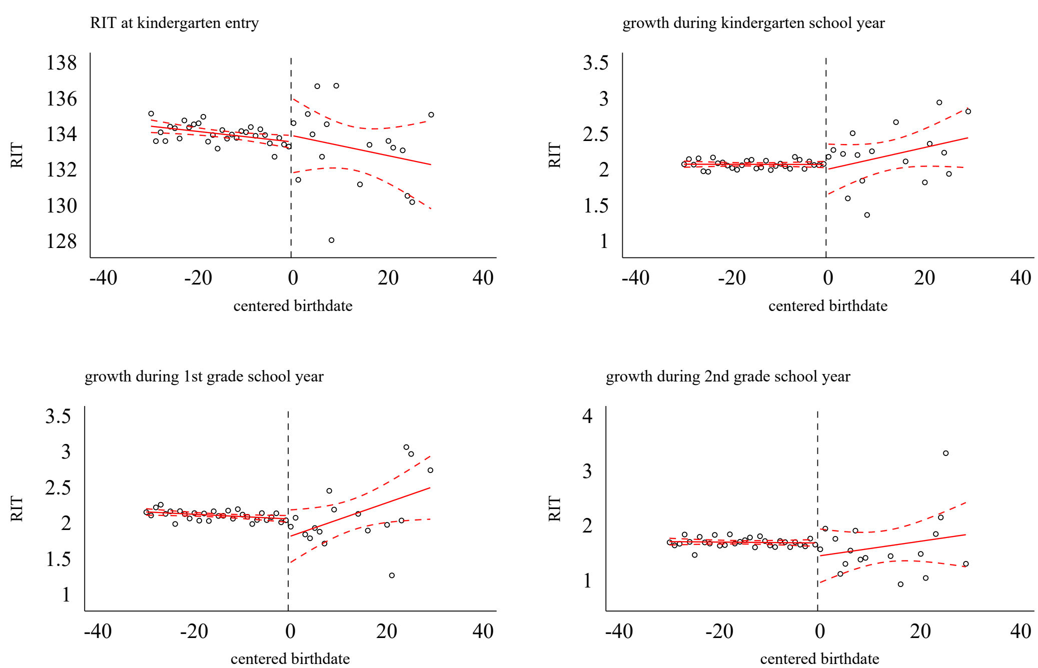


OA5. Effects Heterogeneity (continued)

Compliers and Always-takers vs. Always-takers, Reading



Compliers and Never-takers vs. Never-takers, Reading



OA6. Summer learning loss estimates for Models 1-6

Summer learning loss estimates corresponding to the HLM results presented in Table 3

	Math			Reading		
Variable	(I)	(II)	(III)	(I)	(II)	(III)
			Growth in Summer after Kindergarten			
A	**-0.14 (0.08)**	**-0.15 (0.08)**	**-0.15 (0.08)**	**-0.11 (0.09)**	**-0.11 (0.09)**	**-0.11 (0.09)**
Control Group	-1.47 (0.09)	-1.59 (0.09)	-1.61 (0.10)	-1.17 (0.10)	-1.30 (0.10)	-1.33 (0.11)
State 1	0.39 (0.07)	0.37 (0.07)	0.29 (0.08)	0.48 (0.07)	0.49 (0.08)	0.43 (0.09)
State 2	1.08 (0.10)	1.07 (0.10)	0.96 (0.11)	0.82 (0.09)	0.84 (0.10)	0.76 (0.10)
C2	-0.06 (0.05)	-0.06 (0.05)	-0.06 (0.05)	0.12 (0.06)	0.11 (0.06)	0.11 (0.06)
C3	-0.42 (0.06)	-0.42 (0.06)	-0.42 (0.06)	-0.13 (0.06)	-0.13 (0.06)	-0.13 (0.06)
Female		0.16 (0.04)	0.16 (0.04)		0.16 (0.04)	0.16 (0.04)
Black		0.09 (0.06)	0.28 (0.07)		0.15 (0.08)	0.30 (0.10)
Hispanic		0.08 (0.07)	0.11 (0.08)		0.01 (0.09)	0.05 (0.10)
Other Race		0.15 (0.06)	0.19 (0.06)		0.09 (0.07)	0.13 (0.07)
% FRPL			0.37 (0.26)			0.10 (0.29)
% White			0.13 (0.15)			0.08 (0.16)
% Black			-0.69 (0.20)			-0.52 (0.25)
			Growth in Summer after 1st Grade			
A	**-0.58 (0.09)**	**-0.58 (0.09)**	**-0.58 (0.09)**	**-0.33 (0.12)**	**-0.33 (0.12)**	**-0.33 (0.12)**
Control Group	-3.17 (0.11)	-3.77 (0.11)	-3.73 (0.12)	-2.73 (0.12)	-3.12 (0.13)	-3.07 (0.13)
State 1	0.40 (0.10)	0.36 (0.10)	0.25 (0.11)	0.50 (0.10)	0.49 (0.10)	0.40 (0.12)
State 2	1.54 (0.14)	1.64 (0.13)	1.65 (0.14)	1.23 (0.15)	1.25 (0.15)	1.19 (0.16)
C2	-0.35 (0.06)	-0.36 (0.06)	-0.36 (0.06)	-0.12 (0.07)	-0.13 (0.07)	-0.13 (0.07)
C3	-0.78 (0.08)	-0.78 (0.08)	-0.77 (0.08)	-0.40 (0.10)	-0.42 (0.11)	-0.42 (0.11)
Female		0.36 (0.04)	0.36 (0.04)		0.19 (0.06)	0.19 (0.06)
Black		0.92 (0.08)	0.98 (0.08)		0.64 (0.10)	0.69 (0.11)
Hispanic		0.84 (0.08)	0.84 (0.09)		0.47 (0.12)	0.42 (0.12)
Other Race		0.40 (0.09)	0.41 (0.09)		0.48 (0.11)	0.48 (0.11)
% FRPL			1.32 (0.40)			-0.05 (0.46)
% White			0.03 (0.20)			-0.45 (0.24)
% Black			-0.33 (0.21)			-0.52 (0.29)

OA6. Summer learning loss estimates for Models 1-6 (continued)

Summer learning loss estimates corresponding to the HLM results presented in Table 4

	Math			Reading		
Variable	(IV)	(V)	(VI)	(IV)	(V)	(VI)
			Growth in Summer after Kindergarten			
A	-0.23 (0.10)	-0.14 (0.09)	-0.16 (0.08)	-0.30 (0.11)	0.01 (0.10)	-0.12 (0.10)
A*State 1	0.05 (0.09)			0.19 (0.10)		
A*State 2	0.33 (0.12)			0.53 (0.13)		
A*Female		-0.02 (0.07)			-0.24 (0.08)	
A*Black			0.03 (0.10)			-0.23 (0.11)
A*Hispanic			0.04 (0.10)			0.34 (0.12)
			Growth in Summer after 1st Grade			
A	-0.67 (0.11)	-0.56 (0.10)	-0.51 (0.10)	-0.29 (0.14)	-0.28 (0.14)	-0.32 (0.13)
A*State 1	0.09 (0.11)			-0.08 (0.13)		
A*State 2	0.27 (0.15)			0.00 (0.18)		
A*Female		-0.04 (0.09)			-0.10 (0.12)	
A*Black			-0.11 (0.11)			0.02 (0.15)
A*Hispanic			-0.22 (0.12)			-0.13 (0.17)