I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17088

Measuring and Correcting Monotonicity Bias: The Case of School Entrance Age Effects

Itay Attar
Danny Cohen-Zada
Todd Elder

JUNE 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17088

Measuring and Correcting Monotonicity Bias: The Case of School Entrance Age Effects

Itay Attar
Ben-Gurion University

Danny Cohen-Zada
Ben-Gurion University and IZA

Todd Elder
Michigan State University

JUNE 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Measuring and Correcting Monotonicity Bias: The Case of School Entrance Age Effects*

Instrumental variables estimators typically must satisfy monotonicity conditions to be interpretable as capturing local average treatment effects. Building on previous research that suggests monotonicity is unlikely to hold in the context of school entrance age effects, we develop an approach for identifying the magnitude of the resulting bias. We also assess the impact on monotonicity bias of bandwidth selection in regression discontinuity (RD) designs, finding that "full sample" instrumental variables estimators may outperform RD in many cases. We argue that our approaches are applicable more broadly to numerous settings in which monotonicity is likely to fail.

JEL Classification: C21, C26, C1, I2, I28

Keywords: monotonicity, selection, entrance age, regression discontinuity, instrumental variable

Corresponding author:
Danny Cohen-Zada
Department of Economics
Ben-Gurion University
Beer-Sheva 84105
Israel
E-mail: danoran@bgu.ac.il

* We are grateful for helpful comments from Daniele Paserman, Scott Imberman, and numerous seminar participants.

I. Introduction

The assumption of monotonicity, which implies that an instrument weakly influences treatment intensity in the same direction for all members of a population, plays a vital role in the interpretability of instrumental variable (IV) and fuzzy regression discontinuity (RD) estimators. Despite its importance, few applied researchers assess whether monotonicity holds in practice; Fiorini and Stevens (2021) found that among 22 papers using IV or RD methods recently published in the *American Economic Review*, half did not mention monotonicity at all.

In contrast to the lack of attention paid to monotonicity among empiricists, several recent methodological studies have developed approaches to detect violations of monotonicity and to analyze conditions under which Local Average Treatment Effects (LATEs) can be identified even in its absence. Dahl et al. (2023) introduce the idea of "local" monotonicity, defined over subsets of values of potential outcomes. They argue that under certain assumptions, a variant of a LATE can be identified even at potential outcome values for which compliers and defiers co-exist. De Chaisemartin (2017) discusses scenarios in which monotonicity is likely to fail, including when instruments are based on the widely used "judge fixed effects" strategy. Like Dahl et al. (2023), De Chaisemartin argues that identification of well-defined estimands is still possible under weaker conditions than monotonicity. Similarly, Słoczyński (2021), Chan et al. (2022), and Frandsen et al. (2023) describe conditions in which IV estimators converge to interpretable estimands under weaker versions of monotonicity than that originally described by Imbens and Angrist (1994).

Our paper extends this existing literature by introducing methods to detect and quantify monotonicity bias. We focus primarily on estimators of the effects of school entrance age on student outcomes such as test scores and grade retention. Barua and Lang (2016) first raised concerns about monotonicity in this setting, especially in the context of the quarter-of-birth instruments used by Angrist and Krueger (1991), and Fiorini and Stevens (2021) argue that monotonicity generally does not hold in the specifications commonly used in the literature. We propose a procedure to eliminate monotonicity bias, which is typically both statistically and economically significant in this setting. We also illustrate the impacts on monotonicity bias of bandwidth choice and of the inclusion of trends in running variables in RD designs.

We emphasize that the failure of monotonicity is not unique to the school entrance age setting, but more generally in RD designs when the running variable is measured coarsely, the treatment is continuous, and the relevant discontinuity in the treatment variable is of the opposite sign of the gradient of the treatment with respect to the running variable. When this "sawtooth" relationship between the treatment and the running variable emerges, monotonicity will typically fail in the absence of perfect compliance. To illustrate this pattern, Figure 1 shows actual and expected school entrance ages as a function of a running variable – the number of days between a child's birth date and the cutoff date for entrance into that year's first grade class – for a birth cohort of Israeli students, described in more detail below. Those who comply with the cutoff law will enter school in the current academic year if they reach age six on or before the cutoff date, but they must wait an additional year to start school if they turn six afterward. For these compliers, entrance age decreases linearly with the running variable on either side of the cutoff, but there is a sharp positive discontinuity at the cutoff point itself. For an example, those born 14 days before the cutoff date begin first grade at the age represented by point *C*, while those born 14 days after the cutoff date must wait until the following year, entering school at the age represented by point *B*.

In practice, compliance with entrance age requirements is imperfect. Numerous studies have shown that children born just before cutoff dates are disproportionately likely to "redshirt" by delaying school entry until the following year.[1] For these children, the corresponding entrance ages are *A* and *B*, respectively. Similarly, another group of non-compliers consists of those who would begin first grade in the academic year that they turn six, regardless of whether they are born before the entrance cutoff or not; for this group, the corresponding entrance ages are *C* and *D*, respectively. The sawtooth pattern in the

[1] Molnár (2024) documents that in Hungary, the share of children who delay entry increases from 6 percent among those born far from the entrance cutoff date to as high as 60 percent for those born in the month immediately before the cutoff. Similarly, Elder and Lubotsky (2009) and Dhuey et al. (2019) find that American children born just before entrance cutoff dates are disproportionately likely to delay entry.

figure is what drives the failure of monotonicity – moving from before the cutoff to afterward increases school entrance age for compliers but reduces it for both groups of non-compliers. This pattern appears in several other contexts, such as those involving class size caps as exogenous variation in class sizes (see, e.g., Figure 1 of Angrist and Lavy, 1999, and Figure 5 of Urquiola and Verhoogen, 2009), mandatory insurance deductibles and healthcare expenditures (Figure 2 of Remmerswaal et al., 2023), and asset size and corporate governance (Figure 3 of Black et al., 2006).

In our primary application, we use administrative records from the Israeli Ministry of Education that allow us to separately identify biases due to both the failure of monotonicity and the potential non-random timing of births. Because Israeli entrance cutoff dates are based on the Jewish calendar, which does not map one-to-one with the Gregorian calendar, each cohort of Israeli students faces a slightly different cutoff date in December. We leverage this variation to control for the timing of births while allowing for variation in birth dates relative to the cutoffs. Importantly, we describe how researchers can identify the magnitude of monotonicity bias even without this unusually rich data environment. In such cases, dichotomizing the treatment variable and rescaling the estimates eliminates monotonicity bias. Finally, we show that a "full-sample" IV approach can often lead to lower bias than RD designs; this counterintuitive finding arises because the full-sample IV approach allows for the inclusion of linear trends in the running variable even when it is measured coarsely, eliminating monotonicity bias.

II. The Failure of Monotonicity in the School Entrance Age Setting

To provide more structure on the environment described above, we consider a case in which a researcher has access to a student's month of birth and wishes to estimate the effect of school entrance age (*SEA*) on an outcome *Y*:

(1) $$Y_i = \alpha_{CT} + \beta_{CT} SEA_i + \varepsilon_i,$$

where β_{CT}, the coefficient of interest, captures the impact of *SEA* on *Y*, the "*CT*" subscript denotes that *SEA* is a continuous treatment, and *i* indexes students. For simplicity, we abstract from other observable student characteristics. As numerous studies have argued (see, e.g., Bedard and Dhuey, 2006; Elder and

Lubotsky, 2009; Dobkin and Ferreira, 2010; Black et al., 2011; Fredriksson and Öckert, 2014; Depew and Eren, 2016; Cook and Kang, 2016; Landersø et al., 2017, 2020; Persson et al., 2021), *SEA* is likely to be correlated with ε, generating bias in OLS estimates of β_{CT}. As a result, researchers have typically used instrumental variables strategies based on the timing of a child's birth relative to school entry cutoff dates.

Defining a binary instrument Z that equals 0 if a child is born in the month before the cutoff and 1 if born in the following month, monotonicity implies that a switch from $Z = 0$ to $Z = 1$ influences entrance age in weakly the same direction for all children. Below we use data from seven birth cohorts who each face different cutoff dates, so for illustrative purposes we assume a hypothetical cutoff date of January 1. Among those who comply with the entrance age requirement, those born in December are, on average, 5 years and 8.5 months old (i.e., roughly 5.71 years old) when they enter first grade on September 1, assuming for simplicity that the distribution of birth dates is uniform within a month.[2] Those born in January are, on average, 6 years and 7.5 months old (6.63 years old) when they begin first grade in the following academic year.

However, as noted above, many children do not begin school as soon as they are legally eligible. Such redshirting behavior reflects parents' efforts to prevent their children from being among the youngest students within a grade, as children near the bottom of the within-grade age distribution have substantially increased likelihoods of later grade repetition and of being diagnosed with socioemotional disorders such as ADHD (see, e.g., Elder, 2010; Evans et al., 2010; Layton et al., 2018). Among this group, the average entrance age is 6.71 if $Z = 0$ and 6.63 if $Z = 1$. Among those in the third group, who begin school in the academic year that they turn six whether they are born before the entrance cutoff or not, the average entrance age is 5.71 if $Z = 0$ and 5.63 if $Z = 1$. This latter phenomenon is less prevalent than delayed entry because public schools typically will not allow students to enter early, but sufficiently motivated families can circumvent entrance cutoff laws (see, e.g., Rinn et al., 2018).

[2] Formal schooling in Israel starts in first grade, not in kindergarten, so school entrance laws refer to entry into first grade.

Panel A of Table 1 illustrates the average entrance ages of the three groups described above, with the compliers shown in the upper-right cell, the redshirters in the lower-right, and those who always enter in the academic year they turn six in the upper-left. The lower-left cell describes a fourth group: those who would enter school a year late if they were born in December (at an average age of 6.71 years) but would instead enter a year early if they were born in January (at an average age of 5.63 years). It is difficult to conceive of optimizing behavior that would be consistent with this pattern, and we view it as sufficiently unlikely that it is ignorable. Nonetheless, even in its absence, monotonicity fails because of the existence of the other two groups of non-compliers.

Although a failure of monotonicity is not necessarily problematic, such as when the returns to *SEA* are homogenous in the population, the bias in this case is potentially large. First, if there is essential heterogeneity (see, e.g., Heckman, Urzua, and Vytlacil, 2006), implying that those who delay entry are those with the highest return to doing so, then redshirters will have disproportionately high returns to *SEA* compared to compliers. This is likely to be the case, as Heckman, Urzua, and Vytlacil (2006) argue that essential heterogeneity is consistent with optimizing behavior in most contexts. Second, as noted above, up to 60 percent of children born in the month just before cutoff dates delay their entry to the following year, implying that the noncompliers are empirically relevant: Z is positively correlated with *SEA* for many children and negatively correlated with *SEA* for many *other* children. As a result, the failure of monotonicity is likely to have large impacts in the setting of school entrance age effects.

III. Eliminating Monotonicity and Selection Biases

We reemphasize that we are not the first to argue that monotonicity can fail in the context of school entrance age effects. Barua and Lang (2016) described the issue in the context of quarter-of-birth instruments, and Fiorini and Stevens (2021) developed tools to detect monotonicity using the *SEA* setting as a motivating example. Our contribution is to extend the logic of testing whether monotonicity holds to estimating the magnitude of the resulting bias and describing a procedure to eliminate it. To do so, we first consider a specification similar to (1), but where we use a binary version of *SEA* rather than *SEA* itself. For concreteness, we define this variable as equaling 1 if SEA is greater than 6.2 years and zero otherwise:

(2) $$Y_i = \alpha_{BT} + \beta_{BT} 1(SEA_i > 6.2) + \varepsilon_i,$$

where we now use the "BT" subscript to denote that this model refers to a binary treatment.[3]

Panel B of Table 1 illustrates why monotonicity holds here, unlike in specification (1). The upper-right cell again corresponds to the compliers: their SEA_i is less than 6.2 when they are born in December and greater than 6.2 when they are born in January. The redshirters in the lower-right cell always enter school after age 6.2, now corresponding to the traditional characterization of "always-takers" because the binary treatment $1(SEA_i > 6.2)$ equals 1 whether $Z = 0$ or $Z = 1$. Similarly, the children who always enter in the year they turn six, shown in the upper-left cell, are "never-takers" in that their treatment always equals 0 whether $Z = 0$ or $Z = 1$. Thus, monotonicity is only violated in this case by the "defiers" in the bottom-left cell. This characterization highlights why dichotomizing the treatment variable eliminates monotonicity bias: it produces always- and never-takers. In contrast, in the continuous treatment model, all children who were not compliers were defiers – no child's value of SEA_i was invariant to the value of Z.

The drawback of using a binary version of SEA_i is that β_{BT} cannot be interpreted as the effect of a year of entrance age because $[E(SEA_i|SEA_i > 6.2) - E(SEA_i|SEA_i \leq 6.2)] < 1$. To scale β_{BT} and β_{CT} identically, we first estimate the relationship between the continuous and the binary SEA measures,

(3) $$SEA_i = \delta_0 + \delta_1 1(SEA_i > 6.2) + \eta_i,$$

and then calculate the adjusted β_{BT} coefficient as $\hat{\beta}_{BT} / \hat{\delta}_1$. This coefficient will have the same scale as β_{CT} but will be free of monotonicity bias.

[3] Figure 1 illustrates why we use a value of 6.2 in specification (2): those who always begin school in the year that they turn 6 have an entrance age below 6.2 (represented by the horizontal red line in the figure) regardless of where their birthdate falls in our estimation samples, and redshirters always enter after age 6.2. Below we use a maximum bandwidth of 150 days, as we exclude children born in June due to ambiguity about the cohort to which they should be assigned.

Selection Bias Due to Nonrandom Distribution of Births

Although estimates based on expressions (2) and (3) are free of monotonicity bias, identification is threatened by the possibility of births being nonrandomly distributed throughout the year. Previous studies have shown that in some settings, children born in December are observably different than children born in January, in part due to tax-based incentives (Dickert-Conlin and Chandra, 1999). More generally, observable characteristics of children and their parents vary systematically throughout the calendar year (see, e.g., Buckles and Hungerman, 2013; Dickert-Conlin and Elder, 2010; Attar and Cohen-Zada, 2018), so children born before the cutoff dates may also be different from those born afterward on unobservable dimensions.

To this point, our discussion has used hypothetical cutoffs and data, but we make use of a unique setting in our empirical work below. We study Israeli children who began school before 2015, and the Israeli school entrance cutoffs during that period allow us to estimate the effect of *SEA* on outcomes while controlling for date of birth. Specifically, the school entry cutoff date was always on the first day of the fourth Jewish month of *Tevet*. Because the Jewish and Gregorian calendars are not identical, this date falls on different Gregorian calendar dates in various years, typically in December; for example, in 2022 the first day of *Tevet* was December 25, and in 2023 it was December 13 (see Attar and Cohen-Zada, 2018, for further discussion of the variation in cutoffs across years). As a result, we can nonparametrically control for unobserved characteristics of students born on different dates of the year by incorporating indicators κ_d and τ_c for each (Gregorian) date of birth, *d*, and birth cohort, *c*, respectively:

(4) $$Y_{idc} = \alpha_0 + \beta_{BT} 1(SEA_{idc} > 6.2) + \kappa_d + \tau_c + \varepsilon_{idc}.$$

We next show that when we include these indicators, we obtain identical estimates regardless of whether we use a binary treatment, as in (4), or a continuous treatment, as in

(5) $$Y_{idc} = \alpha_{CT} + \beta_{CT} SEA_{idc} + \kappa_d + \tau_c + \varepsilon_{idc}.$$

To see this equivalence, note that for a given date of birth, a student may have only two possible entrance ages, where one of them is greater than 6.2 and the other is below 6.2, and the difference between these two entrance ages is exactly one year.[4] Thus, we can write the continuous *SEA* as a linear function of a set of indicators η_d for each date of birth and the binary treatment $1(SEA_{idc} > 6.2)$, where the coefficient on the binary treatment equals one:

$$SEA_{idc} = \eta_d + 1(SEA_{idc} > 6.2). \tag{6}$$

Substituting (6) into (5) yields

$$\begin{aligned} Y_{idc} &= \alpha_{CT} + \beta_{CT}[\eta_d + 1(SEA_{idc} > 6.2)] + \kappa_d + \tau_c + \varepsilon_{idc} = \\ &= \alpha_{CT} + \beta_{CT}1(SEA_{idc} > 6.2) + \varphi_d + \tau_c + \varepsilon_{idc}, \end{aligned} \tag{7}$$

where $\varphi_d = \kappa_d + \beta_{CT}\eta_d$.

Comparing expressions (4) and (7) implies $\beta_{CT} = \beta_{BT}$, so that β_{CT} is free of monotonicity bias.[5] Moreover, both estimates are free of selection biases because the cohort and date-of-birth indicators control for nonrandom selection on birth dates. Specifically, unlike typical RD specifications, which rely on the assumption that within a narrow bandwidth, children born on different dates do not systematically differ on unobserved dimensions, expression (7) instead relies on the weaker assumption that such differences may exist but are constant across birth cohorts. We denote these unbiased estimates as β^*. It is important to note that one can estimate the magnitude of monotonicity bias even without including the cohort and date-of-birth indicators by comparing estimates of β_{CT} from expression (1), which incorporate both selection and monotonicity biases, to those of β_{BT} / δ_1 based on (2) and (3), which incorporate only selection bias. The resulting estimate of monotonicity bias is the sample analog of $\beta_{CT} - (\beta_{BT}/\delta_1)$.

[4] In our empirical results below, consistent with Table 1, we include students who enter school either "on time" or one year later if they are born before the cutoff date, and who enter either "on time" or one year earlier if born after the cutoff date; these restrictions eliminate only 1.6% of the students who enter school at ages below 5.2 and above 7.2 years. We do so because these children, by definition, are not compliers, in that they do not enter school as soon as legally eligible, so they do not contribute to LATEs. Second, in the Israeli case (and we suspect in most others), these children are systematically unrepresentative of the population. See Section IV below for more details.

[5] Panel C of Table 1 further illustrates why monotonicity holds when date-of-birth indicators are included. Consider the lower-right cell, which again corresponds to redshirters who would enter late if they were born before the cutoff date and on time if born afterward. Because both circumstances correspond to the same birth date (we chose December 15 for illustrative purposes in the table), the child enters school at 6.71 years of age regardless of whether the cutoff date is before or after December 15. Similar logic holds for the "always enter before age 6" group represented in the upper left of the table: those children begin school at age 5.71 if they are born on December 15, regardless of whether that date occurred before the entrance cutoff date or afterward.

IV. Data and Empirical Results

IV.1 Data Creation

We use administrative records from the Israeli Ministry of Education, spanning 2002 to 2006 and covering fifth- and eighth-grade students in Jewish localities. Each record includes a child's exact birth date, school entry year, gender, parental education, number of siblings, parental birthplaces, an indicator for whether the child was born in Israel, and indicators for whether the student attended a religious or secular public school.

These administrative records were merged with data from the Growth and Effectiveness Measures for Schools (GEMS), a nationally administered Israeli examination for fifth- and eighth-grade students. This dataset includes math and Hebrew test scores and the years that the exams were taken, allowing us to identify those who repeat grades before fifth and eighth grades. We standardized raw scores within grade-subject-year cells, setting mean scores to zero and standard deviations to one.

We define a cohort based on an academic year's entrance cutoff, including children born on June 15 six years prior to the beginning of the academic year up through the following June 14. This approach positions December at the midpoint of each birth cohort. In our analyses, we exclusively consider cohorts in which we have data for students on both sides of the entrance cutoff; for example, when estimating models of fifth grade test scores (starting in 2002), the first birth cohort we consider is those students born from June 15, 1991, to June 14, 1992. In total, our dataset includes 128,695 observations for fifth graders and 128,818 for eighth graders.[6]

From the remaining observations, we excluded 3,815 students (1.49 percent) with entrance ages above 7.2 years and 274 observations (0.11%) with entrance ages below 5.2 years, resulting in sample sizes of 126,234 for fifth graders and 125,413 for eighth graders. Importantly, many of these students

[6] As in Attar and Cohen-Zada (2018), we found that roughly 2.5 times as many children are reported to have been born on January 1 compared to the average daily number of births on other dates throughout the year. Given that none of our cutoffs fall on this date, we suspect that births coded on January 1 include those in which actual birth dates are unknown; for example, 937 births coded on January 1 correspond to children who are new immigrants from Ethiopia, compared to roughly 4.3 births per day on all other days of the year. We thus removed all January 1 births 1,777 observations from the dataset, leaving us with 127,987 fifth graders and 127,749 eighth graders.

with entry ages above 7.2 or below 5.2 are not typical, in that many of them are immigrants who likely delayed school entry due to language issues; immigrants make up roughly thirteen percent of the students in the 5.2-7.2 age range, but 59.1 percent of the students who fall outside that range.

We have access to data on school entry year for approximately 93.5 percent of students, enabling us to determine if they repeated or skipped grades. The rate of grade retention from first grade to fifth grade is 1.52 percent, and the rate of grade skipping is 0.19 percent. Similarly, from first grade to eighth grade, the grade retention rate is 2.65 percent, and the grade skipping rate is 0.28 percent. Because these rates are relatively low, and to avoid introducing sample selection bias, we retain the 6.5 percent of students for whom we cannot observe the year of school entry. We estimate their entrance age based on their observed age at the beginning of the school year during which they were tested, under the assumption that they neither repeated nor skipped grades. This approximation is expected to introduce error in roughly 0.19 percent of the total observations ($= 0.065 \times (0.0265 + 0.0028) \times 100$); we include an indicator in our estimated models for whether entrance age is observed or approximated.

We primarily use an RD approach below, focusing on an interval of +/-28 days around the cutoff points. This results in a sample of 19,758 fifth graders and 18,770 eighth graders. Of these students, 18,566 fifth graders were tested in math and 18,152 in Hebrew, and 16,912 eighth graders were tested in math and 16,973 in Hebrew. In addition, we observe whether 19,029 fifth graders and 17,023 eighth graders repeated a grade. We present summary statistics from this discontinuity sample in Table 2.[7]

IV.2 Baseline Estimates Based on Monthly Data

In Table 3, we present fuzzy regression discontinuity estimates of the effects of *SEA*, using the +/- 28-day discontinuity sample. We use a binary measure of whether a student is born before the cutoff as the excluded instrument. Column (1) shows the unbiased estimate β^*, obtained from specification (4)

[7] Below we also analyze specifications based on full-year samples, although we exclude children born in June due to ambiguity about which cohort to assign them. Specifically, we use a bandwidth of 150 days around the cutoff point, including 104,249 fifth-grade students and 102,863 eighth-grade students.

(or, equivalently, specification (7)). In columns (2)-(4) we report the estimates $\hat{\beta}_{CT}$, $\hat{\beta}_{BT}$, and $\hat{\beta}_{BT}/\ \hat{\delta}_1$ from expressions (1), (2), and (3) above, respectively.

We analyze fifth-grade Hebrew scores in the top row of the table. The unbiased estimate β^* implies that a one-year increase in *SEA* increases average Hebrew scores by 0.268 standard deviations. The estimate of β_{CT} is 0.329, implying that its estimated bias – which includes both selection and monotonicity biases – is 0.061 (= 0.329 – 0.268). Similarly, $\hat{\beta}_{BT}/\ \hat{\delta}_1$ is 0.288, so the estimated monotonicity bias in $\hat{\beta}_{CT}$ is 0.041 (= 0.329 – 0.288) and the estimated selection bias in $\hat{\beta}_{CT}$ is 0.020. Columns (5)-(7) show the estimates of these combined, selection, and monotonicity biases, along with their standard errors (which we estimate via 500 bootstrap replications clustered on date of birth).[8] The final column lists the magnitude of the monotonicity bias as a proportion of the estimate of β^*; for Hebrew scores, the monotonicity bias is 15.3 percent of the magnitude of the unbiased estimate.

The remaining rows of the table present analogous estimates for 5th grade math scores, Hebrew and math scores in 8th grade, and indicators of grade repetition by 5th and 8th grades. In all cases, monotonicity biases are both statistically and practically significant, ranging from roughly 13 to 16 percent of the unbiased estimates of β^*.[9] Notably, the estimates of monotonicity biases are also larger in absolute value than those of selection biases in five of the six cases, with 8th grade math scores being the lone exception.

All estimates in Table 3 were based on +/- 28-day discontinuity samples. To assess how monotonicity biases vary with bandwidth, Figure 2 shows the estimated absolute values of monotonicity bias for additional bandwidths in multiples of 28 days. For all six outcome variables, the estimated biases increase as the bandwidth expands; they are roughly twice as large using bandwidths of 140 days compared to bandwidths of 28 days. This phenomenon occurs because, as bandwidth increases, the

[8] In all cases here and in the tables below, first-stage *F*-statistics are larger than 270, implying that inference based on *t*-statistics is valid without the adjustments proposed in Lee et al. (2022).

[9] We use students' exact dates of birth to calculate the estimates in Table 3. In Appendix Table A1, we instead use a value of *SEA* based on the assumption that all births in a fictional month of 28 days take place at the midpoint of that month. This captures the likely approach that empiricists would use when they only have access to month of birth. For all six outcomes, the resulting estimates are similar to those shown in Table 3.

magnitude of the violation of monotonicity for a given noncomplier increases. For example, for a 28-day bandwidth, the redshirters born before the cutoff are roughly 28 days older at school entry, on average, than redshirters born after the cutoff (as shown in Panel A of Table 1). Using a 140-day bandwidth, this average age difference is 140 days. Thus, each redshirter contributes more bias due to the failure of monotonicity as the size of the bandwidth grows.[10]

As additional evidence on the potential violation of monotonicity, we also consider the stochastic dominance test of Angrist and Imbens (1995). The intuition of that test in the *SEA* case is that if monotonicity holds, the CDFs of *SEA* for those born before and after cutoff dates cannot cross. Using the +/- 28-day discontinuity sample, we show the CDFs for those born before and after cutoff dates in the top panel of Appendix Figure A1. The figure shows that the CDFs cross twice, with the noticeable intersection at entrance ages between 6.5 and 6.75 reflecting the empirical relevance of delayed entry, which causes some students who are born before the cutoff to have higher entrance ages than those born after the cutoff. As argued above, this type of violation likely appears in many other settings in which the relationship between the average treatment and a running variable has a sawtooth pattern.

To illustrate this phenomenon in an additional setting, in Panel B we show an analogous figure for the case of Maimonides' Rule and class sizes. Maimonides Rule imposes a maximum class size of 40 students, which Angrist and Lavy (1999) use to generate an IV strategy to estimate the effects of class size on student outcomes. We use Angrist and Lavy's (1999) estimation sample for fifth grade test scores (posted at https://economics.mit.edu/people/faculty/josh-angrist/angrist-data-archive), and include schools with enrollments between 36 and 45 students to focus on the discontinuity between 40 and 41. Again, the CDFs intersect, reflecting that some schools with enrollments below 41 open multiple classrooms even though Maimonides' Rule does not require them to do so. Such schools have smaller class sizes when

[10] An opposing effect tends to reduce monotonicity bias as bandwidth grows: the proportion of redshirters in the estimation sample decreases. As noted above, those born in the last month before a cutoff date are disproportionately likely to delay entry, and this phenomenon weakens as the distance between the birth and cutoff dates increase. Thus, there are proportionately fewer redshirters in a five-month estimation window than in a one-month window. The findings in Figure 2 imply that this opposing effect is dominated by the primary effect described above: that the magnitude of the violation of monotonicity for each noncomplier increases with bandwidth.

school enrollments are below 41 than when enrollments are 41 or above, resulting in a failure of monotonicity.

IV.3 Monotonicity Bias and Trends in the Running Variable

To this point, we have assumed that researchers only have access to monthly data to estimate β_{CT} and β_{BT} (although we used information on exact date of birth to estimate β^* in Table 3). We next consider settings in which researchers observe exact birth dates. We define a running variable *RC* as the number of days that a child's birth date falls relative to the cutoff date, which is negative for children born before the cutoff and positive for those born afterward. Previous authors such as Fiorini and Stevens (2021) have argued that the inclusion of a linear trend in *RC* eliminates monotonicity bias if that trend is correctly specified.[11] We can strengthen this argument in the school entrance age context: including linear trends will eliminate monotonicity bias regardless of whether the trend is correctly specified in the model of outcomes. To see why, note that a child's actual *SEA* is a linear function of *RC*:

(8) $$SEA_i = \gamma_0 - \gamma_1 RC_i + 1(SEA_i > 6.2),$$

where γ_0 is the *SEA* (in years) for a child born on the cutoff date who entered school in that year. For example, if the entrance cutoff is December 1, the first day of school is September 1, and the child was born on November 28, then $RC_i = -3$ and $\gamma_0 = 5.75$. Among compliers, SEA_i varies one-for-one with RC_i, so that children born on November 27 have $RC_i = -4$ and are one day older when they begin first grade than children born on November 28. Because SEA_i is measured in years, $\gamma_1 = \frac{1}{365.25}$. Moreover, the slope on RC_i is identical on either side of the cutoff, allowing us to write the SEA_i as an additively separable linear function of RC_i and $1(SEA_i > 6.2)$. If a child instead delays entry, SEA_i increases by one year, as represented by the $1(SEA_i > 6.2)$ term. Note that expression (8) is an identity, rather than an estimating equation like (3) above.

[11] Specifically, Fiorini and Stevens write that to best approximate the trend in the model of outcomes, "[a]n ideal fuzzy RD design would have data relying on date rather than month of birth, use a small bandwidth and include a trend using local linear regression that is allowed to be different on each side of the threshold. That is possibly a robust solution in the school entry age setting…" (p. 1512).

To see how including a linear running variable eliminates monotonicity bias in this setting – and more generally, when the slope of the treatment among compliers with respect to the running variable is identical on either side of a cutoff – recall that there is no monotonicity bias in the binary specification given by (2) above. This is also the case if we include a linear running variable in that specification:

(2') $$Y_i = \alpha_{BT} + \beta_{BT} 1(SEA_i > 6.2) + \varphi RC_i + \varepsilon_i,$$

using the same arguments as in the context of Table 1, Panel B. Rearranging expression (8) and substituting into (2') yields

(9) $$\begin{aligned} Y_i &= \alpha_{BT} + \beta_{BT}[SEA_i - \gamma_0 + \gamma_1 RC_i] + \varphi RC_i + \varepsilon_i \\ &= [\alpha_{BT} - \gamma_0 \beta_{BT}] + \beta_{BT} SEA_i + [\beta_{BT}\gamma_1 + \varphi] RC_i + \varepsilon_i. \end{aligned}$$

Thus, the coefficient on SEA_i in the second line of (9) is identical to the coefficient on $1(SEA_i > 6.2)$ in (2'), and because the latter coefficient is free from monotonicity bias, so is the former.[12] In other words, including a linear trend in a model that uses the continuous treatment variable SEA_i eliminates monotonicity bias, and this equivalency holds regardless of whether the outcome model is correctly specified. It also holds for any choice of bandwidth and regardless of whether one follows typical practice by allowing for different slopes on the running variable on either side of the cutoff date. We again emphasize that this finding is specific to settings, like the school entrance age case, in which the treatment is a linear and additively separable function of the running variable.

We present estimates based on specifications that include a linear running variable in Table 4. As in Table 3, column (1) shows the estimates of β^*. As expected, and unlike in Table 3, here the estimates from models that use continuous (column (2)) and binary (column (3)) measures are equivalent and free of monotonicity bias in all six cases. Column (4) presents estimates of selection bias, given by the

[12] Relatedly, when one includes linear controls for RC_i in a specification using the binary treatment variable, there is no need to rescale the estimate of β_{BT}. To see why, note that after conditioning on RC_i, expression (3) becomes $SEA_i = \delta_0 + \delta_1 1(SEA_i > 6.2) + \delta_2 RC_i + \eta_i$. Expression (8) implies that the coefficient δ_1 in that model is identically one, so that the "adjusted" $\beta_{BT}/\delta_1 = \beta_{BT}$. In Appendix A, we show that even if δ_1 is not identically equal to 1, including a linear function of the running variable will eliminate monotonicity bias as long as the gradient of the treatment variable with respect to the running variable is identical on either side of a discontinuity for compliers. This result again holds regardless of whether the outcome model is correctly specified.

difference between the estimates in columns (2) (or (3)) and (1), and column (5) presents the magnitude of that bias as a fraction of β^*. Note that these estimates are not identical to those in Table 3 because the inclusion of the running variable changes the point estimates of selection bias. Finally, in Appendix Table A2 we present analogous estimates from specifications that allow for different slopes on the running variable on either side of the cutoff dates. In all six cases, the point estimates are nearly identical to those in Table 3, with minor differences stemming from modest variation in the estimates of selection bias.

IV.4 Additional Options when Researchers Have Access to Only Coarse Measures of Running Variables

As we argued above, it is common for researchers to only have access to data on month of birth, rather than exact birth date. In the case of estimating *SEA* effects, a simple solution for eliminating monotonicity bias involves dichotomizing the SEA_i variable, but in other contexts it might not be possible to apply this solution. We next consider two approaches that researchers have typically taken when they only have access to a coarse running variable (such as month of birth). First, one can use RD designs using data from a month or set of months surrounding the cutoff date, where the excluded instrument is an indicator for whether the birth falls after the cutoff. These are the designs described in Section IV.2, and the resulting estimates are generally biased due to both selection and the failure of monotonicity. Second, one can use a full-sample IV approach, which uses data from the entire year and controls for monthly (as opposed to daily) trends in the running variable RC_i. In these models, the first stage is

(10) $$SEA_i = \gamma_0 + \gamma_1 1(RC_i > 0) + \gamma_2 RC_i + \gamma_3[RC_i \times 1(RC_i > 0)] + \varsigma_i,$$

where RC_i is now measured as distance in months from the cutoff date, and $1(RC_i > 0)$ is an indicator equal to 1 if the birth is after the entrance cutoff and 0 otherwise.

We suspect that most researchers would prefer the RD design *a priori*, based on the idea that identification is likely to be "cleanest" in models that use only those born in the neighborhood of a discontinuity point. Although this preference seems reasonable, there is a distinct advantage to using the full-sample IV model: it will generally satisfy monotonicity, unlike the RD design. To see why, recall that specifications using linear trends in the running variable are free of monotonicity bias in the *SEA* case (and in other settings involving identical gradients of the expected treatment with respect to a running variable on both sides of a discontinuity point). This argument did not depend on the units of measurement of the running variable, so even a monthly measure of RC_i will eliminate monotonicity. On the other hand, the full-sample specification may be especially susceptible to selection bias due to nonrandom sorting of births across months and seasons; for example, Buckles and Hungerman (2013) present compelling evidence that observable characteristics of mothers and children vary across birth months in the US, suggesting that unobservable characteristics might follow the same pattern.

To assess the tradeoffs between these two options, in Table 5 we present estimates from full-sample IV specifications that include all births within 150 days of the entrance cutoff date. We measure RC_i as distance in months from the cutoff, treating each 30-day window as a fictional "month". We define SEA_i based on the midpoint of that fictional "month", assigning this value to all children born in the 30-day window. Columns (2) and (3) show that the estimates are identical regardless of whether *SEA* is measured as a binary or multinomial variable, consistent with both specifications being free of monotonicity bias. Column (4) shows the estimated selection bias, and column (5) shows the magnitude of that selection bias as a proportion of the estimate β^*. In three of the six cases, the selection bias is less than 10 percent of β^* and is smaller than the analogous monotonicity bias in the RD specifications shown in Table 3. Taken together, these estimates suggest that full-sample IV performs no worse than the RD specifications.[13]

We present additional evidence about the relative performance of the two options in Figure 3. In each panel, we show the variation across bandwidth (in months) in the absolute value of the estimated monotonicity bias in RD designs using a continuous *SEA* measure with no trend in the running variable, analogous to the estimates in Table 3, along with the selection bias in models that include the monthly running variable. In most cases, the monotonicity bias in the former model is larger than the selection bias

[13] In Appendix Table A3, we show point estimates and standard errors for analogous models that allow the monthly trends in RC_i to differ before and after cutoff dates. The estimates are similar in magnitude to those shown in Table 5, although in the case of "held back prior to 8th grade" the estimated selection bias is slightly larger and marginally statistically significant in Table A3.

in the latter model. The estimated monotonicity bias grows as the bandwidth increases (as shown in Figure 2 above), while there is no clear pattern across outcomes in the effect of bandwidth on the estimated selection bias in models including RC_i. Researchers using models with no monthly trend would likely use the one-month bandwidth, but even those estimates have larger monotonicity biases than the selection biases in the full-sample models using 5-month bandwidths models for 5th grade Hebrew scores, held back prior to 5th grade, and held back prior to 8th grade.

Finally, in Appendix Table A4 we include a summary of the monotonicity and selection biases (in absolute value) for three different specifications: RD models using the +/- 28-day discontinuity sample without trends in the running variable, similar RD models that add trends (either linear or piecewise) in the running variable, and the +/- 5-month "full-sample" specification that includes monthly trends in the running variable. The monotonicity biases in the first model, shown in column (3), are larger than the selection biases in columns (7) and (8) for three of the six outcomes, again implying that specifications including data far from cutoffs may perform as well or better than those using only months straddling the cutoff.

V. Discussion and Conclusions

This paper makes three primary contributions. First, we show that when researchers have access to relatively coarse measures of a running variable, the typical practice of including only those located in the closest neighborhood possible to discontinuity points (in the school entrance age case, this corresponds to those born in the months straddling the cutoff date) can yield empirically relevant monotonicity biases. For each of the six outcomes we study, the estimated monotonicity biases are roughly one-sixth of the magnitude of the corresponding unbiased point estimates. We show that using a dichotomized version of the treatment variable eliminates monotonicity bias even when researchers do not have access to our rich data environment. This dichotomization strategy is a general solution, in that it will deliver estimates free of monotonicity bias in the prototypical case in which the relationship between the expected value of a treatment and a running variable has a sawtooth pattern.

Relatedly, we also consider the relative performance of options available to empiricists when they only have access to coarse running variables. We show that the commonly-used design – again, using only births in months that straddle school entrance cutoffs – does not lead to systematically lower bias than an alternate strategy of including births in all months in a cohort, which eliminates monotonicity biases by allowing for the inclusion of monthly trends in the running variable. We suspect that many applied researchers view selection bias as a more salient threat to identification than violations of monotonicity, but our findings suggest that this view in not always justified.

Finally, we show that including linear functions of running variables eliminates monotonicity bias in the school entrance age context, as well as in other settings in which the average value of the treatment among compliers is a linear function of a running variable with identical slopes on either side of a discontinuity point. Remarkably, this result holds regardless of whether the outcome model is misspecified, such as when the true model is a complex nonlinear function of the underlying running variable. Although this finding is less general than those described above, it highlights the broader point that prior knowledge of how the treatment is determined may be sufficient to eliminate monotonicity bias even if researchers must be agnostic about the functional form of outcome models.

Taken together, our findings extend the recent literature by emphasizing the potential gains to developing estimation strategies that are robust to violations of monotonicity. They also imply that monotonicity should play a role in future methodological work on optimal bandwidth selection in RD designs, as the empirical relevance of monotonicity depends critically on the choice of bandwidth.

References:

Angrist, J.D. and Imbens, G.W., 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, 90(430), pp.431-442.

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

Angrist, J. and Lavy, V., 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly journal of economics*, *114*(2), pp.533-575.

Attar, I. and Cohen-Zada, D., 2018. The effect of school entrance age on educational outcomes: Evidence using multiple cutoff dates and exact date of birth. *Journal of Economic Behavior & Organization*, 153, pp.38-57.

Barua, R. and Lang, K., 2016. School entry, educational attainment, and quarter of birth: A cautionary tale of a local average treatment effect. *Journal of Human Capital*, *10*(3), pp.347-376.

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

Black, B.S., Jang, H. and Kim, W., 2006. Does corporate governance predict firms' market values? Evidence from Korea. *Journal of Law, Economics, and Organization*, 22(2), pp.366-413.

Black, S.E., Devereux, P.J. and Salvanes, K.G., 2011. Too young to leave the nest? The effects of school starting age. *The review of economics and statistics*, 93(2), pp.455-467.

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

Chan, D.C., Matthew Gentzkow, and Chuan Yu. 2022. Selection with Variation in Diagnostic Skill: Evidence from Radiologists. *Quarterly Journal of Economics* 137 (2): 729–83.

Cook, P.J. and 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), pp.33-57.

Dahl, C.M., Huber, M. and Mellace, G., 2023. It's never too LATE: A new look at local average treatment effects with or without defiers. *The Econometrics Journal*.

De Chaisemartin, C., 2017. Tolerating defiance? Local average treatment effects without monotonicity. *Quantitative Economics*, 8(2), pp.367-396.

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

Dhuey, E., Figlio, D., Karbownik, K., and Roth, J., 2019. School Starting Age and Cognitive Development. *Journal of Policy Analysis and Management*, 38(3):538–578.

Dickert-Conlin, S. and Chandra, A., 1999. Taxes and the Timing of Births. *Journal of political Economy*, *107*(1), pp.161-177.

Dickert-Conlin, S. and Elder, T., 2010. Suburban legend: School cutoff dates and the timing of births. *Economics of Education Review*, *29*(5), pp.826-841.

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

Elder, T.E., 2010. The importance of relative standards in ADHD diagnoses: evidence based on exact birth dates. *Journal of health economics*, *29*(5), pp.641-656.

Elder, T.E. and Lubotsky, D.H., 2009. Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers. *Journal of human Resources*, *44*(3), pp.641-683.

Evans, W.N., Morrill, M.S. and Parente, S.T., 2010. Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children. *Journal of health economics*, *29*(5), pp.657-673.

Fiorini, M. and Stevens, K., 2021. Scrutinizing the Monotonicity Assumption in IV and fuzzy RD designs. *Oxford Bulletin of Economics and Statistics*, *83*(6), pp.1475-1526.

Frandsen, B., Lefgren, L. and Leslie, E., 2023. Judging judge fixed effects. *American Economic Review*, *113*(1), pp.253-277.

Fredriksson, P. and Öckert, B., 2014. Life-cycle effects of age at school start. *The Economic Journal*, 124(579), pp.977-1004.

Heckman, J.J., Urzua, S. and Vytlacil, E., 2006. Understanding instrumental variables in models with essential heterogeneity. *The review of economics and statistics*, *88*(3), pp.389-432.

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

Layton, T.J., Barnett, M.L., Hicks, T.R. and Jena, A.B., 2018. Attention deficit–hyperactivity disorder and month of school enrollment. *New England Journal of Medicine*, *379*(22), pp.2122-2130.

Landersø, R., Nielsen, H.S. and Simonsen, M., 2017. School starting age and the crime-age profile. *The Economic Journal*, 127(602), pp.1096-1118.

Landersø, R.K., Nielsen, H.S. and Simonsen, M., 2020. Effects of school starting age on the family. *Journal of Human Resources*, 55(4), pp.1258-1286.

Lee, D.S., McCrary, J., Moreira, M.J. and Porter, J., 2022. Valid t-ratio Inference for IV. *American Economic Review*, 112(10), pp.3260-3290.

Molnár, Tímea Laura, 2024. Can Academic Redshirting Shrink the Education Gender Gap? Causal Evidence on Student Achievement and Mental Health. Central European University.

Persson, P., Qiu, X. and Rossin-Slater, M., 2021. *Family spillover effects of marginal diagnoses: The case of ADHD* (No. w28334). National Bureau of Economic Research.

Remmerswaal, M., Boone, J., Bijlsma, M., and Douven, R.C., 2019. Cost sharing designs matters: A comparison of the rebate and deductible in healthcare. *Journal of Public Economics,* 170, 83-97.

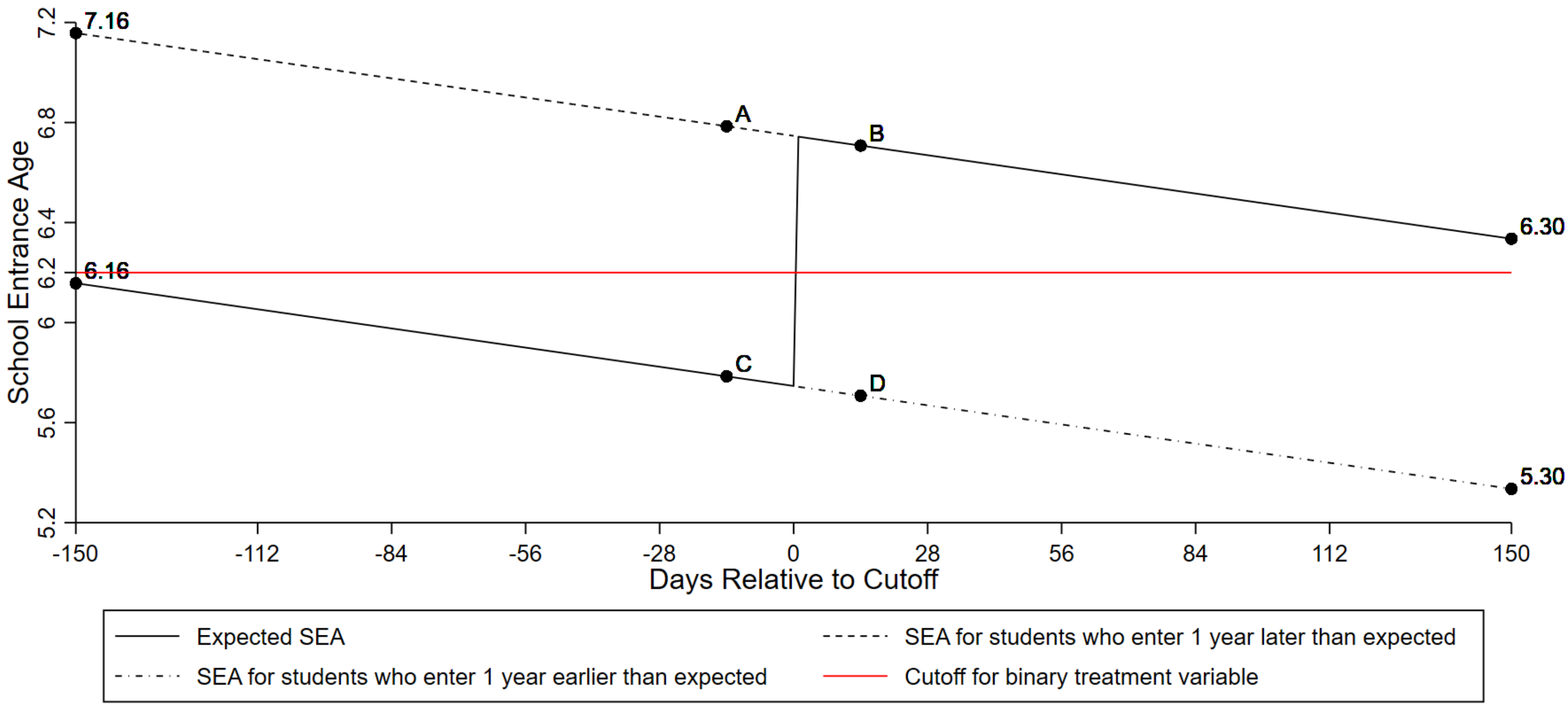
Remmerswaal, M., Boone, J. and Douven, R., 2023. Minimum generosity levels in a competitive health insurance market. *Journal of Health Economics*, 90, p.102782.

Rinn, A.N., Mun, R.U. and Hodges, J., 2018. 2019 State of the States in Gifted Education. *National Association for Gifted Children and the Council of State Directors of Programs for the Gifted.*

Słoczyński, T., 2021. *When Should We (Not) Interpret Linear IV Estimands as LATE?* (No. 9064). CESifo.

Urquiola, M. and Verhoogen, E., 2009. Class-size caps, sorting, and the regression-discontinuity design. *American Economic Review*, 99(1), pp.179-215.

Figure 1: Expected and Actual School Entrance Ages for a Cohort of Israeli Students



Notes: The solid black line in the figure denotes actual school entrance ages among one cohort of Israeli children who fully comply with the school entrance cutoff (i.e., those children who enter first grade as soon as legally eligible to do so), with the *X*-axis measuring their birth date relative to the cutoff. This representative cohort, born between July 1, 1994, and May 31, 1995, faced a cutoff date of December 3. A complier whose birthday is November 19, 14 days before that cutoff date, entered first grade on September 1, 2000, at an age of 5.79 years, as represented by point "C" in the figure.

Figure 2: Monotonicity Biases Based on Monthly Data with No Included Trends in the Running Variable

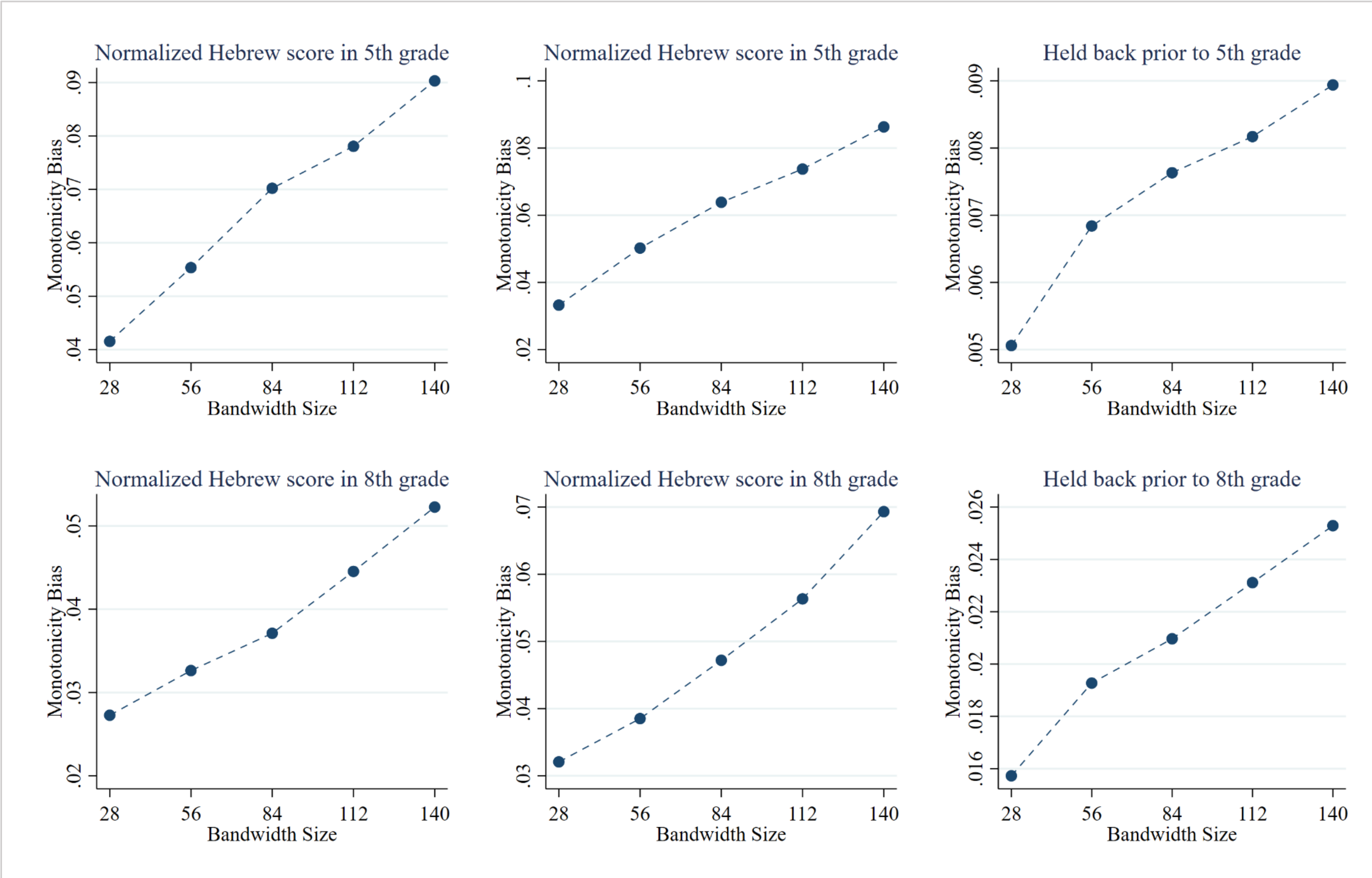


Figure 3: Selection and Monotonicity Biases as a Proportion of $\boldsymbol{\beta}^*$ across Various Specifications

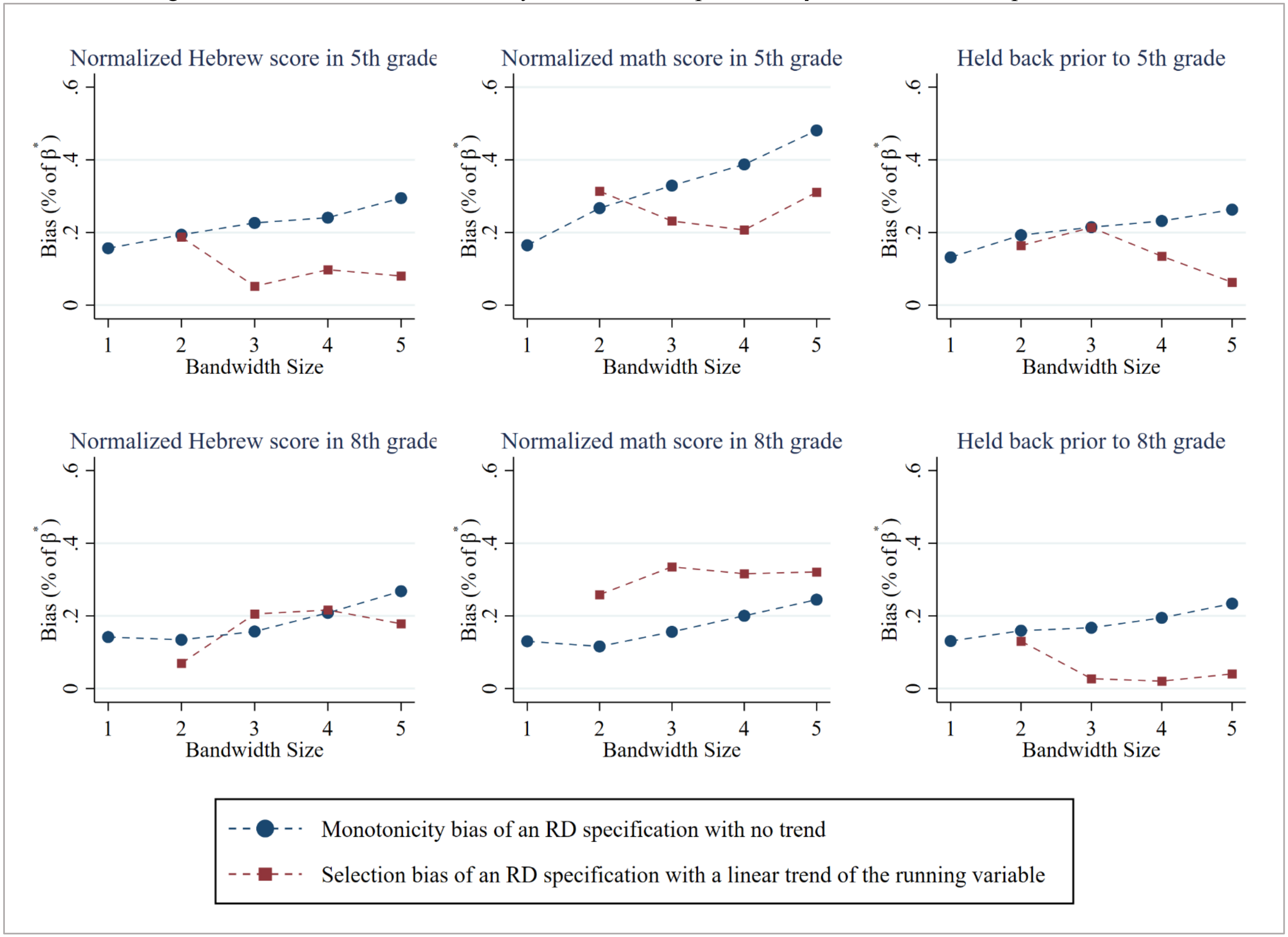


Table 1: Actual and Counterfactual Entrance Ages in Different Specifications

Panel A: Actual and Counterfactual Entrance Ages for a Child Born in December Compared to January (RD Design)

		Born After Cutoff	
		Enter Early	**Enter On time**
Born Before Cutoff	**Enter On time**	(5.71, 5.63)	(5.71, 6.63)
	Enter Late	Unlikely	(6.71, 6.63)

Panel B: Actual and Counterfactual Entrance Ages for a Child Born in December Compared to January (RD Design) Using a Binary Treatment Variable $1(SEA > 6.2)$

		Born After Cutoff	
		Enter Early	**Enter On Time**
Born Before Cutoff	**Enter On time**	(0,0)	(0,1)
	Enter Late	Unlikely	(1,1)

Panel C: Actual and Counterfactual Entrance Age for a Child Born on December 15 (including date-of-birth indicators)

		Born After Cutoff	
		Enter Early	**Enter On Time**
Born Before Cutoff	**Enter On time**	(5.71, 5.71)	(5.71, 6.71)
	Enter Late	Unlikely	(6.71, 6.71)

Table 2: Summary Statistics for RDD Sample (±28 Days)

	5th Grade (n=19,758)		8th Grade (n=18,770)	
Outcome Variables	**Mean**	**SD**	**Mean**	**SD**
Retained in school prior to GEMS exam	0.017	0.129	0.037	0.190
Normalized Hebrew score	0.041	0.983	0.031	0.989
Normalized Math score	0.029	0.988	0.021	0.996
Age Variables				
After cutoff	0.490	0.500	0.483	0.500
Entrance age	6.406	0.444	6.397	0.444
Background Variables				
Father education (0-8 Years)	0.037	0.190	0.043	0.203
Father education (9-11 Years)	0.129	0.335	0.145	0.352
Father education (12 Years)	0.380	0.485	0.356	0.479
Father education (13-16 Years)	0.235	0.424	0.249	0.433
Father education (17+ Years)	0.110	0.312	0.107	0.310
Mother education (0-8 Years)	0.029	0.168	0.033	0.179
Mother education (9-11 Years)	0.102	0.302	0.116	0.320
Mother education (12 Years)	0.400	0.490	0.388	0.487
Mother education (13-16 Years)	0.274	0.446	0.282	0.450
Mother education (17+ Years)	0.097	0.296	0.104	0.305
Number of siblings (0-1)	0.328	0.469	0.309	0.462
Number of siblings (2)	0.340	0.474	0.281	0.449
Number of siblings (3+)	0.208	0.406	0.190	0.392
Male	0.496	0.500	0.493	0.500
Attended secular public school	0.755	0.430	0.795	0.404
Father born in Asia or Africa	0.228	0.420	0.283	0.451
Father born in Americas	0.078	0.268	0.091	0.288
Father born in Israel	0.575	0.494	0.585	0.493
Mother born in Asia or Africa	0.214	0.410	0.258	0.437
Mother born in Americas	0.089	0.284	0.093	0.291
Mother born in Israel	0.603	0.489	0.639	0.480
Student born in Israel	0.895	0.306	0.843	0.363

The number of observations for the outcome variables are lower than the number reported in the table. Specifically, 18,566 fifth graders were tested in math and 18,152 in Hebrew, 16,912 eighth graders were tested in math and 16,973 in Hebrew, and we observe whether 19,029 fifth graders and 17,023 eighth were retained in a grade.

Table 3: Assessment of the Monotonicity and Selection Biases for a Discontinuity Sample (±28 days) with No Trend in the Running Variable (and *SEA* Based on Exact Birth Dates)

	$\boldsymbol{\beta^*}$	$\boldsymbol{\beta^{CT}}$	$\boldsymbol{\beta^{BT}}$	**Adjusted** $\boldsymbol{\beta^{BT}}$ (β^{BT}/δ_1)	**Combined Bias (2) - (1)**	**Selection Bias (4) - (1)**	**Monotonicity Bias (2) - (4)**	$\boldsymbol{\frac{MB}{\beta^*}}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hebrew score in 5th grade								
School Entrance Age	0.268***	0.329***	0.275***	0.288***	0.061	0.020	0.041***	15.3%
N=18,152	(0.099)	(0.039)	(0.033)	(0.035)	(0.089)	(0.089)	(0.005)	
Math score in 5th grade								
School Entrance Age	0.205**	0.263***	0.219***	0.229***	0.058	0.025	0.033***	16.1%
N=18,566	(0.082)	(0.038)	(0.031)	(0.031)	(0.074)	(0.074)	(0.005)	
Held back prior to 5th grade								
School Entrance Age	-0.039***	-0.041***	-0.034***	-0.036***	-0.002	0.003	-0.005***	12.8%
N=19,029	(0.012)	(0.005)	(0.004)	(0.004)	(0.012)	(0.012)	(0.001)	
Hebrew score in 8th grade								
School Entrance Age	0.183	0.206***	0.171***	0.179***	0.023	-0.004	0.027***	14.8%
N=16,973	(0.117)	(0.033)	(0.027)	(0.028)	(0.109)	(0.109)	(0.005)	
Math score in 8th grade								
School Entrance Age	0.250*	0.239***	0.197***	0.207***	-0.010	-0.043	0.032***	12.8%
N=16,912	(0.133)	(0.041)	(0.034)	(0.035)	(0.123)	(0.123)	(0.006)	
Held back prior to 8th grade								
School Entrance Age	-0.123***	-0.127***	-0.106***	-0.111***	-0.004	0.012	-0.016***	13.0%
N=17,023	(0.022)	(0.006)	(0.005)	(0.006)	(0.021)	(0.021)	(0.001)	
Date-of-year fixed effects	***Yes***	***No***	***No***	***No***				
Controls	***Yes***	***Yes***	***Yes***	***Yes***				

Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Assessment of Selection Biases for a Discontinuity Sample (±28 days) with Linear Daily Trends in the Running Variable

	β^{*}	β^{CT}	β^{BT}	**Selection Bias (3) - (1)**	$\frac{SB}{\beta^{*}}$
	(1)	(2)	(3)	(4)	(5)
Hebrew score in 5th grade					
School Entrance Age	0.268***	0.297***	0.297***	0.029	10.9%
N=18,152	(0.099)	(0.091)	(0.091)	(0.042)	
Math score in 5th grade					
School Entrance Age	0.205**	0.254***	0.254***	0.049	23.9%
N=18,566	(0.082)	(0.077)	(0.077)	(0.049)	
Held back prior to 5th grade					
School Entrance Age	-0.039***	-0.044***	-0.044***	-0.005	12.6%
N=19,029	(0.012)	(0.011)	(0.011)	(0.006)	
Hebrew score in 8th grade					
School Entrance Age	0.183	0.068	0.068	-0.115	-63.0%
N=16,973	(0.117)	(0.096)	(0.096)	(0.060)	
Math score in 8th grade					
School Entrance Age	0.250*	0.193*	0.193*	-0.057	-22.7%
N=16,912	(0.133)	(0.101)	(0.101)	(0.071)	
Held back prior to 8th grade					
School Entrance Age	-0.123***	-0.127***	-0.127***	-0.004	3.0%
N=17,023	(0.022)	(0.019)	(0.019)	(0.010)	
Date-of-year fixed effects	***Yes***	***No***	***No***		
Linear trend of RC	***No***	***Yes***	***Yes***		
Controls	***Yes***	***Yes***	***Yes***		

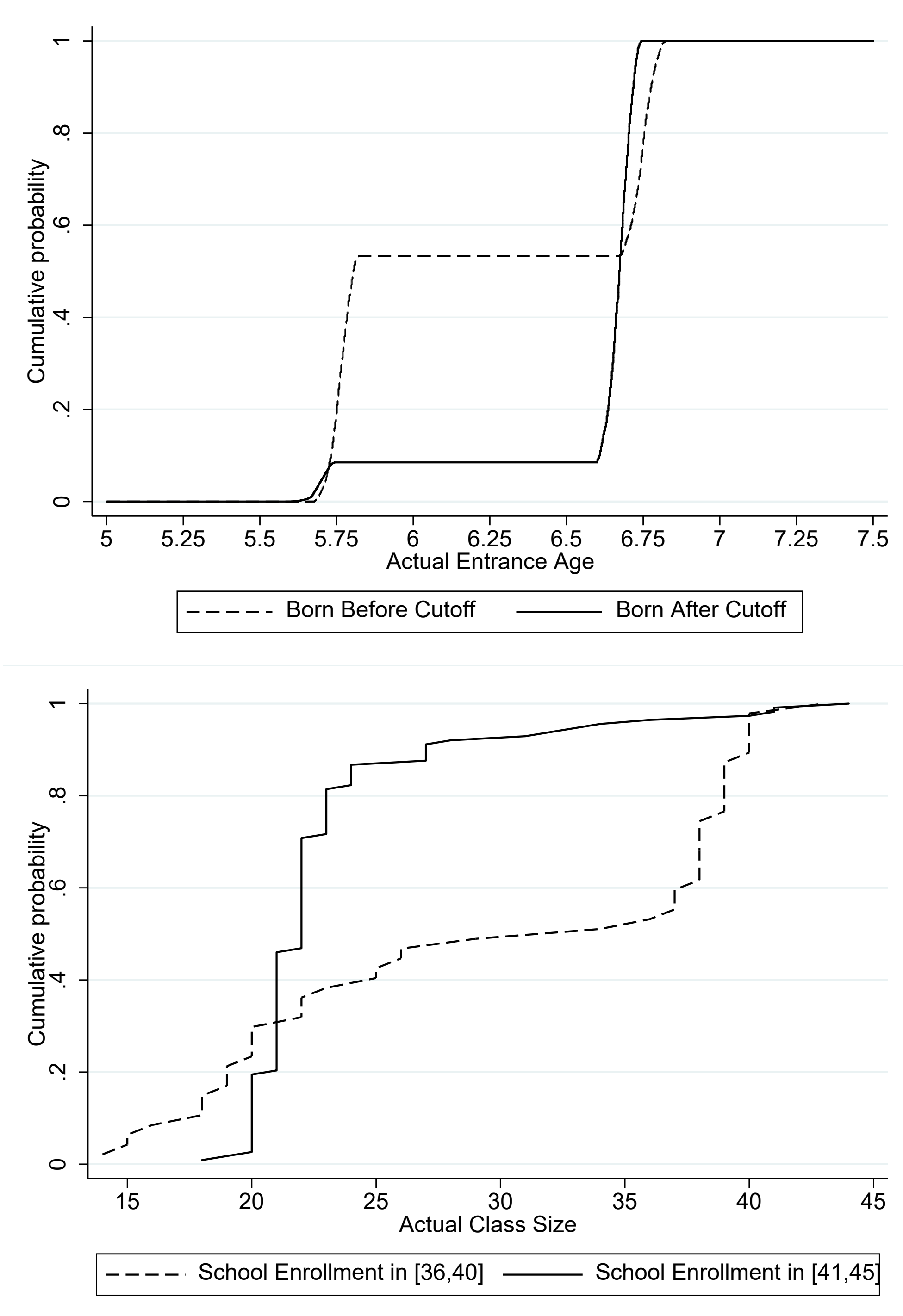
Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Assessment of Selection Biases for a Bandwidth of 150 days with a Linear Monthly Trend in The Running Variable (Treating Each 30 days as a Fictional Month and Assigning All Children in That Month the Same *SEA*)

	$\boldsymbol{\beta^*}$	$\boldsymbol{\beta^{CT}}$	$\boldsymbol{\beta^{BT}}$	**Selection Bias (3) - (1)**	$\boldsymbol{\frac{SB}{\beta^*}}$
	(1)	(2)	(3)	(4)	(5)
Hebrew score in 5th grade					
School Entrance Age	0.326***	0.300***	0.300***	-0.026	-8.1%
N=96,001	(0.084)	(0.027)	(0.027)	(0.078)	
Math score in 5th grade					
School Entrance Age	0.194***	0.255***	0.255***	0.060	31.0%
N=97,888	(0.057)	(0.028)	(0.028)	(0.054)	
Held back prior to 5th grade					
School Entrance Age	-0.036***	-0.039***	-0.039***	-0.002	6.3%
N=100,609	(0.009)	(0.004)	(0.004)	(0.010)	
Hebrew score in 8th grade					
School Entrance Age	0.208**	0.171***	0.171***	-0.037	-18.0%
N=93,213	(0.084)	(0.024)	(0.024)	(0.080)	
Math score in 8th grade					
School Entrance Age	0.297***	0.202***	0.202***	-0.095	-32.3%
N=92,738	(0.107)	(0.028)	(0.028)	(0.098)	
Held back prior to 8th grade					
School Entrance Age	-0.115***	-0.110***	-0.110***	0.005	-4.2%
N=93,000	(0.014)	(0.008)	(0.008)	(0.014)	
Date-of-year fixed effects	***Yes***	***No***	***No***		
Linear trend of RC	***No***	***Yes***	***Yes***		
Controls	***Yes***	***Yes***	***Yes***		

Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Figure A1: Stochastic Dominance Tests of Angrist and Imbens (1995) for the School Entrance Age and Maimonides Rule Cases



Notes: The top panel displays Angrist and Imbens's (1995) stochastic dominance test using the +/- 28-day discontinuity sample described in the text. The bottom panel displays the test using data from Angrist and Lavy (1999), focusing on schools with enrollment in the [36,45] window.

Appendix A

Here we prove that including linear slopes in the running variable eliminates monotonicity bias when the gradient of the treatment with respect to the running variable is identical on either side of a discontinuity. First, we assume that we can write the outcome of interest as a function of the dichotomized treatment $BT_i = 1(CT_i > \lambda)$, where λ is a constant:

(A1) $Y_i = \alpha_{BT} + \beta_{BT} BT_i + \phi RC_i + u_i$.

We further assume that we can write the continuous treatment as

(A2) $CT_i = \gamma_0 + \gamma_1 RC_i + \gamma_2 1(CT_i > \lambda)$,

where λ is such that for all compliers, $CT_i > \lambda$ for those to the right of the cutoff and $CT_i < \lambda$ for those to the left of the cutoff, and such a value of λ exists for at least some value of the bandwidth. Note that there is no error term in equation (A2), as it is deterministic. Rearranging (A2) using the fact that $BT_i = 1(CT_i > \lambda)$ and substituting into (A1),

(A3) $Y_i = \alpha_{BT} + \beta_{BT} \left[\frac{CT_i - \gamma_0 - \gamma_1 RC_i}{\gamma_2}\right] + \phi RC_i + u_i$

$$= [\alpha_{BT} - \frac{\beta_{BT}\gamma_0}{\gamma_2}] + \beta_{BT}\frac{CT_i}{\gamma_2} + [\phi - \frac{\gamma_1 \beta_{BT}}{\gamma_2}]RC_i + u_i .$$

Thus, in this model of Y_i as a function of CT_i and RC_i, the population coefficient on CT_i is $\frac{\beta_{BT}}{\gamma_2}$.

Recalling that we need to rescale the parameter β_{BT} to make it comparable to a continuous model, a generalization of expression (3) in the main text is

(A4) $CT_i = \delta_0 + \delta_1 BT_i + \delta_2 RC_i + \eta_i$.

Comparing (A4) to (A2) implies $\delta_1 = \gamma_2$, so that the rescaled parameter of the binary treatment $\equiv \frac{\beta_{BT}}{\delta_1} = \frac{\beta_{BT}}{\gamma_2}$; thus, the population slope on CT_i in a continuous treatment model is identical to the rescaled parameter from the binary treatment model. Note that this is a generalization of the SEA_i case in the main text in that it relaxes the assumption that $\gamma_2 = 1$.

Table A1: Assessment of the Monotonicity and Selection Biases for a Discontinuity Sample (±28 days) with No Trend in the Running Variable (All Children Born in the Same Side of the Cutoff Are Assigned the Same *SEA* Based on the Middle of that Fictional Month)

	β^*	β^{CT}	β^{BT}	**Adjusted** β^{BT} (β^{BT}/δ_1)	**Combined Bias (2) - (1)**	**Selection Bias (4) - (1)**	**Monotonicity Bias (2) - (4)**	$\frac{MB}{\beta^*}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hebrew score in 5th grade								
School Entrance Age	0.268^{***}	0.335^{***}	0.275^{***}	0.288^{***}	0.066	0.019	0.047^{***}	17.5%
N=18,152	(0.099)	(0.040)	(0.033)	(0.032)	(0.092)	(0.093)	(0.006)	
Math score in 5th grade								
School Entrance Age	0.205^{**}	0.267^{***}	0.219^{***}	0.229^{***}	0.062	0.024	0.038^{***}	18.3%
N=18,566	(0.082)	(0.038)	(0.031)	(0.033)	(0.079)	(0.078)	(0.006)	
Held back prior to 5th grade								
School Entrance Age	-0.039^{***}	-0.042^{***}	-0.034^{***}	-0.036^{***}	-0.003	0.003	-0.006^{***}	14.6%
N=19,029	(0.013)	(0.005)	(0.004)	(0.004)	(0.012)	(0.012)	(0.001)	
Hebrew score in 8th grade								
School Entrance Age	0.183	0.209^{***}	0.171^{***}	0.179^{***}	0.026	-0.005	0.031^{***}	16.8%
N=16,973	(0.117)	(0.033)	(0.027)	(0.029)	(0.109)	(0.110)	(0.005)	
Math score in 8th grade								
School Entrance Age	0.250^{*}	0.243^{***}	0.197^{***}	0.207^{***}	-0.007	-0.043	0.036^{***}	14.6%
N=16,912	(0.133)	(0.042)	(0.034)	(0.037)	(0.131)	(0.132)	(0.007)	
Held back prior to 8th grade								
School Entrance Age	-0.123^{***}	-0.129^{***}	-0.106^{***}	-0.111^{***}	-0.006	0.012	-0.018^{***}	14.5%
N=17,023	(0.022)	(0.006)	(0.005)	(0.005)	(0.021)	(0.022)	(0.001)	
Date-of-year fixed effects	***Yes***	***No***	***No***	***No***				
Controls	***Yes***	***Yes***	***Yes***	***Yes***				

Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Table A2: Assessment of the Selection Bias for a Discontinuity Sample (±28 days) with Piecewise Linear Trend in the Running Variable

	$\boldsymbol{\beta^*}$	$\boldsymbol{\beta^{CT}}$	$\boldsymbol{\beta^{BT}}$	**Selection Bias (3) - (1)**	$\boldsymbol{\frac{SB}{\beta^*}}$
	(1)	(2)	(3)	(4)	(5)
Hebrew score in 5th grade					
School Entrance Age	0.268***	0.294***	0.294***	0.026	9.8%
N=18,152	(0.099)	(0.091)	(0.091)	(0.043)	
Math score in 5th grade					
School Entrance Age	0.205**	0.254***	0.254***	0.049	24.0%
N=18,566	(0.082)	(0.077)	(0.077)	(0.050)	
Held back prior to 5th grade					
School Entrance Age	-0.039***	-0.044***	-0.044***	-0.005	13.1%
N=19,029	(0.012)	(0.011)	(0.011)	(0.007)	
Hebrew score in 8th grade					
School Entrance Age	0.183	0.069	0.069	-0.114	-62.4%
N=16,973	(0.117)	(0.095)	(0.095)	(0.059)	
Math score in 8th grade					
School Entrance Age	0.250*	0.185*	0.185*	-0.065	-26.1%
N=16,912	(0.133)	(0.096)	(0.096)	(0.081)	
Held back prior to 8th grade					
School Entrance Age	-0.123***	-0.128***	-0.128***	-0.005	3.8%
N=17,023	(0.022)	(0.019)	(0.019)	(0.011)	
Date-of-year fixed effects	***Yes***	***No***	***No***		
Linear trend of RC	***No***	***Yes***	***Yes***		
Controls	***Yes***	***Yes***	***Yes***		

Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Table A3: Assessment of the Selection Bias for a Bandwidth of 150 days with a Piecewise Linear Monthly Trend of the Running Variable (Treating Each 30-day Period as a Fictional Month, while *SEA* is Based on the Exact Date of Birth)

	$\boldsymbol{\beta^*}$	$\boldsymbol{\beta^{CT}}$	$\boldsymbol{\beta^{BT}}$	**Selection Bias (3) - (1)**	$\boldsymbol{\frac{SB}{\beta^*}}$
	(1)	(2)	(3)	(4)	(5)
Hebrew score in 5th grade					
School Entrance Age	0.326***	0.297***	0.297***	-0.029	-8.9%
N=96,001	(0.084)	(0.025)	(0.025)	(0.077)	
Math score in 5th grade					
School Entrance Age	0.194***	0.247***	0.247***	0.053	27.2%
N=97,888	(0.057)	(0.025)	(0.025)	(0.054)	
Held back prior to 5th grade					
School Entrance Age	-0.036***	-0.034***	-0.034***	0.002	-6.4%
N=100,609	(0.009)	(0.003)	(0.003)	(0.010)	
Hebrew score in 8th grade					
School Entrance Age	0.208**	0.166***	0.166***	-0.043	-20.6%
N=93,213	(0.084)	(0.022)	(0.022)	(0.077)	
Math score in 8th grade					
School Entrance Age	0.297***	0.190***	0.190***	-0.107	-36.1%
N=92,738	(0.107)	(0.026)	(0.026)	(0.098)	
Held back prior to 8th grade					
School Entrance Age	-0.115***	-0.092***	-0.092***	0.022*	-19.5%
N=93,000	(0.014)	(0.004)	(0.004)	(0.013)	
Date-of-year fixed effects	***Yes***	***No***	***No***		
Linear trend of RC	***No***	***Yes***	***Yes***		
Controls	***Yes***	***Yes***	***Yes***		

Notes: Standard errors based on 500 bootstrap replications and clustered at the date of birth level are shown in parentheses. In all models, we control for the children's background characteristics described in the text. "*", "**", and "***" denote significance at the 10%, 5%, and 1% levels, respectively.

Table A4: Summary of Monotonicity and Selection Biases for Various Specifications

	RDD Sample (±28 Days)						Full year (±5 months)	
	No Trend				Linear Trend	Piecewise Linear Trend	Linear Monthly Trend	Piecewise Linear Trend
Bias as proportion of β^*	**Combined Bias**	**Sum of Biases (Abs)**	**Monotonicity Bias**	**Selection Bias**	**Selection Bias**	**Selection Bias**	**Selection Bias**	**Selection Bias**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hebrew score in 5th grade	22.8%	22.8%	15.3%	7.5%	10.9%	9.8%	8.1%	8.9%
Math score in 5th grade	28.3%	28.3%	16.1%	12.2%	23.9%	24.0%	31.0%	27.2%
Held back prior to 5th grade	5.1%	20.5%	12.8%	7.7%	12.6%	13.1%	6.3%	6.4%
Hebrew score in 8th grade	12.6%	16.9%	14.8%	2.2%	63.0%	62.4%	18.0%	20.6%
Math score in 8th grade	4.0%	30.0%	12.8%	17.2%	22.7%	26.1%	32.3%	36.1%
Held back prior to 8th grade	3.3%	22.8%	13.0%	9.8%	3.0%	3.8%	4.2%	19.5%