DISCUSSION PAPER SERIES

IZA DP No. 14134

Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design

Krzysztof Karbownik
Umut Özek

FEBRUARY 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14134

Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design

Krzysztof Karbownik
Emory University and IZA

Umut Özek
American Institutes for Research

FEBRUARY 2021

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

Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design*

Using a regression discontinuity design generated by school-entry cutoffs and school records from an anonymous district in Florida, we identify externalities in human capital production function arising from sibling spillovers. We find positive spillover effects from an older to a younger child in less affluent families and negative spillover effects from a younger to an older child in more affluent families. These results provide empirical evidence that educational policies could create both positive and negative within-family externalities depending on the characteristics of the affected households.

JEL Classification: D13, I20, J13

Keywords: school starting age, sibling spillovers, human capital externalities

Corresponding author:
Krzysztof Karbownik
Department of Economics
Emory University
Rich Memorial Building
1602 Fishburne Drive
30307 Atlanta, Georgia
USA

E-mail: krzysztof.karbownik@emory.edu

* We appreciate feedback from David Figlio, Krzysztof Kalisiak and Paco Martorell as well as conference and seminar participants at APPAM, Aarhus University, Michigan State University, National Bureau of Economic Research Education Meeting, Northwestern University, University of Copenhagen and University of Hong Kong. We are grateful to the anonymous county in Florida for providing the data used in the analysis. Views expressed here are those of the authors and do not necessarily reflect those of the anonymous district or the institutions to which the authors are affiliated.

1 Introduction

Social scientists and policymakers have long been interested in understanding the human capital production function and the role it plays in equality of opportunity. One question that gained a lot of interest in recent decades relates to potential externalities that changes in human capital generate. The two streams of research in this area examine spillovers due to societal interactions (Manski 2000) and within-family spillovers (Black and Devereux 2011). Understanding these externalities is important because if the hypothesized associations reflect causality, then positive (negative) shocks to an individual can propagate beyond the direct effects that are typically considered in program evaluations, and can lead to greater declines (increases) in inequality. In this paper, we study a particular type of within-family externality that has received relatively little attention in the extant literature due to data limitations and identification challenges – sibling spillovers.

Theoretically, we consider two main mechanisms through which the educational achievement of one sibling could influence the other regardless of birth order. First, sibling spillover effects could arise due to direct interactions between siblings. This channel could be more prevalent in less affluent, single-parent households in which the older child serves as the provider/role model for the younger sibling. Second, the poor academic performance of one sibling could lead to parents diverting resources, including time or money, away from other children (compensatory behavior, e.g., Pitt et al. (1990) or Conley (2008)) or from the struggling child (reinforcing behavior, e.g., Becker and Tomes (1976) or Grätz and Torche (2016)); or this poor academic performance could lead parents to engage in even more complex reallocation mechanisms (Yi et al. 2015; Leight and Liu 2020). Thus, theoretically, both spillovers from an older to a younger child and spillovers from a younger to an older child are possible; and we are able to causally estimate both using quasi experimental variation and a regression discontinuity (RD) design.

Empirically, establishing causality in social interactions in general is difficult due to well-known problems with simultaneity, correlated unobservables, and reflection (Manski 1993; Manski 2000); and even experimental studies often yield highly context-specific results (Sacerdote 2014). All these problems are likely present when studying the effect of one sibling on another since these children typically grow up in the same household, share common traits (e.g., cultural values) as well as experiences, and are genetically related.[1]

We estimate sibling spillovers in the context of a universally used education policy – school starting age. Previous work has documented large effects of the policy on children's development, with children who are relatively older than their classmates having improved cognitive outcomes (Bedard and Dhuey 2006; McEwan and Shapiro 2008; Elder and Lubotsky 2009; Dhuey et al. 2019), better leadership skills (Dhuey and Lipscomb 2008), lower likelihood of disability identification (Dhuey and Lipscomb 2010; Elder 2010; Evans et al. 2010), improved college outcomes (Hurwitz et al. 2015), reduced criminal activity (Depew and Eren 2016; Cook and Kang 2016; Landersø et al.

[1] It is worth noting that thus far only a few papers (Black et al. 2011; Dhuey et al. 2019; Landersø et al. 2020; Zang et al. 2020; Persson et al. 2021) have been able to link siblings in the context of school-entry policies, which illustrates how hard it is to combine credible variation stemming from natural experiments with samples including information on individuals beyond treatment and control groups.

2017), and higher wages (Kawaguchi 2011; Fredriksson and Öckert 2014). As noted above, these developmental effects on the focal child could spill over to other children in the family through direct interactions between siblings or indirectly through parental decisions triggered by the effects on focal child. In this study, we consider both of these channels as sibling spillovers as long as they are driven by the effect of the policy on the outcomes of the focal child.[2] Understanding these potential externalities, both their direction and magnitude, should be of interest given the sheer number of families affected by the policy in question.

To examine the spillover effects, we make use of detailed school records from a large, anonymous district in the state of Florida, which enable us to link all children to their place of residence, providing information on household characteristics and sibship composition. Using these data, we first confirm the developmental effects of school starting age policies on focal children. In particular, we find that students born right after the school starting cutoff (and hence more likely to be the oldest in their cohort) significantly outperform students born on the other side of the cutoff (by 16 percent to 23 percent of the standard deviation) on standardized tests in elementary and middle school.[3] We then examine the spillover effects by comparing the educational outcomes of students whose siblings were born in the days before and after the school starting cutoff in an RD design.

We find evidence for within-family spillovers running from an older to a younger child that are concentrated in less affluent households. In particular, our reduced-form RD estimates indicate that students in impoverished families whose older siblings were born right after the school-entry cutoff score about 15 percent of the standard deviation higher on standardized tests in elementary and middle school grades compared to students whose older siblings were born right before the cutoff. We

[2] In principle, school entry policies could also influence the outcomes of other children in the family through channels other than their effects on broadly understood human capital of the focal child. For example, eligibility to start school a year earlier could, through parental employment decisions, affect family income or time that siblings spend with parents; and these in turn could affect the educational outcomes of other children in the family. If these alternative channels exist in our context, our analysis will more likely reveal the "family" effects rather than "sibling spillover" effects of school entry policies. Several studies examined the effects of school entry policies on maternal labor supply or family stability (Landersø et al. 2020; Gelbach 2002; Fitzpatrick 2010a) yielding mixed results. For example, Gelbach (2002), using U.S. Census data, finds positive effects of eligibility to start school a year earlier on the labor supply of mothers and negative effects on public assistance take-up among less affluent households. On the other hand, Landersø et al. (2020) document that mothers of old-for-grade children have increased labor supply and earnings and are more likely to remain married in Denmark, although it is not clear whether these findings apply to the U.S. context given the differences between the two countries in term of access to childcare and public assistance programs. While we are unable to provide direct evidence on this mechanism due to data limitations, we conduct a number of exploratory exercises and present suggestive evidence that it is unlikely to be a major mechanism behind our main findings. For this reason we tend to favor "sibling spillover" over "family effects" interpretation of our results, albeit we acknowledge that to the extent that our indirect evidence is insufficient one can think about the results as implying broader externalities of the school-entry policies on the family. Irrespective of the exact interpretation, the fact is that consequences of the policy extend beyond directly affected child.

[3] It is important to note that these findings reflect how well the focal child is performing in school compared to other students in the same grade, which could be due to school entry age effects, testing age effects, or the effect of being older in peer group (Cascio and Whitmore Schanzenbach 2016). As such, they suggest that the relative test performance of students born right after the cutoff is higher (compared to their peers at the same grade level) than the relative test performance of students born right before the cutoff. While this finding may not necessarily imply that old-for-grade students have higher human capital (as proxied by test scores) compared to young-for-grade students holding age constant, it could be the relevant signal for the parents and other children in the family about the academic ability of the focal child, and this in turn could drive the role model effect and parental resource reallocation decisions.

do not find similar spillover effects for siblings in more affluent households, however; in higher SES families, older siblings appear to be at a disadvantage if their younger brother or sister is born after the school-entry cutoff compared to before. These negative effects, albeit smaller and not as robust as the older-to-younger spillovers in terms of statistical significance, are particularly pronounced when siblings are spaced closer together or when the older child is struggling academically, and appear only after the younger sibling is tested in school for the first time, providing suggestive evidence that school entry policies affect other children in the family through their effect on the educational performance of the focal child rather than any effects on the family resources during the gap year imposed by delayed school enrollment.[4] No support for younger-to-older spillovers is found in poorer households.

Our work contributes to a growing literature on sibling spillovers and within-family externalities of education policies. It is closely related to work by Nicoletti and Rabe (2019), Qureshi (2018b), Joensen and Nielsen (2018), Landersø et al. (2020), Zang et al. (2020), Persson et al. (2021), all of who examine sibling and family spillovers generated by educational shocks or policies in developed countries.[5] Using data from England and North Carolina, respectively, Nicoletti and Rabe (2019) and Qureshi (2018b) employ fixed-effects strategies to estimate spillover effects of having a higher performing sibling in the household. In the former paper, the older sibling's human capital is affected through his/her classroom peers, while the latter study examines the spillover effects of exposure to an experienced teacher. Both papers find evidence for positive spillovers when the older child is affected and conclude that these operate primarily through direct exposure and interaction with a higher ability sibling rather than through parental resource reallocation.[6] In both cases, causality is

[4] It could also be the case that family outcomes, independent of effects operating through human capital of the focal child, are affected beyond the gap year during which the old-for-grade child stays at home for an extra year compared to young-for-grade child. In such a case investigating socioeconomic status of the family only around the start of focal child's schooling would be insufficient. Landersø et al. (2020) provide some mixed evidence on this possibility. On the one hand, they only find statistically significant positive effects of being old-for-grade on maternal employment at very early stages of education, exactly around the gap year that then decline as child ages. Note that this is in stark contrast to the U.S. setting where Fitzpatrick (2010b) finds zero while Gelbach (2002) finds negative effects of being old-for-grade. Irrespective, if fade-out pattern such as the one in Denmark is present in our data then our test during the year gap should provide an upper bound on potential family effects due to changing labor supply. On the other hand, Landersø et al. (2020) documents no effects on marriage early on in child's schooling and these become statistically significant only once child reaches high school which is a time period beyond what we consider in this paper. Overall, we view the data patter documented in Danish data, albeit opposite to what was documented in the U.S., as supporting the notion of testing for family effects during the years surrounding focal child's school-entry age. In Figure A3 we show that family poverty is unaffected up to three years before and up to two years after younger child's school entry age. This supports the notion that externalities in our context operate primarily through human capital of the focal child rather than broader family effects unrelated to this human capital channel.

[5] There is more extensive literature on spillovers generated by health shocks. Studies using U.S. data focus on spillovers from having a disabled sibling (Fletcher et al. 2012; Black et al. 2017), while studies using Danish data consider disability (Black et al. 2017), attention deficit/hyperactivity disorder (Breining 2014), or additional care in the neonatal intensive care unit (Breining et al. 2015). Others examine externalities driven by the iodine supplementation campaign in Tanzania (Adhvaryu and Nyshadham 2016), the immunization campaign in Turkey (Alsan 2017), the deworming program in Kenya (Ozier 2018), experiencing a disease before age 3 in China (Yi et al. 2015), or influenza pandemics in the U.S. (Parman 2015). In settings outside of education or health, Bingley et al. (2020) study sibling spillovers from military conscription in Denmark, and Heissel (2020) studies sibling spillovers from teenage pregnancies in Florida.

[6] Another paper, Qureshi (2018a), uses gender segregation in Pakistani schools to measure the effect of older-sibling schooling on younger siblings, and finds that an increase in older sisters' schooling has a positive effect on

established through fixed-effects models with individual level controls (school-by-cohort-by-subject and child fixed effects to identify peer effects in Nicoletti and Rabe (2019) and school-by-year fixed effects paired with lagged achievement controls to identify teacher experience effects in Qureshi (2018b)) that necessarily require stronger assumptions than natural experiments, such as the one explored in this paper. Furthermore, both studies rely on relatively small and contemporaneous effects on the focal child to study spillovers while only Qureshi (2018b) examines effects for both younger and older siblings. We complement these studies by (1) using a policy that has a large effect on focal siblings' achievement that persists through their elementary and middle school careers; (2) looking beyond contemporaneous effects when either the older or the younger child in the family is affected; and (3) employing a regression discontinuity design.

Another recent study by Joensen and Nielsen (2018) utilizes Danish data and instrumental variable estimation, wherein the instrument generates quasi-random variation in the propensity that an older sibling takes an advanced math or science course in high school.[7] The study finds that older siblings' course participation increases the likelihood that their younger sibling also takes an advanced math or science course. Due to the nature of the policy in question, however, they are unable to explore spillovers from younger to older children. Furthermore, the intervention itself takes place later in schooling, which could yield results different from ours if human capital production function is dynamic in nature and childhood shocks are more consequential than events in early adolescence (Cunha and Heckman 2007; Cunha et al. 2010).

Landersø et al. (2020) is similar to our work in that they also utilize school entry policies as a quasi-experimental variation to study within-family spillovers using Danish data. They show that a child's advantage in school driven by school entry policies affects both their parents and older siblings. Our study diverges from Landersø et al. (2020) in several important dimensions. First, we are able to examine the older-to-younger sibling spillovers in an RD setting which was not feasible in the Danish context. Moreover, unlike in the Danish study, socioeconomic status of the family is at the center of our investigation. Examining spillover effects in both directions and comparing them in high-SES and low-SES families is important because it allows us, under specific assumptions, to disentangle direct (e.g., operating through direct interactions between siblings) and indirect (e.g., operating through parental decisions) channels. Second, we study sibling spillovers in the U.S. context, which could lead to different conclusions given the institutional and socioeconomic differences between the two countries; and our results are indeed different from Landersø et al. (2020). For example, Landersø et al. (2020) find positive spillover effects from the younger to the older sibling when the older sibling is close to their ninth grade exit exam around the school start of the younger sibling. In contrast, we find negative spillover effects from the younger to the older sibling concentrated among high SES families. Furthermore, we document economically meaningful positive estimates when the older child in a family is affected - a sample that they have

their younger brothers' literacy and schooling. It is not clear, however, if these results could carry over to a developed country context in which universal schooling is available for all and gender segregation is minimal.

[7] Relatedly, Gurantz et al. (2020) and Goodman et al. (2020), examine older-to-younger spillovers in high school exam taking and college choices, respectively.

not investigated. Thus, we view the two papers as important complements to each other highlighting the fact that the same educational policy could yield very different externalities in two developed countries with divergent institutional settings.

In a paper written subsequent to the current study, that is closest in spirit to our work, Zang et al. (2020) also examine sibling spillovers in an RD design using school starting cutoffs and data from North Carolina. Similar to our findings, they find positive spillover effects from the older to the younger sibling that are considerably more pronounced in low-SES families albeit these effects, unlike in our case, are limited to middle school grades. On the other hand, they do not find spillover effects from the younger to the older sibling, however, in this analysis they do not incorporate SES stratification. Beyond this important distinction (we likewise do not find statistically significant negative effects in younger-to-older analysis on average), the discrepancy in findings could also be explained by differences in context (between Florida and North Carolina), using different proxies for socioeconomic status or differences in sample selection (e.g., we use all individuals attending public schools while they restrict their sample to students born in North Carolina who are either non-Hispanic Whites or non-Hispanic African-Americans). Nonetheless, evidence from North Carolina and Florida should naturally be treated as complementary.

We document intra-family spillovers in the U.S. educational context using a regression discontinuity design and use this identification strategy to study this phenomenon in both directions in the same data set. Taken together, our results have important implications for understanding human capital production function, household dynamics, and program evaluation. First, our causal estimates depart significantly from the observed correlations between sibling outcomes, strongly supporting initial endogeneity concerns and necessitating credible quasi-experimental designs. Second, these findings suggest that human capital production function should depend not only on the child's own investments and those of his/her parents but also on interactions with other family members. Finally, our causal estimates provide evidence that an educational policy affecting millions of children in the U.S. and globally could create both positive and negative within-family externalities depending on the characteristics of the affected households. To the extent that this finding generalizes to other educational interventions, failing to account for these spillovers in cost effectiveness analyses could lead to misleading policy recommendations.

2 Theoretical considerations

To help interpret our empirical findings, in this section, we present a simple model of human capital production that is consistent with our empirical findings and that incorporates spillovers arising through two possible channels as well as socioeconomic status of the family. In that, given our empirical findings, for now we ignore the possibility of school-entry policies creating intra-family spillovers that do not operate through the effects on human capital of the focal child, but we come back to this important assumption below. In this model, we define human capital in a broad sense, beyond simple measurements of cognitive skills such as test scores. For example, we view traits like leadership or general non-cognitive skills as components of broadly understood human capital that can affect the outcomes and experiences of the focal child. Following prior work (Becker 1993; Currie and Almond 2011; Yi et al. 2015), and to highlight the salient features of our application, we make the following assumptions: (1) there are only two school-aged children in each household (labeled j and k, and j is older child in the family); (2) there are two types of households – constrained (corresponding to families whose children were ever eligible for free or reduced priced lunch (FRPL)) and unconstrained (corresponding to never-FRPL-families)[8]; (3) parents treat their children in a neutral way (i.e., they do not discriminate based on sex, birth order or other innate characteristics); and (4) test scores provide a reliable proxy for children's broader human capital accumulation. We come back to this last point and the plausible effects on other outcomes when mapping predictions of the model onto our empirical application.

Human capital production function is of the following form:

$$\Psi = \mathit{\Psi}(i, \mu_j, \mu_k) \tag{1}$$

where i denotes parental investments; μ_j is a shock to older child j and μ_k is a shock to younger child k.[9] Parents value human capital of their children as well as their own current consumption. Thus, their utility function can be written as:

$$\mathrm{U} = U(C, \Psi_j, \Psi_k) \tag{2}$$

where C is parental consumption while Ψ_j and Ψ_k are accumulated human capitals of the two children in a family. In reality, parents might care about the adult incomes of their children (Becker 1993), but, for simplicity, we assume that human capital is the sole determinant of income. Parents also face the following budget constraint:

$$i_j + i_k + C = Y \tag{3}$$

where i_j and i_k are parental investments in their children and Y is household wealth or resources. Parents maximize their utility function in (2) subject to this budget constraint and production technology of their children by choosing $i = (i_j, i_k)$ and C. Under the standard assumptions on the utility function and the production function, the optimal human capital investment is:

$$i^* = i(\mu_j, \mu_k, Y) \tag{4}$$

[8] This is borrowed from Becker (1993) who examines the family resource allocation separately for low-income and high-income households.

[9] Here Ψ is a value produced according to a function $\mathit{\Psi}$. More traditionally this production function would also include endowments, individual level characteristics, schooling inputs, and household level characteristics (Yi et al. 2015). We chose not to include them in the model for transparency since in our setting they are orthogonal to the school-entry policy. Human capital of a particular child in this model depends on both his/her own school-entry effect and his/her sibling's school-entry effect, but we only consider one degree of contagion (i.e., affected child can have an effect on their brother or sister, through, for example, mentoring, but this direct effect does not feed back to affected child's human capital). An intuitive interpretation of our school-entry policy is to think about it as a change in a child's perceived ability that is an input into his/her human capital production. Thus, a shock in this context should be understood as a causal effect of school-entry policy on focal child's achievement, relative to other children in the cohort, and this information is revealed to both parents and siblings once testing of the focal child commences in grade 3.

Let's now consider how human capital of the sibling (older or younger) is affected by the effect of the school-entry policy (μ) on the focal child in the family (older or younger):

$$\frac{d\Psi_j}{d\mu_k} = \frac{\partial \Psi_j}{\partial \mu_k} + \frac{\partial \Psi_j}{\partial i} \cdot \frac{\partial i_j}{\partial \mu_k} \quad \text{and} \quad \frac{d\Psi_k}{d\mu_j} = \frac{\partial \Psi_k}{\partial \mu_j} + \frac{\partial \Psi_k}{\partial i} \cdot \frac{\partial i_k}{\partial \mu_j} \tag{5}$$

The term on the left-hand side of each equation corresponds to the total effect of the policy for child k (j) on the human capital of sibling j (k), which, in our case, is proxied by test scores. This total effect can be decomposed into a direct effect (the first term on the right-hand side of each equation) and an indirect effect (the second term on the right-hand side of each equation). For the latter term, consistent with prior literature, we assume that marginal productivity of investment ($\frac{\partial \Psi}{\partial i}$) is positive for both children. The signs of the direct effect and the second term of the indirect effect are generally ambiguous, and below we elaborate on some plausible scenarios. Under all these scenarios, we assume that $\frac{\partial \Psi_j}{\partial \mu_j} > 0$ and $\frac{\partial \Psi_k}{\partial \mu_k} > 0$ in the case of causal effects of school entry policy. That is, based on the evidence presented in the prior literature (reviewed in Section 1) and our findings in Figure A1, we assume that being born after the school-entry cutoff, and thus being oldest at the start of school, has positive effects on test scores of all students, irrespective of family SES.[10]

We expect the direct spillover to be positive if higher achieving students serve as better mentors or role-models for their siblings and behave altruistically towards them (Buhrmester 1992).[11] Because older siblings have, on average, greater knowledge and experience than younger siblings, we expect this direct channel to be stronger if the affected child is the older one in the pair: (i) $\frac{\partial \Psi_k}{\partial \mu_j} > \frac{\partial \Psi_j}{\partial \mu_k} = 0$.[12] It is also conceivable that these direct effects, mentoring in particular, could be more pronounced in less affluent households (especially those that are headed by a single parent) where the older sibling could be more likely to take over parental responsibilities.

The sign of the indirect effect, however, potentially depends on households' socioeconomic status

[10] We cannot empirically assess if in our sample the school-entry effects for a focal child persist into adulthood, and the literature investigating long-run effects in this context has been mixed (e.g. Hurwitz et al. (2015) or Fredriksson and Öckert (2014) vs. Dobkin and Ferreira (2010) or Black et al. (2011)); but what matters in our empirical application is that the policy induces September born children to have higher human capital (here proxied by test scores) when we measure outcomes of their siblings i.e., in grades 3 to 8. Figure A1 clearly shows that although some fade-out is present, even the smallest estimated effect - in seventh grade - implies 17 percent of a standard deviation advantage of September born children.

[11] There are two competing mechanisms in this context when considering the tutoring channel. On the one hand, older siblings born after the cutoff perform better academically compared to their classmates (this is also the primary channel through which the role model effect is expected to operate) and they are closer in grade to their younger siblings which could make them more effective tutors. On the other hand, older siblings born before the cutoff have one more year of schooling and could have higher human capital at any point in time thus offsetting any negative effects of being young-for-grade. Therefore, ultimately it is an empirical question whether these spillover effects are positive or negative and our data support the former rather than the latter channel. To the extent that this additional year of schooling of the young-for-grade (control) students has positive effects on their own and their siblings' human capital, it would lead to an underestimate of our coefficients of interest and thus they should be treated as a lower bound.

[12] Alternatively, if a higher achieving focal child leads to feelings of envy or lower self-esteem, this could adversely affect achievement of their sibling (Lavy et al. 2012). We view this negative channel as plausible in both older-to-younger and younger-to-older settings, and we note that it will simply diminish any positive effects from mentoring/role-modeling. In an extreme case where younger child is affected by the policy and there is no mentoring/role-modeling at all we may observe negative direct spillovers. Similar consequences could arise due to sibling rivalry.

to a larger degree. It is plausible to expect that families that are budget constrained are not able to invest in children beyond covering their basic consumption (i.e., $Y = C$ in equation 3), while those that are not budget constrained can invest in children.[13] The former case simplifies the problem because parents do not have a budget slack to make differential investments in children. Furthermore, even if low-SES households could invest in their children, we expect differential investment in these households to be at much lower levels and to be less effective because quality of the inputs is lower (Chiswick 1988). Thus, focusing on the lower bound, we consider the following condition for indirect effects in low-SES households: $(ii)\ \frac{\partial \Psi_j}{\partial i} \cdot \frac{\partial i_j}{\partial \mu_k} = \frac{\partial \Psi_k}{\partial i} \cdot \frac{\partial i_k}{\partial \mu_j} = 0$.

In the absence of the school-entry policy effect (μ), if the siblings have similar perceived abilities in high-SES households, we expect parents to invest the same amount in their children ($i_j^* = i_k^*$) because of the neutrality assumption. On the other hand, if the observed relative test scores of one child is higher than the other, then parents may engage in compensatory or reinforcing behavior. Compensatory behavior implies that both $\frac{\partial i_j}{\partial \mu_k}$ and $\frac{\partial i_k}{\partial \mu_j}$ are positive while the latter implies that both of these terms are negative. In other words, in response to an increased achievement of child k (j), parents will either increase (compensate) or decrease (reinforce) investments in child j (k).

Why would parents, in particular those in high-SES families, ever divert resources from an under-performing to an over-performing child? Becker (1993) provides one explanation for this, assuming that parents maximize adult incomes of their children and the investment strategy is multidimensional. Namely, parents could divert resources toward the most productive child (in our setting, the one with exogenously shifted test scores), and then compensate the other children in the family with either investments in non-human capital goods or provide them with a compensatory lump sum transfer later in life. Interestingly, descriptive evidence based on Health and Retirement Survey suggests that parents who engage in differential spending on post-secondary schooling of their children do not engage in later life offsetting of such differences using cash transfers (Haider and McGarry 2018). On the other hand, Grätz and Torche (2016) show that advantaged parents provide more cognitive stimulation to higher-ability children, and lower-class parents do not respond to ability differences; a behavioral pattern consistent with our argument above. More generally,

[13] Here budget constraint combines both financial and time dimensions, and households which we consider budget-constrained have maximum income of 185 percent of the federal income poverty levels. In the year 2000, this indicated an annual income of 31,543 USD for a family with two children (i.e., household size of four). For the same year, Lino (2000) reports that the total annual expenditure on children (below age 15) in families with before-tax income less than 38,000 USD was in the range of 6,280 to 7,380 USD, and it was about half of the total expenditure among households with before-tax income of more than 64,000 USD. Furthermore, among school-aged children considered in our empirical application (ages 6 to 14), households making more than 64,000 USD spend 2.8 to 3.5 times more on childcare and education than households making less than 38,000 USD. Since our families are almost 20% poorer than those considered as low-SES by Lino (2000), we expect these expenditure differences to be even more striking, lending further credibility for the assumption that in these households $Y = C$. Finally, both Guryan et al. (2008) and Kalil et al. (2012) document that low-SES parents spend much less time on childcare than high-SES parents, especially along quality-adjusted dimensions (Vinopal and Gershenson 2017). Given the evidence in aforementioned studies we expect time and monetary investments to move in the same direction. Were low-SES families substituting money with time (ignoring the quality adjustment issue highlighted by Vinopal and Gershenson (2017)) then we would underestimate older-to-younger spillovers under reinforcement and overestimate those under compensatory behavior. Conversely, for younger-to-older spillovers, we would expect negative effects on older child under reinforcement and positive effects under compensatory behavior. Were high-SES families substituting money with time, then we would observe attenuated indirect effect and thus our estimates should be treated as lower-bounds.

parents will reinforce if marginal rates of return on human capital investments exceed the difference in their marginal utilities (i.e., only if efficiency outweighs equity). Otherwise, they will engage in compensatory behavior and direct resources towards the struggling child.

Overall, the theoretical predictions on how a higher-achieving children affect their siblings are ambiguous. Under some of the scenarios discussed above, however, we can draw testable hypotheses that correspond to our empirical setting and available data (assuming that school-entry policies operate through human capital of the focal child rather than through family effects unrelated to focal child's human capital). In particular, in low-SES households we expect positive effects if the older child is affected by school-entry policy and null effects if the younger child is affected by the school-entry policy (combine (*i*) and (*ii*) above). Conversely, in high-SES families, the effect depends on whether parents compensate or reinforce in response to the effect of the policy on focal child achievement. When the older child is affected by the school-entry policy, positive direct effect (per (*i*) above) will be either offset towards zero by reinforcing behavior or magnified by compensatory behavior. If the younger child is old-for-grade, however, and there is no direct effect, then we expect negative effects if parents engage in reinforcement and positive effects if they behave in a compensatory way.

A subtle point that requires additional clarification relates to what we understand by children's human capital, and in particular whether it includes, for example, non-cognitive skills and disability. This is important given that prior literature documented effects of school-entry policies on behavioral outcomes and disability (Elder 2010; Evans et al. 2010; Dee and Sievertsen 2018) as well as on leadership skills (Dhuey and Lipscomb 2008). We view these traits as components of broadly understood human capital and expect them to generally be affected in the same way as cognitive skills (or at least we did not find evidence suggesting otherwise). Namely, being old for grade confers cognitive (Dhuey et al. 2019) as well as non-cognitive (Dhuey and Lipscomb 2008) advantage and it decreases disability likelihood (Elder 2010). Our data do not include high quality measures of non-cognitive skills (e.g., from military testing) but the data support positive cognitive and negative disability effects. Figure A1 presents evidence that indeed the policy has persistent effects on the cognitive development of focal children while in Table A1 we show that it reduces the likelihood of disability diagnoses - this is especially true for disabilities that are related to behavioral problems. To the extent that disability itself has consequences for other children in the family beyond the affected child (Black et al. 2017), our spillovers from school-entry policy could envelop reduced likelihood of being diagnosed with a disability for the focal child. Given these estimates and findings in prior literature, this data limitation motivates our focus on reduced-form estimates which in this particular context should still be interpreted as sibling spillovers where the effects on the focal child operate through improvements in multiple developmental domains.

Thus far we have discussed implications of the model assuming that the policy affects broadly understood human capital of the focal (treated) child in the family, and then spills over to siblings either through (1) direct interactions between children or (2) indirectly through parental investment decisions triggered by the effect on the focal child. As noted in the introduction, however, such exclusion restriction might not hold if school entry cutoff rule affects siblings through parental choices independent of focal child's human capital; for example, through changes in parental labor supply or family stability (Landersø et al. 2020). Consequences of such adjustments can theoretically have both advantageous or detrimental effects on other children in the family; for instance, if mother returns to labor market earlier, then this could generate additional income for the household but simultaneously limit the time children spend with their mother. We explore the plausibility of these alternative channels in Section 4.2 and conclude that in our particular application - to the extent that we can measure it - they do not seem to be quantitatively meaningful threats to the interpretation of reduced-form effects as stemming from sibling spillovers. Having said that, to the extent that our individual- and neighborhood-level measures of poverty and socioeconomic status of the household may be insufficient to fully rule out the effects on the family not operating though focal child's human capital, it would be preferable to interpret our results more broadly intra-family externalities rather than sibling spillovers.

3 Data and descriptive statistics

Our primary dataset comes from school records in an anonymous, large, and diverse school district in the state of Florida. School districts in Florida are at the county level and are the primary local administrative units responsible for public education. The dataset covers all individuals born between 1970 and 2002 who attended public schools in the district in school years between 1989/90 and 2004/05. There are 311,248 unique children in this dataset. From these data, we know the name, demographic information, and home address of each child, which enables us to form sibling pairs. We define children as siblings if they consistently co-reside at the same address and have the same last name. We also relax this definition and include children who consistently co-reside at the same address but who have different last names, and our results remain unchanged. Our primary outcome variable of interest in this analysis is student test scores in grades 3 through 8, which are recorded as national percentiles and range from 1 to 99.[14] These are all based on standardized tests and principals or teachers do not have input in the grading of the tests. We restrict the sample in our main analysis to adjacent sibling pairs born between 1976 and 1996, who were tested at least once in grades three through eight during our sampling frame.[15] We drop (1) sibling pairs whose

[14]During the time period investigated in this paper, the district used two types of tests: Comprehensive Test of Basic Skills (CTBS) between school years 1989/90 and 1998/99 and Florida Comprehensive Assessment Test (FCAT) between school years 1999/00 and 2004/05. Across these two subsamples, we observe consistently reported scores in mathematics and reading. Since in each case these are national percentiles, they can be compared across the two assessments. One important difference between these two tests, however, is that the former is a low-stakes test, while the latter is a high-stakes test with important implications for schools and students. We do not detect discontinuity in the likelihood of ever taking the low-stakes test in either the older-to-younger or the younger-to-older spillover analyses, and the estimates are 2.28 (SE of 1.61; mean of Y of 63.19) and -1.29 (SE of 1.41; mean of Y of 80.01) for older-to-younger and younger-to-older samples, respectively.

[15]We focus on sibling pairs born between 1976 and 1996 because we require observations of at least two births per family, which creates small cells in both the very early and very late birth cohorts. Furthermore, we observe test scores only in grades 3 through 8, and children from very early or very late cohorts are observed only in grades not covered by testing.

age spacing is greater than eight years because we cannot credibly assert their sibship relationship, and (2) twins and higher order multiplets because they have the same birth date and, thus, no variation in our treatment.

Columns 1 and 2 of Table 1 present the descriptive statistics for all students born between 1976 and 1996 and those students who come from families with at least two children observed in our data, a minimum requirement for sibling spillovers to exist. Nearly 75 percent of students in the district are White, and about 50 percent come from families whose children have ever been classified eligible for free or reduced price lunch (FRPL) – our primary measure of affluence.[16] At the same time, households with two or more children are somewhat more likely to be White and economically disadvantaged. This change in racial composition when focusing on larger families is not a unique feature of our school district but is likewise true for the whole state (see e.g., Breining et al. 2020b). On the other hand, limiting the sample to students who were tested at least once between third and eight grades (column 3) does not further alter the average characteristics. Comparison of these numbers with state averages obtained using the Common Core of Data from school years 1998/99 to 2004/05 reveals that the district is more White (74 percent vs. 53 percent) and slightly less affluent (50 percent vs. 45 percent on FRPL) than the rest of Florida. Furthermore, based on our analyses using American community Survey from 2015, somewhat larger share of K-12 students in the district enroll in private schools (for which we do not have any information) compared to other districts in Florida.[17] It is also the case that our district has slightly higher fraction (54 percent vs. 51 percent in other counties in Florida) of families with two or more children. We observe 263,811 children born in these cohorts in our school records, but limiting the sample to families with two or more children reduces the sample size to 110,872 while a subset of 90,352 students have at least one test score in grades 3 through 8. There are three main reasons why we may not observe test scores for children. First, while children are consistently tested in grades 3 through 8 during our sampling frame, some students are observed only in non-tested grades. Second, if students leave the school district prior to the commencement of testing, we cannot track them across borders. Third, before the 1999/00 school year, students were given a low-stakes test, and during that time frame it was

[16] In particular, we classify a family as low SES if a child in the family was identified as FRPL eligible in school records for at least one school year. During the time frame we examine in our study, students were eligible for FRPL at school if their household income fell below the 130 (eligibility threshold for free lunches) or 185 (eligibility threshold for reduced priced lunches) percent of the federal income poverty levels. When we look at the families who are identified as ever FRPL eligible in our data, we see that their children, on average, are classified as FRPL eligible for 63 percent of the school years during which we observe them in our data. At most, however, 7 percent of our sample is consistently eligible for FRPL across all years; therefore, we do not have power to investigate spillovers separately for families facing very deep poverty (Michelmore and Dynarski 2017).

[17] It is plausible to expect that our analysis sample is negatively selected as we are only able to observe children who attend public schools in the anonymous district. Indeed, based on our calculations using the American Community Survey from 2015, households with at least one child enrolled in a private school have higher median income compared to households where all children were enrolled in public schools (roughly $100,000 versus $55,000 in 2015 dollars) and have better educated parents (for example, 50 percent of fathers in the former group have a college degree or higher compared to 29 percent for the latter). These differences are only slightly higher in the anonymous district compared to the rest of Florida and thus it should provide a good representation, in that dimension, of the whole state. Importantly, these differences should not affect the internal validity of our estimates and indeed in Tables 2, A2, and A3 we document that our treatment and control groups are balanced on the background characteristics (including various SES measures) across the RD threshold.

easier for students to avoid, through exceptions, being tested.[18]

Subsequent columns of Table 1 present characteristics of some of our estimation samples.[19] First, columns 4 and 5 document the composition of our preferred sample that maximizes the number of children we can observe with test scores. Column 4 describes the younger sibling characteristics in the older-to-younger sibling spillover sample, while column 5 describes older sibling characteristics in the younger-to-older sibling spillovers sample. Comparing the characteristics of these spillover samples with all tested children in families with two or more children in column 3 reveals that our preferred estimation samples are very similar. We use these larger samples as our preferred specification to maximize statistical power; however, as columns 6 and 7 reveal, our choice does not appear to generate much of a selection issue when we compare our empirical sample to a plausibly more preferable sample that restricts the analysis to the first two births in each family. This gives us confidence that our results are not driven by focusing on families for whom we combine sibling contrasts across multiple parities.

4 Empirical approach

4.1 Regression discontinuity

We are interested in estimating the causal effect of school-entry cutoff experienced by one of the children in the family on his/her closest sibling, either younger or older. Thus, our RD design compares the outcomes of siblings of children born right before and right after the school-entry cutoff.[20] More formally, we estimate the discontinuity in a child's outcome at the school-entry cutoff faced by his/her older or younger sibling:

$$\beta_{RF} = lim_{S \downarrow 0} E[Y_i | S_j] - lim_{S \uparrow 0} E[Y_i | S_j] \tag{6}$$

where Y_i denotes the outcome of child i. S_j denotes the difference between the birth date of child i's older or younger sibling j and the school-entry cutoff date they faced (i.e. our running or

[18] We formally tested to see whether there exists a concerning discontinuity in the likelihood of having at least one test score in our sample at the student's own school-entry cutoff. Using our preferred bandwidth of 60 days, we found an RD coefficient of -0.303 (0.549), wherein the outcome variable is an indicator multiplied by 100. Given that we observe test scores for roughly 73 percent of the students in our cohorts, this discontinuity is not only statistically insignificant but also trivial in magnitude.

[19] Descriptive statistics are virtually identical when we consider our preferred bandwidth of +/- 60 days for columns 4 to 7 of Table 1. Since descriptive statistics suggest that families in our empirical sample are slightly different from all families with two or more children in the district, we have also re-estimated our preferred models from Table 3 weighted with demographic characteristics (race, ethnicity, and socioeconomic status) of all households with two or more children (column 2 of Table 1). These results point to similar - even more precisely estimated - coefficients thus suggesting that our results may be more broadly generalizable. Since we do not have data for the whole state, we cannot perform similar re-weighting at state-level but we have also re-weighted the main results with characteristics of a full population of districts students (column 1 of Table 1), and the results likewise remain unchanged.

[20] Current policy in Florida stipulates that children must reach age 5 on or before September 1 of the school year to be eligible to attend kindergarten. However, the school-entry cutoff has changed over time, and our oldest cohorts were subject to different criteria. In particular, for those born in 1976, the cutoff was November 1; for those born in 1977, it was October 1; and for those born in 1978 and later, it was September 1. In our preferred sample, we use the cutoff in place at the time the focal child would have turned 5 years old; however, we also show that our main results are robust to using a sample wherein all children were subject to the September 1 cutoff.

forcing variable), with non-positive values indicating dates before the cutoff, so that observations with $S_j \leq 0$ have siblings who were eligible to start school a year before those with $S_j > 0$. For example, for cohorts facing the September 1 cutoff, $S_j = -5$ implies that child j was born on August 27, while $S_j = 5$ implies that child j was born on September 6. For simplicity, let us call a child who receives the treatment (j) the focal child and his/her older/younger sibling for whom we observe spillovers the sibling (i). The parameter β_{RF} is thus the difference in average outcomes of siblings with focal children born just before or just after the school-entry cutoff.

Because our running variable is discrete, following Lee and Card (2008), we estimate β_{RF} parametrically. That is, in our main analysis, we estimate equations of the following form using ordinary least squares (OLS):

$$Y_{ij} = \alpha + \beta_{RF} A_j + k(S_j) + k(S_j) \cdot A_j + \varepsilon_{ij} \tag{7}$$

where Y_{ij} is the outcome for sibling i of treated focal child j, $k(S_j)$ is a linear function of focal child birth date (e.g., -2 for two days before the cutoff or 1 for one day after the cutoff), $A_j \equiv 1(S_j > 0)$ is an indicator for having a focal child who was born after the school-entry cutoff, and $k(S_j) \cdot A_j$ is an interaction between these two variables allowing for different slopes of the linear function before and after the policy cutoff. In a subset of regressions, we add control variables to this equation that include the focal child's school-entry cohort indicators and gender, as well as the sibling's year of birth, month of birth, and gender and race indicators. Our focal child's school-entry cohort fixed effects are at the cutoff year level and control for time-specific shocks that may affect children around a given cutoff (both before and after) differentially across cohorts e.g., hurricane Andrew in 1992 cutoff cohort. In that, these fixed effects do not control for an additional year of schooling that focal children born before the cutoff get as compared to those born after the cutoff. We also include the age spacing between the siblings in days, and when analyzing test scores, we further control for grade and test type indicators. We cluster the standard errors at our running variable – the relative birth day (Lee and Card 2008) – although in Table A11 we test seven different strategies for standard errors computation (Kolesár and Rothe 2018). In our main analysis, we limit the sample to within 60 days of the school starting cutoff (i.e. $-59 \leqslant S_j \leqslant 60$). This bandwidth choice is consistent with other papers using RD design in the context of school-entry cutoff (Cook and Kang 2016), and is at the upper end of a range of bandwidths suggested for various outcomes by data driven, non-parametric bandwidth selection procedures (Calonico et al. 2017). In Section 6, we discuss the robustness of our findings to different bandwidths and functional forms. In particular, we show results using bandwidths ranging from 20 to 150 days, and estimates using a quadratic specification for $k(\cdot)$. We also generate estimates using the non-parametric optimal bandwidth selection procedure developed by Calonico et al. (2017).

4.2 Validity of the research design

In the proposed empirical framework, $\hat{\beta}_{RF}$ provides the causal effect of the focal child's (older or younger) eligibility to start school a year earlier as long as predetermined factors are smooth at the cutoff. Table 2 shows estimated discontinuities in seven variables that should be continuous around the focal sibling's school-entry cutoff date. Panel A presents the older-to-younger spillovers balance (i.e., the discontinuity in younger sibling characteristics at the older sibling's school-entry cutoff date) whereas panel B presents younger-to-older spillovers balance (i.e., the discontinuity in older sibling characteristics at the younger sibling's school-entry cutoff date). Out of the 14 estimates presented in Table 2, none is statistically distinct from zero at conventional levels, and none of the estimates exceeds 5 percent of the dependent variable mean.[21] Importantly, there are no discontinuities in our measure of socioeconomic status whether we treat it as a permanent feature of the family (column 6) or allow it to vary across school years (column 7).[22] Appendix Table A2 replicates this analysis using the non-parametric method (Calonico et al. 2017), and reaches a similar conclusion. Furthermore, Appendix Table A3 shows that balance holds within samples stratified by time-invariant measure of SES, a finding important for our main heterogeneity analysis. Finally, time varying measure of poverty is smooth among families ever classified as those on free or reduced price lunch. This is important because families ever on free or reduced price lunch could experience more transitions in- and out-of-status than families at the margin of becoming eligible for the subsidy.

Further evidence in support of smoothness can be found by examining the distribution of the running variable in our samples (McCrary 2008). We present this graphical evidence in Figure 1 for the distribution of sibling observations around focal child school-entry cutoff and for the analysis of both older-to-younger and younger-to-older spillovers. Because each daily bin in our sample contains relatively few observations (from 94 to 182), and therefore the daily graph is very noisy, we chose to bin the running variable every 5 days for expositional purposes. Nonetheless, we also

[21] We also examined balance in age at school entry measured in days for children for whom we observe test score outcomes. Sample in this analysis is necessarily smaller than the one used in Table 2 as it requires observing a child in both grade 1 as well as in at least one tested grade. Akin to Table A3 we further consider in this analysis not only full sample but also subsamples stratified by our time-invariant measure of SES. Out of six estimates only one (younger-to-older spillover among families never on free or reduced price lunch) is statistically significant but its effect size implies a trivial imbalance of 0.8%. The remaining estimates are even smaller at 0.01% to 0.4%. Finally, when we replicate our main results from Table 3 using this reduced sample, with or without controlling for focal child's age at school start, we obtain results similar to our preferred estimates. Likewise, controlling for age at school start in grade 1 for the sibling of interest (rather than focal child) is not consequential for our estimates, but reduces sample size and thus is not our preferred specification. We also examined the discontinuity in attrition rates across grades as this could affect the internal validity of our results if, for example, siblings of focal children born right after the cutoff are more or less likely to remain in school records either pre- or during testing grades. We execute this analysis by defining dependent variables as probabilities of exiting the sample up to seven years after the first observation. Out of 42 RD estimates (7 periods $\times$ 3 samples (all, ever on FRPL, and never on FRPL) $\times$ 2 analyses (older-to-younger and younger-to-older)) none is statistically significant and they are all small ranging from -0.9 to 2.6 percentage points. Thus, we conclude that differential attrition as a function of either older or younger child's school-entry cutoff is unlikely to affect our estimates. Another balancing/falsification check that one could conduct is to examine birth weight of both focal children and their siblings. Although we do not have access to this information in our data, Dhuey et al. (2019) document no differences in birth weight between children conceived right before and right after the school entry cutoff and they further document no interaction between health at birth and old-for-grade advantage. Based on their findings using state-wide Florida data, we do not view differential health at birth as a threat to our identification.

[22] This distinction matters to the extent that SES could change due to parental responses to older or younger child's school entry cutoff as documented by Gelbach (2002) and Landersø et al. (2020). In the empirical sample, the Pearson correlation between the two variables is 0.67.

provide a formal test, run on daily data for smoothness of density (Cattaneo et al. 2018). We cannot reject the hypothesis of no discontinuity in the density of the distribution at the cutoff, and the p-values are 0.495 and 0.535 for panels A and B of Figure 1, respectively.[23] Overall, we conclude that our data do not exhibit particularly worrisome discontinuities at the policy cutoff that could invalidate the inference and bias our results.

In our empirical framework, similar to any analysis of sibling spillover effects, the analytic sample only includes students with at least one younger or older sibling. This could be problematic when interpreting our reduced-form effects if there is a discontinuity in the likelihood of having a younger sibling at the school-entry cutoff of the focal sibling. For example, if the older child is born right after the school-entry cutoff and perhaps exhibits higher maturity upon starting kindergarten, this could motivate the mother to have a younger baby or have the baby faster, and part of the spillover could envelope the effect of spacing between children. That said, such potential effects should be minimized in our sample since the realization of the treatment for the older child is deferred and occurs after the age of 5 (more likely at the time of first standardized testing in grade 3). Therefore, any kind of fertility considerations would affect only families with larger spacing between births. Nonetheless, to be on the safe side, we formally test the fertility channel in Appendix Table A4, in which we show the likelihood of observing younger sibling as a function of school-entry discontinuity in the older sibling's birth date. Since in our preferred sample we pool siblings across multiple parities to maximize statistical power, we show these potential fertility transitions across three margins: from one to two, from two to three, and from three to four. In all cases, we find statistically insignificant fertility estimates, and in the two larger samples, these insignificant effect sizes are small. At parity three, the relative effect size is large but the sample size is small, and our results are robust to dropping these higher order fertility transitions. Therefore, we do not find support for selection into the sample in the older-to-younger analysis, which bolsters our confidence in these results. Of course, this potential selection issue does not occur when studying younger-to-older spillover effects, because the birth of the older sibling precedes that of the younger.

It is worth noting that our reduced-form estimates measure the effect of the older (younger) child being eligible to start school a year earlier on the outcomes of the younger (older) sibling in a family. As such, these are intent-to-treat effects that, by themselves, are unable to answer how the outcomes of children with siblings who start school a year earlier fare. Scaling the reduced-form effects by the effects of the school-entry policy on the age at which the focal child starts school (i.e., estimating equation 7 with the age of child j at school-entry as the dependent variable) provides an estimate of the effect of the school-entry policy on the likelihood of starting school a year later, provided that the usual instrumental variable assumptions are met. Estimating this effect is useful not only for calculating instrumental variable estimates of the effect of age at school-entry, but also

[23] It can be argued that since our running variable is discrete rather than continuous, we should implement a test that corrects for this data feature (Frandsen 2017). When doing so, assuming $\kappa = 0$, we obtain p-values of 0.229 and 0.003 for older-to-younger and younger-to-older spillovers, respectively. The latter finding of significant discontinuity is puzzling to us given predetermined fertility in this sample but could be driven by e.g., differential geographic mobility. Therefore, to be on the safe side, in Section 6, we also present results wherein we exclude data points near the cutoff. Our results remain unchanged.

for understanding the extent to which the school-entry policies are followed. For example, in the extreme case, if we observed non-zero reduced-form effects but no effect on age at school-entry, one might be worried that our estimates are picking up spurious correlations.

For a subset of cohorts, our school district data allow for direct measurements of the effect of birth date on age at school-entry for the focal child in the family. Ninety-three percent of older focal children within our preferred bandwidth of +/- 60 days start kindergarten on time (90 percent of students born before the cutoff and 97 percent of students born after the cutoff start school on time), and among younger focal children, these numbers are almost identical. Overall, we find that students who were born right before the school starting cutoff were 0.87 years younger when they started kindergarten compared to their peers who were born right after the cutoff (panel I of Appendix Figure A2).[24] These statistics suggest that for the cohorts we study, the school-entry policies were binding and had sizable effects on school-entry age. In fact, our compliance statistics are much higher than those reported by, for example, Dobkin and Ferreira (2010) or Dhuey et al. (2019).

Aside from the standard regression discontinuity checks, as mentioned in Section 2, in order to interpret our findings as sibling spillover effects, we also need to rule out any effects of school-entry policies that affect siblings through parental decisions but without operating through focal child's human capital. This could take place if being able to send the focal child to school one year earlier affects household resources (e.g., through increased income) or time spent with children (e.g., through mother or father not staying at home).[25] To the extent that these mechanisms are more consequential for lower-SES households, where the resources are more limited, we would expect the RD effects on focal child and by extension the spillover estimates in these families to be downward biased. In contrast, Figure A1, if anything, documents slightly larger positive effects for focal children in households ever on free or reduced price lunch. While we are unable to observe parental

[24] It is important to note that if there was no redshirting/fast-tracking, this discontinuity would be exactly one year. Panel II of Appendix Figure A2 suggests that the gap in school starting age is slightly larger for students from low-SES families (0.90 years versus 0.82 years), indicating that compliance with the policy is somewhat lower for high-SES families. The graph also indicates that this difference in school starting age comes from children born after rather than before school-entry cutoff, implying fast-tracking rather than redshirting as the primary explanation for the observed gap. In fact, SES gap in redshirting is less than 1 percentage point compared to almost 10 percentage points for fast-tracking, which suggests that in higher-SES families the resource burden of keeping a child another year in a household could be somewhat reduced for some families whose children are born after the school-entry cutoff. If students coming from higher-SES families are more likely to be fast-tracked then our focal child estimates should provide a lower bound, and indeed they are somewhat smaller than for children coming from lower-SES families as documented in Figure A1. This should lead to smaller spillover effects in the more affluent families. High levels of compliance with the policy, together with RD design and findings in prior literature (Dhuey et al. 2019), ease our concerns regarding monotonicity violations due to differential redshirting/fast-tracking (Barua and Lang 2016). Our results are also very similar in a sample of families whose children start school on-time, however, due to potential selection this test should be treated as only suggestive evidence. Given high compliance rate, if we found very different results in the sample that enters school on time it would have been worrisome but in this empirical application we do not. This gives us some confidence that the results are not confounded by the small share on families that do not comply with the policy. Nevertheless, to gauge what the reduced-form effect might look like with full compliance, one could multiply our estimates by 1.1 in the analysis presented later in the paper.

[25] When the policy affects the younger child in the pair, it could have direct consequences for older sibling who is already at school and undergoes testing. On the other hand, when the policy affects the older child in the pair, then consequences for a younger sibling could be developmental and not manifest until testing starts in a few years down the road.

employment or household income directly in school records, we present some indirect evidence on this plausible alternative channel using FRPL eligibility as a proxy of household poverty.

First, Table 2 shows that the time-varying measure of household poverty (as measured by FRPL eligibility in school records) for the sibling is balanced at the school-entry cutoff of the focal child, suggesting that students whose siblings were just-eligible to start school a year earlier were no more likely to come from low-income families than students whose siblings were born on the other side of the cutoff. This also holds true for time-varying poverty measure among families even on FRPL (Table A3) who could experience more in- and out-of-status transitions. Second, we investigate the FRPL eligibility and characteristics of the neighborhood of residence of the older sibling in the years before, during, and after his/her younger sibling turns 5 (and they become eligible to start schooling).[26] If eligibility to start school a year earlier for the focal child indeed affects parental employment (and hence family income), one would expect a discontinuity in FRPL eligibility of the sibling at the focal child's school starting cutoff in the year the focal child turns 5. Table A5 presents the discontinuity in the FRPL eligibility and neighborhood SES measures of the older child at the younger child's school-entry cutoff during the year before the younger child turns 5 (Panel A) whereas Panel B repeats the same analysis during the year the younger child turns 5. Figure A3 then focuses on a smaller sample and presents event study in free or reduced price lunch status. Irrespective of the exact approach, we do not find any statistically significant changes in family affluence at the time when younger child is eligible to enter kindergarten.

Third, in Table A6 we provide balancing check akin to analyses in Table 2 but based on alternative proxies for family socioeconomic conditions which we measure at micro-neighborhood of residence.[27] Irrespective of birth order of the focal child we do not observe any discontinuities in the affluence of neighborhood at the time when their siblings attend school. Furthermore, there is likewise no significant discontinuities when we narrow the bandwidth of investigation around younger child's school entry age (Table A5). Were the ability to send a child to school a year earlier to affect family resources in a meaningful way, we would expect these families to move, for example, to a better neighborhood. We observe no such pattern in our data and all estimates are not only statistically insignificant but are also very small.

Fourth, in Table A7 we analyze parental compliance with school-entry cutoff rules for the younger child as a function of their older sibling's eligibility to start school. It is plausible that the older sibling's eligibility to start school affects parents' decision to fast-track or redshirt the younger child, which could in turn affect the educational outcomes of the younger child. Such effects would, however, stem from parental decisions that are potentially independent of the effect of school entry policy on the focal child, and thus should not be considered as sibling spillover in the context of our

[26] We are unable to execute the same analysis for older children affected by the policy because our time-varying measures of family affluence are tied to school records and younger siblings naturally have not yet started schooling at the school starting cutoff of their older brother or sister.

[27] These neighborhood units reflect division at a very fine level, because the school district divides county into more than 1,000 micro-neighborhoods for the purposes of school assignment and school bus routing and scheduling. While the neighborhoods vary in size, between 50 and 200 students on average live in each unit at any given time. We merge this information with census block data on income and demographics sourced from the American Community Survey.

model. Table A7 reveals no statistically significant effects on the school-entry timing of the younger child irrespective of family affluence.[28]

Based on the above set of results we conclude that in our particular context, unlike in findings offered by Gelbach (2002) or Landersø et al. (2020), there is not much empirical support for quantitatively meaningful effects of the policy on household affluence and resources which could impact siblings beyond direct effects on broadly understood human capital of the focal child in our empirical setting. Therefore, throughout the remainder of the paper, we continue to interpret our reduced-form estimates as documenting sibling spillovers (either direct or indirect). Having said that, to the extent that our evidence is limited due to data availability and there are indeed unobserved changes to family resources caused by school-entry policy operating beyond its effects on focal child's human capital, one may prefer to interpret our results more broadly as reduced-form intra-family externalities rather than more narrowly through the lens of sibling spillovers only.

5 Results

Table 3 presents our main results for averaged mathematics and reading test scores measured as national percentiles and pooled across grades 3 through 8. Panel A examines spillovers from older-to-younger child (i.e., the dependent variable is younger child test scores, and the running variable is the birth date of the older child relative to his/her school-entry cutoff), while panel B examines spillovers from younger-to-older child (i.e., the dependent variable is older child test scores, and the running variable is the birth date of the younger child relative to his/her school-entry cutoff). In columns 1 through 3, we present the reduced-form effects using the entire sample estimated without any controls in column 1; introducing own covariates in column 2 including race, gender, grade, birth year, and birth month fixed effects; and adding focal sibling characteristics in column 3 including the sibling's gender, birth cohort fixed effects (controlling for time-specific cohort shocks rather than additional year of schooling that focal children born after the cutoff receive), and the age difference between the two siblings. Columns 4 through 6 repeat the same analysis using sibling pairs in families whose children were identified as FRPL eligible at least once in our sample, and columns 7 through 9 utilize sibling pairs in never-FRPL eligible families. Several findings are worth highlighting in Table 3.[29]

First, we find significant positive spillovers from the older child to the younger. In particular, the findings in the first three columns of panel A suggest that students whose older siblings were born after the school entry cutoff score 1.3 to 1.7 percentiles (or 5.1 percent to 6.4 percent of the standard deviation) better on standardized tests compared to students whose older siblings fall on

[28] Although some of these insignificant estimates are large in terms of effect sizes, this is driven by the fact that non-compliance with the cutoff rule is so small. On the other hand, for on-time start the estimated effect sizes imply trivial imbalance of at most 1.1 percent.

[29] We have also analyzed OLS relationship between test scores of siblings and focal children in the family. These descriptive, rather than causal, estimates suggest inter-sibling associations in the range of 0.42 to 0.51 percentile points that are highly statistically significant and positive irrespective of the exact sample used. The difference between these correlations and our causal estimates highlights the need for a credible quasi-experimental design.

the other side of the cutoff, and these estimated effects are statistically different from zero at the 5 percent level in columns 2 and 3. Estimated effects presented in columns 4 through 9 of panel A further indicate that these spillover effects are entirely driven by sibling pairs in less affluent families. In ever-FRPL-eligible families, students with an older sibling born after the cutoff score roughly 3.9 percentiles (or 14.8 percent of the standard deviation) better on standardized tests in elementary and middle school grades compared to students with older siblings born right before the cutoff. In contrast, the estimated older-to-younger effect in more affluent families is less than 1 percent of the standard deviation in magnitude in our preferred specification that controls for both own and sibling characteristics. This difference in older-to-younger spillovers between poorer and richer households of 4 percentiles is statistically significant at the 1 percent level. Second, the findings presented in columns 7 through 9 of panel B reveal modest negative spillover effects from the younger child to the older in more affluent families that are statistically significant at 10 percent level. In particular, we find that older children in more affluent households in which the younger child is endowed with the school-entry advantage, perform worse (by 1.8 percentiles or roughly 8 percent of the standard deviation) than older children in households in which the younger child is relatively more disadvantaged due to their school-entry eligibility.[30]

To assess the magnitudes of these spillover effects, it is helpful to compare them to other estimates in the family and education literature, and in particular to those obtained using data from Florida. For instance, an effect size of almost 15 percent of a standard deviation is about three times the effect size of a 10 percent increase in birth weight (Figlio et al. 2014) and it is over 50 percent larger than the size of birth order gap in reading scores (Breining et al. 2020a). More related to the current variation and research question, this effect is about two-thirds the size of own school-entry cutoff effects (Dhuey et al. 2019). Furthermore, it is approximately equivalent to an increase of \$3,000 in Earned Income Tax Credit (EITC) income, or three times the difference in average EITC payout today versus that in the early 1990s (Dahl and Lochner 2012); or three-quarters of the effect of assignment of a small class for kindergarten through grade 3 on test scores; or 50 percent larger than the effect of the same assignment on ACT/SAT performance (Krueger and Schanzenbach

[30] We also examined the spillovers separately for mathematics and reading, and these results are presented in Table A8. Positive estimates in older-to-younger sample with families ever on FRPL are statistically significant in both subjects, whereas in our younger-to-older analysis among never-FRPL families statistically significant effects are driven by reading only. This is consistent with the evidence presented in the literature in the context of teacher value-added which shows that teachers have larger effects on the math performance of their students compared to reading (Hanushek and Rivkin 2010). This literature argues that such divergence can be attributed to reading being affected more by home production inputs. We have also examined the stability of our results across grades and the preferred estimates are relatively stable in grades 3 to 8 for younger-to-older spillovers among never-FRPL families while they modestly increase over grades for older-to-younger spillovers among ever-FRPL families. Importantly, akin to the pooled regressions, we do not estimate any statistically significant effects in the two remaining samples and these estimates are very similar across grades. We present these results in Figure A4 noting that standard errors are larger in this specification and not all grade-by-grade estimates in the two samples of interest remain statistically significant. This is expected since pooling across grades decreases the potential consequences of measurement error, and thus improves precision. In a separate analysis, we also add grade-difference fixed effects (between the two siblings) to the controls – in analyses in which we do not include focal child covariates – to gauge the extent to which our estimated effects are driven by the difference in "time spent with the sibling at the same school" between students whose siblings were born on either side of the cutoff. Estimates remain very similar, and thus we discard this channel as a potential confounder.

2001). In terms of the test score gaps between different student groups in our data, the positive spillover effect in the older-to-younger setting of 3.9 percentile points is roughly 20 percent of the gap between students in low- and high-SES families and 21 percent of the White-minority gap. At the same time, the negative spillover of 1.8 percentile points in the younger-to-older setting among higher-SES families, which is about 55 percent smaller in absolute terms compared to our positive finding, is approximately 9 percent of the socioeconomic status gap or four-fifths of the gender gap in this sample. Overall, we view our estimated effects in both settings as not only plausible but also as economically meaningful.

Comparing our estimates with other studies on sibling spillovers in education, we find that they are larger not only in absolute terms but also in relative terms when compared to first-order effects of peers (Nicoletti and Rabe 2019) or teachers (Qureshi 2018b). For example, Nicoletti and Rabe (2019) estimate spillover effects that are about 11 percent of the main effect of peers while Qureshi (2018b) estimates spillover effects that are up to 30 percent of the main effect of teachers. In our case, the first-order effects of school starting age are 6.6 and 5.4 percentiles for ever- and never-FRPL families (Figure A1), and thus the positive spillover is about two-thirds while the negative spillover is about one-third of the corresponding main effect. At the same time, these seemingly large effect sizes are much smaller than the only statistically significant spillover effect found in Landersø et al. (2020) where the point estimate is almost one standard deviation; as compared to our spillovers of at most positive 15 percent of a SD or negative 8 percent of a SD.[31] On the other hand, our estimates are somewhat larger than those from North Carolina where Zang et al. (2020) finds older-to-younger spillovers in middle school grades of up to 4.0 and 7.5 percent of a SD among all and socioeconomically disadvantaged families, respectively. They do not examine younger-to-older spillovers by SES or after the focal child's starts being tested (a critical sample if one wants to examine potential parental responses in the context of school-entry policies) but their point estimate among all families is 1.8 percent of a SD which is similar to our coefficient of 2.2 percent of a SD, and in both cases these are statistically insignificant.[32]

It is harder to compare our effect sizes to other spillover studies within the educational domain as these focus on course participation (Joensen and Nielsen 2018), exam taking (Gurantz et al. 2020) or college choices (Goodman et al. 2020), but we can also contrast our effect sizes with

[31] It is also worth highlighting that our estimated positive and negative spillover effects are within the 95 percent confidence intervals of all the insignificant test score estimates presented in Landersø et al. (2020). In fact, for certain outcomes like oral Danish examination at 7 to 9 years of spacing between siblings, they are unable to rule out negative effect sizes as large as 69 percent of a SD. Interestingly, this finding is for the same sample for which they find positive and statistically significant effects of almost one standard deviation in mathematics.

[32] Interestingly, the positive results reported in Zang et al. (2020) are limited to middle school grades only and they don't find much of an effect in elementary school grades. In contrast, we find positive and statistically significant spillovers in both elementary and middle school grades. In grades 3 to 5 these are 5.5 and 13.1 percent of a SD while in grades 6 to 8 these are 7.6 and 13.1 percent of a standard deviation for all and socioeconomically disadvantaged families, respectively. Both sets of results are statistically significant at conventional levels. Conversely, irrespective of grade level we do not find any statistically significant effects for high-SES families with trivial estimates of -0.1 and -0.4 percent of a SD for elementary and middle school grades, respectively. Our negative estimates in younger-to-older analysis of high-SES families are driven by middle school grades which makes sense if we expect these effects to appear only after the younger child begins testing in grade three at which time majority of their older siblings would have transitioned to middle school.

spillovers generated by health shocks. For example, Black et al. (2017) estimate negative spillovers of 5 percent of SD from growing up with a disabled younger sibling in Florida while Ozier (2018) estimates positive spillovers of up to 40 percent of a SD from deworming program in Kenyan schools. Our negative estimate is thus very similar to Black et al. (2017) who use state-wide Florida data.

In the younger-to-older exercise, we can also examine the timing of the spillover effect from the younger child to the older one. This is important because if the observed spillover effects are indeed driven by the school-entry policy and its effects on the focal child, one would expect to find no significant discontinuity in the older sibling test scores before the younger child is tested for the first time. Conversely, if our estimates pick up adjustments due to the effects of policy on the household (e.g., maternal labor supply) that do not operate through younger child's human capital (as proxied by test scores), then one would expect statistically significant negative (e.g., if loss of income from postponed maternal employment dominates) or positive (e.g., if additional year with mother at home dominates) estimates on older child's test scores prior to the commencement of testing of the focal child. We present the results of this analysis in Table A9, in which panel A pools all families together, while panels B and C present estimates separately for lower- and higher-SES families. The negative effects documented in Table 3 are entirely driven by test scores of older children measured after the younger child starts being tested, and thus allowing parents to observe proxies of relative ability for both children in the family. These negative effects are also concentrated in more affluent rather than less affluent families, and even though the sample sizes are smaller, we have enough power to reject statistical significance of coefficients across the two periods.[33] This set of results paired with no changes in affluence levels at either individual household or neighborhood level documented in Figure A3 as well as Tables A5 and A6 strongly suggest that the estimates are unlikely to be confounded by effects of the policy on siblings that do not operate through the focal child's achievement.

These findings further suggest that the negative coefficient in Table 3 is likely an underestimate of the younger-to-older spillover effect because it pools together test scores of an older sibling before and after the younger child's school entry age advantage is revealed to parents through testing. This intuition is confirmed in Table A9 where we find a much larger and more precisely estimated coefficient of negative 14.6 percent of the standard deviation in the latter sample when we restrict the older sibling test scores to years after the younger child's third grade test. Note that this effect size is very similar to what we have estimated for older-to-younger spillovers among less affluent families where a younger child is more likely to be treated in all grades. We chose to focus on the smaller negative estimate (representing a total effect) in the main results because it is likely more relevant from the perspective of understanding how school-entry cutoff policies propagate across siblings. Finally, returning to Table 3, we do not find any statistically significant spillover effects from the younger to the older child in less affluent families. It is also worth noting that our

[33]Coefficient in column 2 in panel C is also negative but more than three times smaller than that in column 3. This could happen if we expect some smaller behavioral adjustment already after the younger child starts schooling but before the start of testing e.g., due to differential size or maturity of old-for-grade children. We do not have enough statistical power to divide this pre-period into years before the start of schooling and those after the start of schooling and before the start of testing.

estimated coefficients remain stable with the addition of own and sibling characteristics, lending further credibility to the identifying assumptions discussed in Section 4.2.

To fully understand the variation in our data, Figure A5 presents the daily means and linear fits for the bandwidth of 60 days around the focal sibling school-entry cutoff for the full sample, ever-FRPL-eligible families, and never-FRPL-eligible families; and points toward the same conclusion. These findings confirm a relatively larger discontinuity in the test scores of the younger sibling as a function of the older child's school-entry cutoff among lower SES families and conversely a smaller discontinuity in the opposite direction in the test scores of the older sibling as a function of the younger child's school-entry cutoff among higher SES families.

Our empirical findings are consistent with mechanisms described in Section 2 provided that our analyses above help ruling out family effects of school-entry policy that are not operating through focal child's human capital. First, our results support the hypothesis that the direct effects of having a higher-performing sibling are driven mainly by the role model/mentoring effect that is positive and more prominent in the older-to-younger setting. Second, we find evidence consistent with reinforcing behavior among high-SES households, which could be offsetting the positive direct effect in the older-to-younger setting, and leading to negative spillover effects in the younger-to-older setting. Had compensatory reallocation been at play, we would expect parents to divert resources to the older siblings from the younger sibling who was advantaged by the school starting cutoff, ultimately leading to the higher achievement of the older. However, no such pattern emerged in our data. Finally, the results are also consistent with negligible indirect effects in low-SES households, which, combined with the direct mentoring effect, lead to significant positive spillover effects from an older child to the younger, and no spillover effects in the younger-to-older case.

It is important to highlight again that school-entry eligibility of the focal child in the family could also affect the sibling by altering the time the focal child spends at home. For example, the focal child starting school a year earlier could free up parental resources (e.g., in the form of preschool tuition), which could in turn benefit the sibling. This channel could be particularly relevant for more resource-constrained households, however, the pattern of our results is not consistent with this hypothesis. First, we find that students whose older siblings were born after the cutoff (and hence were more likely to start school a year later) perform better in low-SES households. Second, while our results seem to support this hypothesis in the high-SES, younger-to-older spillover case, we find no such evidence in the low-SES setting, in which one would expect this channel to be stronger. Furthermore, in Table A9 we show that older siblings' test scores are affected only once the younger child starts being tested and their relative ability is revealed. Nonetheless, a caveat remains: we do not have information on actual investments in children, and, therefore, we equate observed differences in outcomes with particular pattern of parental behavior.[34] Berry et al. (2020) is better

[34]Sibling rivalry could also be a potential mechanism behind our observed spillover effects. In this scenario, a sibling could be negatively affected by direct competition for resources with the focal child who is higher-achieving due to school-entry policy. This is plausible but it would have to be that (1) sibling rivalry is more pronounced in more rather than less affluent households, (2) that it is differentially consequential for younger siblings compared with older siblings, and (3) that it manifests only after commencement of testing of the younger focal child. On the other hand, one can argue that competition for resources should be stronger in less affluent households where they are more

suited to answer questions about parental preferences for investment in children as they conduct lab-in-the-field experiment in Malawi to elicit those. They find that although parents minimize inequality in inputs across siblings and maximize total household earnings they are simultaneously not averse to inequality in outcomes across siblings. At least the latter part of their evidence would be consistent with some of our findings but it is not clear whether such an experimental evidence could be generalized to developed country context. Alas, we have to trade off administrative data on "market outcomes" with eliciting preferences in the experimental context.

Several factors are expected to moderate these spillover effects. For example, it is plausible that parents with children who are closely spaced in age are more likely to engage in reinforcing behavior given that it is easier to compare the revealed abilities of their children.[35] Figure A9 presents estimates for our four samples wherein we divide spacing between the two adjacent siblings by the median. Consistent with our hypothesis above, in relatively more affluent families in which the younger focal child starts school later, we observe more negative effects when siblings are spaced more closely together. The results also indicate that the positive older-to-younger spillover effect in less affluent families is stronger among siblings spaced above the median, which makes sense if the older siblings serve as better caregivers/mentors/role models for the younger ones when the age spacing is larger. Similar to our main results in Table 3, we do not find any sizable estimates in either older-to-younger analysis among high-SES families or younger-to-older analysis among low-SES families.

The achievement level of the sibling is another potential moderator of the indirect spillover effects. For example, the negative spillover effect from the younger to the older child in more affluent families could be more pronounced when the older sibling is academically struggling, which implies that the marginal rate of return on investment in the younger child is relatively higher. Figure A10 presents unconditional quantile regression estimates for the 25th (navy bars), 50th (orange bars) and 75th (maroon bars) percentiles of the test score distribution, again for our four samples defined earlier (Firpo et al. 2009). Similar to the mean results, we find the largest positive spillovers from older-to-younger child in lower SES families, and these effects are comparable across the distribution of the younger sibling's ability. Corresponding estimates for higher SES families are

limited. Since we do not have any measures of sibling rivalry in our data and we could not find any empirical evidence in the literature supporting this exact pattern of behaviors matching our results, we find this explanation less plausible than a more standard parental resource reallocation scenario that has been proposed theoretically (Becker 1993) and documented empirically in some contexts (Grätz and Torche 2016). That said, we acknowledge this limitation of our data and leave such an alternative explanation as potentially fruitful avenue for future research.

[35] Another possible moderator of sibling spillovers is gender composition – i.e., same-sex vs. opposite-sex pairs. This comparison is more problematic in our view because of scope for gender-specific investments (Barcellos et al. 2014, Karbownik and Myck 2017) and differential sensitivity of boys and girls to inputs (Bertrand and Pan 2013, Autor et al. 2019). It could further be very context specific and we could not verify if Florida families in our data exhibit any particular gender preferences. Nonetheless, for completeness, in Figures A6, A7 and A8 we also examine differential effects by sex composition of the sibling pair (Figure A6) as well as by either gender of focal child (Figure A7) or of their sibling for whom we observe outcomes (Figure A8). We find larger positive spillover effects from the older to the younger child in low-SES families when children are of the same-sex. Likewise, estimates in this sample are larger for focal male children and male siblings. On the other hand, we find negative effects from the younger to the older child in high-SES families with opposite-sex children, when the focal child is a female, and to a lesser degree if older sibling is a male. Our estimates in the remaining two samples are not consistently signed, never statistically significant, and generally smaller which is in line with the other analyses throughout the paper.

much smaller and positive only in the upper half of the test score distribution. Moving to younger-to-older spillovers, akin to mean estimates, we do not find any sizable effects in low-SES families, but in line with the hypothesis proposed above, results for more affluent households suggest much larger point estimates when the older sibling is struggling academically. All estimates in this sample are negative, but the spillover effect for the 25th percentile of the distribution is almost eight times larger than the effect for the 75th percentile. This result is consistent with Grätz and Torche (2016), who find that advantaged parents provide more cognitive stimuli to higher-ability children, which, in our case, could imply that high-SES parents observe the relatively higher-performing younger sibling whose advantage is driven by the school-entry policy, and allocate more resources toward (reinforce) this child because a higher return on the investment is more likely (Becker and Tomes 1976).

6 Robustness

This section presents a series of analyses to gauge the robustness of our two main test score findings presented in column (6) in panel A and column (9) in panel B of Table 3.[36] We begin by re-estimating our parametric models using alternative bandwidths and both linear and quadratic specifications for $k(\cdot)$. Panel A of Figures 2 and 3 present the results of this robustness check, wherein the blue squares depict the point estimates along with 95 percent confidence intervals for the linear polynomial specification (bandwidth indicated on the x-axis), and orange circles show the corresponding effects for the quadratic specification. These are in line with the main result presented in Table 3. As expected, the quadratic estimates are less precise, and both polynomial specifications are less precise at smaller bandwidths. Nevertheless, almost all coefficients in Figure 2 are statistically significant at the 5 percent level, and the coefficients in Figure 3 are consistently negative except for very small bandwidths in the quadratic specification.[37]

Second, in panel B of Figures 2 and 3, we re-estimate the main results using donut-RD models where we remove observations very close to the focal child school-entry cutoff. The test score results are remarkably stable irrespective of the number of observations that are dropped and, if anything, they modestly grow at the larger donuts. This is consistent with no manipulation at the cutoff and balance in auxiliary covariates.

Third, in panel C of Figures 2 and 3, we present a falsification exercise estimating spillover effects using placebo cutoffs away from the true school-entry cutoff, in which we consider alternative dates up to 100 days on either side of the threshold.[38] This is akin to the randomization inference idea,

[36] For brevity, we focus this discussion primarily on test scores in older-to-younger spillovers among those ever on FRPL and younger-to-older spillovers among those never on FRPL. We present the analyses for the remaining four samples in Figures A11 to A14.

[37] To be precise, in Figure 2 all coefficients in linear specification are statistically significant at least at 10% level while in the quadratic specification, this is true in 23 out of 27 coefficients. In Figure 3, 14 out of 27 coefficients in the linear specification are statistically significant at least at the 10% level while in the quadratic specification none of the coefficients reach conventional levels of significance. This could potentially be explained by overfitting of quadratic specification in relatively smaller samples with modest estimated effect sizes.

[38] For example, if the threshold is September 1, 1985, then in the two most extreme cases, we are comparing siblings

and we label it as such. Note that, even if the true effect at the real cutoff were non-zero, we would still expect to find some "effect" when we move the threshold by a small number of days. Similarly, we would expect some non-zero effects with the alternative cutoffs due to random chance. Nonetheless, if we are finding true effects, we would expect the estimates based on placebo cutoffs to follow roughly a bell-curve shape distribution with the mass centered at zero and away from the preferred RD estimate. This is certainly the case in panel C of Figure 2; however, in Panel C of Figure 3, this pattern is not as distinct. That said, the point estimate from column 9 in panel B of Table 3 corresponds to the 12th percentile of the distribution in the latter graph, which is not far off the p-value of 0.070 reported in the main regression table. Overall, we view these results as supporting the argument that our main spillover effects reflect causality rather than other confounders.

Fourth, in panel D of Figures 2 and 3, we present the main effects estimated using a variety of alternative samples to ensure that the findings presented in Table 3 are not driven by the particular set of observations we use in our analysis to maximize statistical power. Here, we test seven alternative samples, including cases in which the school-entry cutoff is consistently September 1 (second bar); cases wherein we observe full birth order (third bar); cases focusing on the first two births only (fourth bar); cases focusing on births at parity one and two only (fifth bar), cases where we can observe panel of test scores (sixth bar); or cases in which we observe information on early inputs of the focal child (seventh and eighth bars). Irrespective of the sample, in Figure 2 the estimates are in the same ballpark as our initial choice, and they are statistically significant at the 5 percent level in six out of seven cases and at the 10 percent level in all cases. Likewise, in Figure 3, five out of seven specifications are statistically significant at the 10 percent level, and all estimates are consistently negative. Our results are, therefore, robust to reasonable alternative assumptions on which sample of students is included in the analysis.

Given that our results are weaker statistically in some specification in Figure 3, we have also analyzed robustness of the results based on column (3) and panel C of Table A9 where we drop observations for older sibling prior to commencement of testing of younger focal child. As explained in Section 5 this is a sample where we expect the largest negative effects due to potential reinforcement according to predictions from our model outlined in Section 2. These results are presented in Figure A15 and despite 60 percent smaller sample size they imply, as expected, very robust and much larger quantitatively estimates. In particular, in panel A, 23 and 16 out of 27 coefficients are statistically significant at 10% level in linear and quadratic specifications, respectively. The insignificant estimates are concentrated at lower bandwidths which could be driven by the overfitting of the quadratic specification. Subsequent panels (B to D) likewise provide evidence for robustness of the negative estimate.

Fifth, we re-estimate all our main results using the non-parametric optimal bandwidth selection approach proposed by Calonico et al. (2017). These estimates are provided in appendix Table A10,

of focal children born +/- 60 days around July 4, 1985, or focal children born +/- 60 days around October 31, 1985. Note that we do not perform any bootstrapping here but, rather, simply plot a density distribution over 201 coefficients.

and are in line with our parametric estimates. This procedure, however, produces larger standard errors than our polynomial approach; therefore, our result in column 9 of panel B is no longer statistically significant at conventional levels. Larger standard errors might be due to the discrete nature of our running variable, while the data-driven method requires a continuous running variable. Importantly for our interpretation, the point estimates for test scores among higher SES families in younger-to-older spillovers are very similar across the two estimation methods.[39]

The positive spillovers from the older child to the younger child in relatively less affluent families are precisely estimated irrespective of the exact specification or sample. The negative effects from the younger child to the older sibling in families never on free or reduced price lunch, however, are smaller quantitatively and weaker in statistical sense. In the latter case, in the majority of specifications p-values are in the range between 0.05 and 0.10, however, they rarely reach 5% significance levels. As we explained in Section 5 this is partly due to pooling zero effects before the younger child starts being tested and negative effects afterward. Furthermore, these efficiency concerns occur in part because we chose to report as our preferred specification the most conservative standard errors adjustment across all plausible estimation methods. In that, we cluster the standard errors at the level of running variable even though it is discrete.

Such an estimation approach has been questioned recently (Kolesár and Rothe 2018), and thus in Table A11 for completeness we report seven sets of 95% confidence intervals across three estimation strategies. In particular, we compute confidence intervals based on (1) Eicker-Huber-White heteroskedasticity-robust standard errors; (2) standard errors clustered at the running variable level; (3) standard errors clustered at the family level (to account for correlation in errors between pairs of siblings coming from the same household; approximately 6% of our empirical sample); (4) standard errors clustered two-way at the running variable and family levels (to jointly account for specifications (2) and (3) above); (5) standard errors based on heteroskedasticity-robust nearest neighbor variance estimator by Calonico et al. (2017); (6) standard errors based on running variable cluster-robust variance estimator by Calonico et al. (2017); and (7) standard errors for honest confidence intervals based on bounding the second derivative (Kolesár and Rothe 2018).

Considering our main positive finding for older-to-younger spillovers in less affluent families, none of confidence intervals cross zero. In case of the main negative finding for younger-to-older spillovers in more affluent families, two out of seven confidence intervals cross zero, however, in both of these cases the 90% confidence interval excludes zero. Thus, these estimates would be considered statistically significant at 10% level - lowest conventionally considered level of statistical significance. Furthermore, in the former sample, the most positive upper-bound implies an effect size of +21.7 percent of a standard deviation while the least positive lower-bound implies an effect size of +4.8 percent of a standard deviation. In the latter sample, the most negative lower-bound implies an

[39] For 8 outcomes and 32 specifications/samples under investigation across the balancing check and the main results, the optimal bandwidths range from 31 to 67 days, which motivates our preferred bandwidth of +/- 60 days at the upper bound of the range produced by the data-driven procedure from Calonico et al. (2017). Furthermore, the younger-to-older estimate for families never eligible for free or reduced price lunch becomes -3.033 with standard error of 1.277 (p-value of 0.018) when we focus on older sibling test scores after the younger focal child starts being tested (c.f., column 3 and panel C in Table A9).

effect size of -21.8 percent of a standard deviation while the least negative upper bound implies an effect size of +0.3 percent of a standard deviation. As a comparison, the preferred estimated treatment effects in these two samples are positive 14.8 and negative 8.0 percent of a standard deviation, respectively. Thus, these potentially positive effects for younger-to-older spillovers among more affluent families are minuscule and are obtained using the non-parametric method developed by Calonico et al. (2017) which is not well suited for discrete running variables. On the other hand, honest confidence intervals proposed by Kolesár and Rothe (2018) generate 95% confidence intervals that do not cover zero. Importantly, even the more lenient 90% confidence intervals always cover zero in the two remaining samples where we did not find any statistically significant effects in the main specification (older-to-younger spillovers among affluent families and younger-to-older spillovers among impoverished families). We view this discussion on the efficiency of our estimates as important but secondary to whether the coefficients are biased or not and whether the treatment effects are meaningful in economic terms or not. Given our extensive testing of RD-assumptions and benchmarking of estimates discussed in Sections 4.2 and 5, we conclude that our findings reveal unbiased and economically meaningful positive spillovers when the older child in a relatively less affluent family is positively affected by the school-entry policy and smaller negative spillovers when the younger child in a relatively more affluent family is positively affected by the school-entry policy.

7 Conclusions

Common perception and correlational evidence suggest that siblings exert important influences on each other. However, does the academic performance of one sibling causally affect the performance of another? Or do correlations in this case reflect just common unobservables? This is an important policy question: If there indeed is a causal link between the educational achievement of siblings, interventions that improve the outcomes of one sibling could have larger benefits than those typically considered in policy evaluations and further reduce inequality.

In this study, we examine the spillovers in the context of school-entry policies. A number of studies in the last decade have shown that these policies have profound effects on the cognitive and non-cognitive skills of students, with students born after the school-entry cutoff (and hence more likely to be the oldest in their cohorts) significantly outperforming students born before the cutoff. Here, we examine whether and how this policy-driven academic advantage spills over to the siblings of these students. This is important from policy perspective to the extent that school-entry rules affect millions of children in the U.S. and globally and thus even small - positive or negative - spillovers could translate into economically meaningful aggregate educational grains or losses.

Using student-level administrative data that enable us to identify siblings and an RD design, we compare the educational outcomes of students whose older or younger siblings were born in the days before and after the school-entry cutoff. We find statistically significant positive spillover effects from older-to-younger siblings that are entirely driven by children in less affluent families. We do not find any evidence of statistically significant spillovers from younger-to-older children on average or in the less affluent sample. Interestingly, however, we do find suggestive evidence for negative spillover effects among more affluent families (1) that are larger for low-performing older children, (2) that are larger in families where siblings are spaced closer together, and (3) that manifest only after the ability of younger sibling is revealed through testing. These findings, potentially contradictory at first, are consistent with two theoretical mechanisms: mentoring/role-modeling when older sibling is positively affected by school-entry policies and parental reinforcement in more affluent households when either the younger or the older sibling is positively affected by school-entry policies.

It is important to note several limitations of our study. First, as we acknowledge in the introduction, it is possible that school entry policies could affect other children in the family indirectly through parental decisions that are not related to focal child's human capital (e.g., mother coming back to the labor market a year earlier when her younger child starts school a year ahead because of their August birth date), and these effects could vary by family affluence. If this is the case, it may be more appropriate to attribute our findings to the broader effects of school entry policies on family rather than solely to sibling spillover effects. There is evidence in the Danish context suggesting that school entry policies affect parental decisions such as maternal employment (Landersø et al. 2020), yet this evidence for the U.S. is mixed perhaps due to fundamental differences between the two countries in public assistance programs (Gelbach 2002; Fitzpatrick 2010b). Our exploratory analysis reveals evidence against this possibility in the Florida data. For example, we examine the discontinuity in family poverty at the younger child's school starting cutoff, and find no significant change in this discontinuity around the gap year of the younger child. Further, we find no significant spillover effect from the younger to the older sibling until the younger child reaches third grade and is tested for the first time, providing indirect evidence that parental decisions are likely driven by the effect of the policy on the educational outcomes of the focal kid instead of the effect of the policy during the gap year where we would expect "family effects" to manifest the strongest. That said, unlike in Denmark, we are unable to provide more direct evidence on this issue due to data limitations reflected in only limited proxies for family income.

Second, we examine sibling spillovers in an anonymous district using an RD design. As such, there is concern about the generalizability of our findings to other settings or to students away from the school entry cutoff. We somewhat alleviate these concerns by exploring the effects separately for high- and low-SES families indeed documenting important heterogeneity, but even within these categories, students in our district are likely different than students in other settings. For example, as we mention above, students in our district are more likely to be White compared to the public student population in Florida. A larger share of K-12 students in the district enroll in private schools (14 percent) compared to other districts in Florida (11 percent) or the general K-12 student population in the United States (10 percent). Furthermore, the differences between students enrolled in private schools and those enrolled in public schools (along parental education and household income) are larger in the district compared to the rest of the state (based on our calculations using the American Community Survey). It is also the case that our district has somewhat higher fraction (54 percent vs. 51 percent in other counties in Florida) of families with two or more children, a necessary condition for sibling spillovers to exist. That said, district data, unlike nationally representative surveys (e.g., NLSY), allow us to observe exact birth dates of individual students and use an RD approach, thus greatly improving the internal validity of the estimates.

Although the policy in question here is not easily manipulable, it is reasonable to think that our results could generalize to other educational interventions that are targeted toward individual children. In particular, our findings suggest that the same educational policy could create both positive and negative externalities depending on the characteristics of the affected households, putting the "one-size fits all" policy approach into question. Further, our findings suggest that any cost effectiveness analyses that ignore such spillovers likely over- or underestimate the costs/benefits associated with that particular policy. Thus, policymakers should be cautious when designing and implementing such interventions, since their target population could differ from the full population of affected individuals.

References

Adhvaryu, Achyuta and Anant Nyshadham, "Endowments at Birth and Parents' Investment in Children," *Economic Journal*, 2016, *126* (593), 781–820.

Alsan, Marcella, "The Gendered Spillover Effect of Young Children's Health on Human Capital: Evidence from Turkey," NBER Working Paper 23702, 2017.

Armstrong, Timothy and Michal Kolesár, "Simple and honest confidence intervals in nonparametric regression," *arXiv:1606.01200v5*, 2019.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman, "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes," *American Economic Journal: Applied Economics*, 2019, *11* (3), 338–381.

Barcellos, Silvia, Leonardo Carvalho, and Adriana Lleras-Muney, "Child gender and parental investments in India: Are boys and girls treated differently?," *American Economic Journal: Applied Economics*, 2014, *6* (1), 157–189.

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

Becker, Gary, *A treatise on the family*, Harvard University Press, 1993.

_ and Nigel Tomes, "Child endowments and the quantity and quality of children," *Journal of Political Economy*, 1976, *84* (4), S143–S162.

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

Berry, James, Rebecca Dizon-Ross, and Maulik Jagnani, "Not playing favorites: An experiment on parental preferences for educational investment," *Mimeo*, 2020.

Bertrand, Marianne and Jessica Pan, "The trouble with boys: Social influences and the gender gap in disruptive behavior," *American Economic Journal: Applied Economics*, 2013, *5* (1), 32–64.

Bingley, Paul, Petter Lundborg, and Stepanie Vincent Lyk-Jensen, "Brothers in Arms: Spillovers from a draft lottery," *Journal of Human Resources*, 2020, *forthcoming.*

Black, Sandra and Paul Devereux, "Recent developments in intergenerational mobility," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, pp. 1487–1541.

_ , _ , and Kjell Salvanes, "Too young to leave the nest? The effects of school starting age," *Review of Economics and Statistics*, 2011, *93* (2), 455–467.

—, **Sanni Breining, David Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen**, "Sibling spillovers," NBER Working Paper 23062, 2017.

Breining, Sanni, "The presence of ADHD: Spillovers between siblings," *Economics Letters*, 2014, *124* (3), 469–473.

—, **Joseph Doyle, David Figlio, and Krzysztof Karbownik**, "Birth Order and Delinquency: Evidence from Denmark and Florida," *Journal of Labor Economics*, 2020, *38* (1), 95–142.

—, —, —, —, **and Jeffrey Roth**, "Birth Order and Delinquency: Evidence from Denmark and Florida," *Journal of Labor Economics*, 2020, *38* (1), 95–142.

—, **Meltem Daysal, Marianne Simonsen, and Mircea Trandafir**, "Spillover Effects of Early Life Medical Interventions," IZA Discussion Paper 9086, 2015.

Buhrmester, Duane, "The developmental course of sibling and peer relationships," in Frits Boer and Judith Dunn, eds., *Children's sibling relationships: Developmental and clinical issues*, Lawrence Erlbaum Associates, 1992.

Calonico, Sebastian, Matias Cattaneo, Max Farrell, and Rocio Titiunik, "rdrobust: Software for regression-discontinuity designs," *The Stata Journal*, 2017, *17* (2), 372–404.

Cascio, Elizabeth and Diane Whitmore Schanzenbach, "First in the class? Age and education production function," *Education Finance and Policy*, 2016, *11* (3), 225–250.

Cattaneo, Matias, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *Stata Journal*, 2018, *18* (1), 234–261.

Chiswick, Barry, "Differences in education and earnings across racial and ethnic groups: Tastes, discrimination, and investments in child quality," *Quarterly Journal of Economics*, 1988, *103* (3), 517–597.

Conley, Dalton, "Bringing sibling differences in: Enlarging our understanding of the transmission of advantage in families," in Annette Lareau and Dalton Conley, eds., *Social Class: How does it work?*, Russell Sage Foundation, 2008.

Cook, Philip and Songman Kang, "Birthdays, schooling, and crime: Regression-discontinuity analysis of school performance, delinquency, dropout, and crime initiation," *American Economic Journal: Applied Economics*, 2016, *8* (1), 33–57.

Cunha, Flavio and James Heckman, "The technology of skill formation," *American Economic Review*, 2007, *97* (2), 31–47.

—, —, **and Susanne Schennach**, "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 2010, *78* (3), 883–931.

Currie, Janet and Douglas Almond, "Human capital development before age five," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, Elsevier, 2011, pp. 1315–1486.

Dahl, Gordon and Lance Lochner, "The impact of family income on child achievement: Evidence from the Earned Income Tax Credit," *American Economic Review*, 2012, *102* (5), 1927–1956.

Dee, Thomas and Hans Sievertsen, "The gift of time? School starting age and mental health," *Health Economics*, 2018, *27* (5), 781–802.

Depew, Briggs and Ozkan Eren, "Born on the wrong day? School entry age and juvenile crime," *Journal of Urban Economics*, 2016, *96*, 73–90.

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

_ **and** _ , "Disabled or young? Relative age and special education diagnoses in schools," *Economics of Education Review*, 2010, *29* (5), 857–872.

_ **, David Figlio, Krzysztof Karbownik, and Jeffrey Roth**, "School starting age and cognitive development," *Journal of Policy Analysis and Management*, 2019, *38* (3), 538–578.

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

Elder, Todd, "The importance of relative standards in ADHD diagnoses: Evidence based on child's date of birth," *Journal of Health Economics*, 2010, *29* (5), 641–656.

_ **and Darren Lubotsky**, "Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers," *Journal of Human Resources*, 2009, *44* (3), 641–683.

Evans, William, Melinda Morrill, and Stephen Parente, "Measuring inappropriate medical diagnosis and treatment in survey data: the case of ADHD among school-aged children," *Journal of Health Economics*, 2010, *29* (5), 657–673.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth, "The Effects of Poor Neonatal Health on Children's Cognitive Development," *American Economic Review*, 2014, *104* (12), 3921–3955.

Firpo, Sergio, Nicole Fortin, and Thomas Lemieux, "Unconditional quantile regression," *Econometrica*, 2009, *77* (3), 953–973.

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

— , "Revising our thinking about the relationship between maternal labor supply and preschool," *Journal of Human Resources*, 2010, *47* (3), 583–612.

Fletcher, Jason, Nicole Hair, and Barbara Wolfe, "Am I my Brother's Keeper? Sibling Spillover Effects: The Case of Developmental Disabilities and Externalizing Behavior," NBER Working Paper 18279, 2012.

Frandsen, Brigham, "Party bias in union representation elections: testing the manipulation in the regression discontinuity design when the running variable is discrete," *Advances in Econometrics*, 2017, *38*, 281–315.

Fredriksson, Peter and Björn Öckert, "Life-cycle effects of age at school start," *Economic Journal*, 2014, *124* (579), 977–1004.

Gelbach, Jonah, "Public schooling for young children and maternal labor supply," *American Economic Review*, 2002, *92* (1), 307–322.

Goodman, Joshua, Michael Hurwitz, Christine Mulhern, and Jonathan Smith, "O brother, where start thouu?: The impact of older siblings' college choices on younger sibings' college choices," *NBER WP 26502*, 2020.

Grätz, Michael and Florencia Torche, "Compensation or reinforcement? The stratification of parental responses to children's early ability," *Demography*, 2016, *53* (6), 1883–1904.

Gurantz, Oded, Michael Hurwitz, and Jonathan Smith, "Sibling effects on high school exam taking and performance," *Mimeo*, 2020.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney, "Parental Education and Parental Time with Children," *Journal of Economic Perspectives*, 2008, *22* (3), 23–46.

Haider, Steven and Kathleen McGarry, "Parental investments in college and later cash transfers," *Demography*, 2018, *55* (5), 1705–1725.

Hanushek, Eric and Steven Rivkin, "Generalizations about using value-added measures of teacher quality," *American Economic Review*, 2010, *100* (2), 267–271.

Heissel, Jennifer, "Teen fertility and siblings' outcomes: Evidence of family spillovers using matched samples," *Journal of Human Resources*, 2020, *forthcoming.*

Hurwitz, Michael, Jonathan Smith, and Jessica Howell, "Student age and the collegiate pathway," *Journal of Policy Analysis and Management*, 2015, *34* (1), 59–84.

Joensen, Juanna Schroeter and Helena Skyt Nielsen, "Spillovers in Education Choice," *Journal of Public Economics*, 2018, *157*, 158–183.

Kalil, Ariel, Rebecca Ryan, and Michael Corey, "Diverging Destinies: Maternal Education and the Developmental Gradient in Time with Children," *Demography*, 2012, *49* (4), 1361–1383.

Karbownik, Krzysztof and Michal Myck, "Who gets to look nice and who gets to play? Effects of child gender on household expenditure," *Review of Economics of the Household*, 2017, *15* (3), 925–944.

Kawaguchi, Daiji, "Actual age at school entry, educational outcomes, and earnings," *Journal of the Japanese and International Economies*, 2011, *25* (2), 64–80.

Kolesár, Michal and Christoph Rothe, "Inference in regression discontinuity designs with a discrete running variable," *American Economic Review*, 2018, *108* (8), 2277–2304.

Krueger, Alan and Diane Schanzenbach, "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR," *Economic Journal*, 2001, *111* (468), 1–28.

Landersø, Rasmus, Helena Nielsen Skyt, and Marianne Simonsen, "School starting age and the crime-age profile," *Economic Journal*, 2017, *127* (602), 1096–1118.

_ , _ , and _ , "Effects of school starting age on the family," *Journal of Human Resources*, 2020, *forthcoming.*

Lavy, Victor, Olmo Silva, and Felix Weinhardt, "The good, the bad, and the average: Evidence on ability peer effects in schools," *Journal of Labor Economics*, 2012, *30* (2), 367–414.

Lee, David and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

Leight, Jessica and Elaine Liu, "Maternal education, parental investment and non-cognitive characteristics in rural China," *Economic Development and Cultural Change*, 2020, *forthcoming.*

Lino, Mark, "Expenditures on children by families: 2000 annual report," Technical Report 1528-2000, Center for Nutrition Policy and Promotion 2000.

Manski, Charles F., "Indentification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

_ , "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 2000, *14* (3), 115–136.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

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

Michelmore, Katherine and Susan Dynarski, "The gap within the gap: Using longitudinal data to understand income differences in educational outcomes," *AERA Open*, 2017, *3* (1), 1–18.

Nicoletti, Cheti and Birgitta Rabe, "Sibling spillover effects in school achievement," *Journal of Applied Econometrics*, 2019, *34* (4), 482–501.

Ozier, Owen, "Exploiting Externalities to Estimate the Long-term Effects of Early Childhood Deworming," *American Economic Journal: Applied Economics*, 2018, *10* (3), 235–262.

Parman, John, "Childhood Health and Sibling Outcomes: Nurture Reinforcing Nature During the 1918 Influenza Pandemic," *Explorations in Economic History*, 2015, *58*, 22–43.

Persson, Petra, Xinyao Qiu, and Maya Rossin-Slater, "Family spillover effects of marginal diagnoses: The case of ADHD," *NBER WP 28334*, 2021.

Pitt, Mark, Mark Rosenzweig, and Nazmul Hassan, "Inequality in the intrahousehold distribution of food in low-income countries," *American Economic Review*, 1990, *80* (5), 1139–1156.

Qureshi, Javaeria, "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," *Economic Journal*, 2018, *128* (616), 3285–3319.

_ , "Siblings, Teachers and Spillovers on Academic Achievement," *Journal of Human Resources*, 2018, *53* (1), 272–297.

Sacerdote, Bruce, "Experimental and quasi-experimental analysis of peer effects: two steps forward?," *Annual Reviews Economics*, 2014, *6*, 253–272.

Vinopal, Katie and Seth Gershenson, "Re-Conceptualizing Gaps by Socioeconomic Status in Parental Time with Children," *Social Indicators Research*, 2017, *133* (2), 623–643.

Yi, Junjian, James Heckman, Junsen Zhang, and Gabriella Conti, "Early Health Shocks, Intra-Household Resource Allocation and Child Outcomes," *Economic Journal*, 2015, *125*, F347–F371.

Zang, Emma, Poh Lin Tan, and Philip Cook, "Sibling spillovers: Having an academically successful older sibling may be moore important for children in disadvantaged families," *SSNR*, 2020, (3542306).

Tables

Table 1: Descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All children	Children from 2+ families	Col (2) with grades 3 to 8 outcomes	All possible adjacent sibling pairs		First two births	
				Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older
				Panel A. Demographics			
White	74.0	78.1	78.8	78.8	79.1	79.3	79.6
Ever on free or reduced price lunch	50.2	52.3	53.1	53.5	53.6	53.4	53.4
Female	48.3	48.9	49.5	49.7	48.5	49.7	48.3
Birth year	1986.5	1986.4	1986.2	1987.3	1984.4	1988.6	1985.6
				Panel B: Grades 3 to 8 outcomes			
Test score	-		60.9	60.0	62.1	60.7	62.9
Gifted	-		11.6	10.8	12.8	11.2	13.2
Disabled	-		12.1	13.0	10.6	13.2	10.4
N	263,811	110,872	90,352	34,331	33,290	23,659	23,131

Note: Column 1 includes all children attending public schools in the county between 1990 and 2005 who were born between 1976 and 1996; column 2 restricts the sample from column 1 to families where we observe at least two children, column 3 additionally requires at least one test score observation in grades 3 to 8; column 4 is our primary sample for the analyses where older sibling is the focal child and we investigate outcomes for younger child: older sibling birth cohorts 1973 to 1995 while younger sibling birth cohorts 1976 to 1996; column 5 is our primary sample for the analyses where younger sibling is the focal child and we investigate outcomes for the older child: older sibling birth cohorts 1975 to 1996 while younger sibling birth cohorts 1977 to 2000; columns 6 and 7 present subsamples of the previous two columns for the sample of the first two births in family. Running variable in columns 4 through 7 is restricted to +/- 150 days.

Table 2: Discontinuities in background characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	White	Female	Birth month	Birth year	Age difference	Free or reduced price lunch	
						Ever	Annual
				Panel A: Older-to-younger			
Focal child born after school entry cutoff	-0.968	2.096	-0.062	-0.250	9.214	2.114	0.628
	(1.360)	(1.712)	(0.102)	(0.157)	(17.422)	(1.731)	(1.298)
Mean of Y	78.7	49.3	6.6	1987.3	1021.9	53.6	34.282
Implied imbalance (%)	-1.2	4.2	-0.9	0.0	0.9	3.9	1.8
Observations			14,148				114,628
				Panel B: Younger-to-older			
Focal child born after school entry cutoff	1.364	0.279	0.045	-0.075	-13.203	0.954	-0.228
	(1.165)	(1.233)	(0.099)	(0.155)	(18.784)	(1.632)	(1.484)
Mean of Y	78.8	48.8	6.6	1984.3	1019.2	54.1	34.8
Implied imbalance (%)	1.7	0.6	0.7	0.0	-1.3	1.8	-0.7
Observations			13,854				111,418

Note: Local linear regression with bandwidth of +/- 60 days for reduced-form estimates based on equation 7. Sample in panel A is based on all children for whom we observe two adjacent siblings with at least one test score in grades 3 to 8 for the younger child and where the older sibling is a focal child. Sample in panel B is based on all children for whom we observe two adjacent siblings with at least one test score in grades 3 to 8 for the older child and where the younger sibling is a focal child. Single observation per individual. Outcome variables are: indicator for White student, indicator for female student, birth month, birth year, age difference between siblings, indicator for family ever being on free or reduced price lunch, and indicator for contemporaneous free or reduced price lunch status. Indicators are multiplied by 100. Standard errors clustered at running variable at daily level.

Table 3: Main results: sibling spillovers in test scores

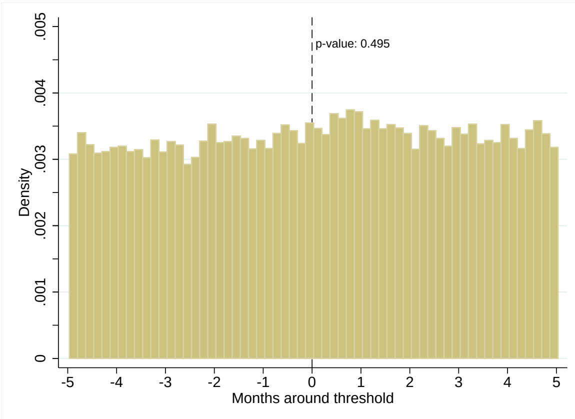
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		All			Ever on free or reduced price lunch			Never on free or reduced price lunch	
					Panel A. Older-to-younger				
Focal child born after school entry	1.347	1.696**	1.694**	3.875***	4.006***	3.943***	-0.387	-0.119	-0.076
	(0.830)	(0.778)	(0.785)	(0.929)	(0.921)	(0.929)	(1.172)	(1.160)	(1.155)
Mean/SD of Y		59.9/26.4			50.6/26.3			70.0/22.6	
Observations		56,701			29,582			27,119	
					Panel B. Younger-to-older				
Focal child born after school entry	-0.561	-0.602	-0.588	0.960	0.692	0.809	-1.983*	-1.755*	-1.773*
	(0.882)	(0.802)	(0.796)	(1.318)	(1.268)	(1.269)	(1.014)	(0.986)	(0.969)
Mean/SD of Y		61.7/26.7			52.3/26.8			72.7/22.2	
Observations		55,813			29,981			25,832	
Own controls		X	X		X	X		X	X
Sibling controls			X			X			X

Note: Local linear regression with bandwidth of +/- 60 days for reduced-form estimates based on equation 7. Sample in panel A investigates spillovers from older (treated) to younger (outcome) children while sample in panel B investigates spillovers from younger (treated) to older (outcome) children. Outcome variable is averaged mathematics and reading scores represented as percentiles of nationally normed distribution. We present both means as well as standard deviations of this variable. Columns 1 through 3 pool all families, columns 4 through 6 are limited to families where any of the children was ever on free or reduced price lunch, while columns 7 through 9 are limited to families where children were never on free or reduced price lunch. Due to grade repetition there are up to eight and up to seven observations per individual in panels A and B, respectively. Mean number of tests are 4.0, 3.9, and 4.1 for columns 1 to 3, 4 to 6 and 7 to 8 in panel A, respectively. Mean number of tests are 4.0, 4.0, and 4.1 for columns 1 to 3, 4 to 6 and 7 to 8 in panel B, respectively. Free or reduced price lunch is measured in school records between 1989/90 and 2004/05. Columns 1, 4 and 7 do not include any controls; columns 2, 5 and 8 control for individual covariates of child for whom we measure the outcome; columns 3, 6 and 9 further include individual controls of the focal child from the pair. Own controls are: indicator for White student, indicator for female student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test. Sibling controls are: indicators for sibling school starting cohorts (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)), indicator for sibling being a female, and age difference between siblings in days. Standard errors clustered at running variable at daily level.

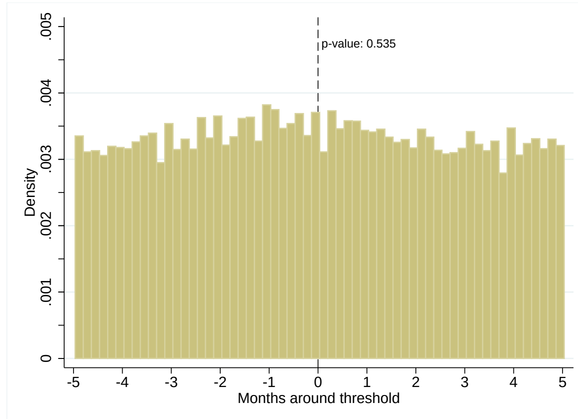
Figures

Figure 1: Distribution of sibling observations around focal child school-entry cutoff

A. Older-to-younger

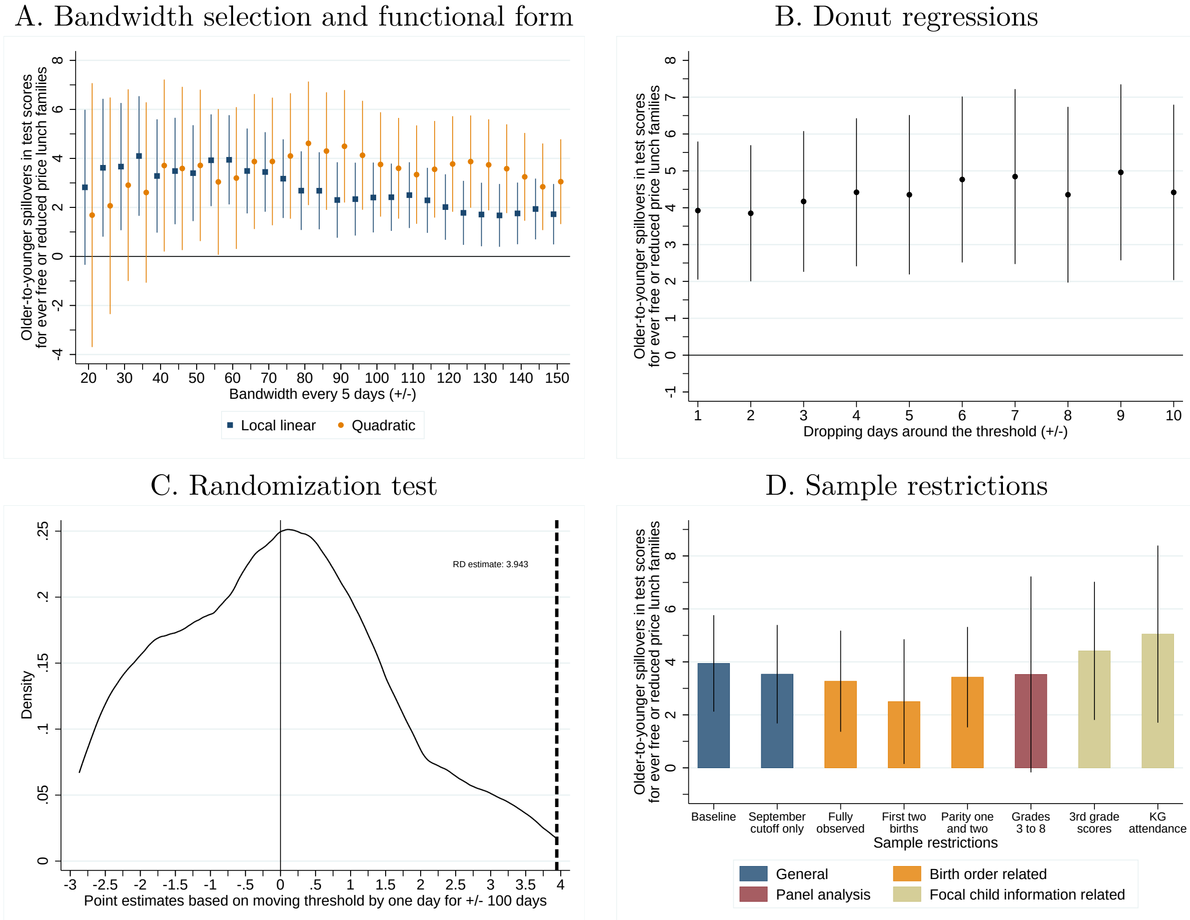


B. Younger-to-older



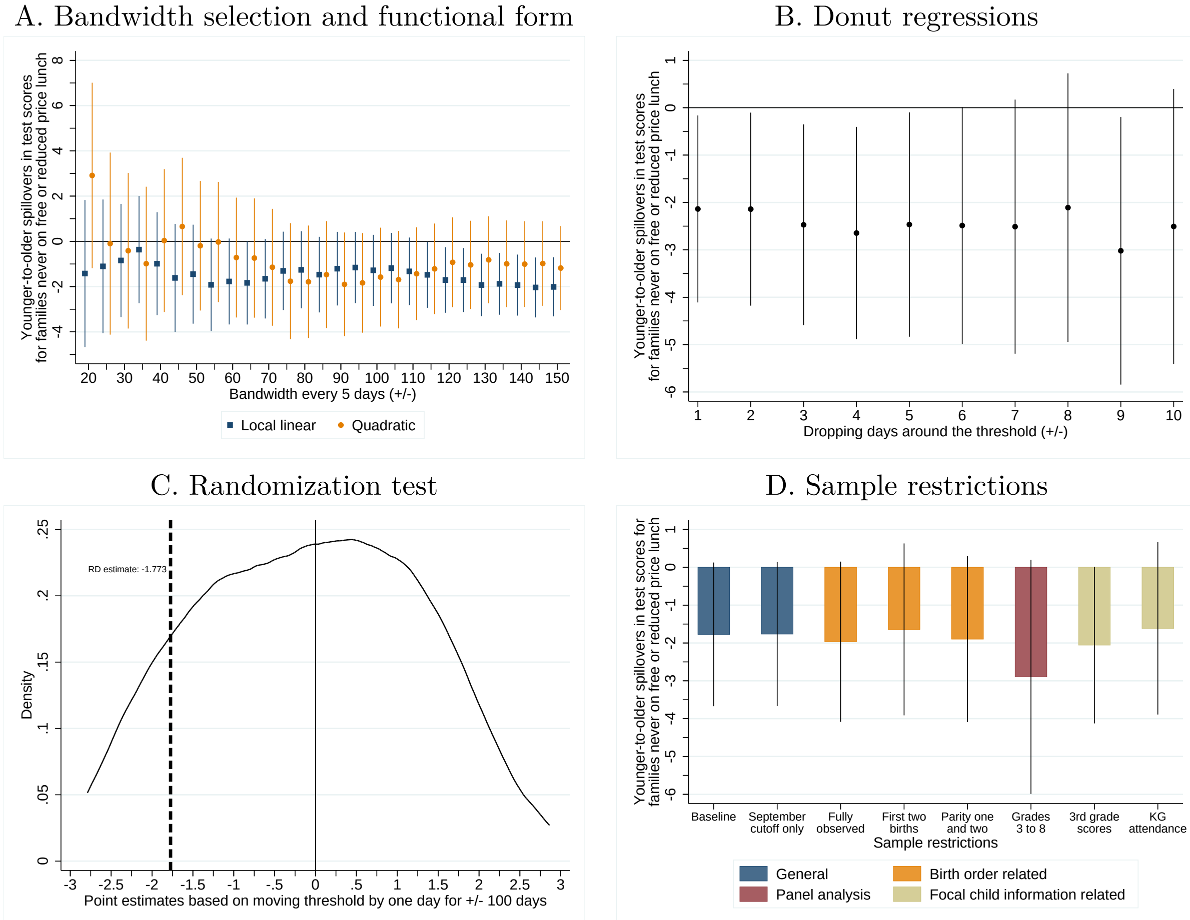
Note: Figures present histogram of density of children born around their sibling's school-entry cutoff (+/- 150 days). Bin width is set to 5 days. Sample in panel A is based on births of younger siblings around the older sibling cutoff while sample in panel B is based on births of older sibling around the younger sibling cutoff. Both samples are based on those used in main analysis in Table 3. Dotted lines indicate cutoff while p-values are based on density test, run at daily level, proposed in Cattaneo et al. (2018).

Figure 2: Robustness of the main results: Older-to-younger spillovers in test scores among families ever on free or reduced priced lunch



Note: This figure presents robustness checks for older-to-younger spillovers in test scores for the sample of families that were ever observed on FRPL. Panel A tests bandwidth and functional form assumptions; panel B presents "donut regressions"; panel C presents randomization inference; and panel D presents estimates in alternative samples. Panel A: dots present point estimates while spikes reflect 95% confidence intervals with standard errors clustered at running variable at daily level; navy squares are based on local linear specification while orange circles are based on quadratic specification; our preferred bandwidth is +/- 60 days and we test bandwidth range from +/- 20 days to +/- 150 days every 5 days. Panel B: point estimates and 95% confidence intervals from local linear regression with +/- 60 days bandwidth when we exclude data points around the threshold; excluded days range from +/- 1 to +/- 10 days. Panel C: density of point estimates coming from 201 regressions where for each regression we move the threshold by one day, either earlier or later; we also include our baseline regression result which is marked with dotted vertical line; each regression uses local linear specification with +/- 60 days bandwidth; the x-axis presents the range of estimates. Panel D: all estimates are based on our preferred specification from panel A, column (6) of Table 3; first bar replicates this estimate; second bar restricts the sample to cohorts where school-entry cutoff is consistently September 1; third bar restricts the sample to cohorts where we can observe full information on birth order from first birth onward; fourth bar limits the sample to first and second borns only; fifth bar limits the sample to parity one and two transitions only; sixth bar limits the main sample to pairs where we observe panel of test scores for the younger child between grades 3 and 8; seventh bar limits the sample to pairs where we observe information on 3rd grade test scores for the focal child; and eight bar limits the sample to pairs where we observe information on kindergarten attendance for the focal child.

Figure 3: Robustness of the main results: Younger-to-older spillovers in test scores among higher SES families



Note: This figure replicates analysis from Figure 2 for the specification from column (9) in panel B of Table 3.

For Online Publication

Appendix Tables

Table A1: School-entry effects on disability for focal children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Older child from older-to-younger pair						Younger child from younger-to-older pair					
	All children		Ever on free or reduced price lunch		Never on free or reduced price lunch		All children		Ever on free or reduced price lunch		Never on free or reduced price lunch	
	Mean	Estimate	Mean	Estimate	Mean	Estimate	Mean	Estimate	Mean	Estimate	Mean	Estimate
Any disability	18.2	-5.045***	23.7	-6.801***	11.8	-3.424**	24.3	-6.371***	29.4	-9.796***	18.3	-2.632
		(1.492)		(2.027)		(1.701)		(1.379)		(1.834)		(1.701)
Behavioral disability	11.1	-4.060***	15.2	-5.434***	6.3	-2.798**	13.7	-3.449***	17.7	-5.274***	8.9	-1.624
		(1.061)		(1.648)		(1.130)		(0.973)		(1.424)		(1.203)
Cognitive disability	1.6	0.223	2.7	0.163	0.5	0.183	2.0	-0.355	3.3	-1.015	0.5	0.251
		(0.360)		(0.608)		(0.361)		(0.444)		(0.846)		(0.355)
Physical disability	7.9	-1.509	9.8	-2.379**	5.8	-0.598	12.0	-3.931***	13.6	-5.513***	10.2	-1.913
		(1.072)		(1.201)		(1.364)		(1.078)		(1.497)		(1.557)
Other disability	0.2	-0.462***	0.2	-0.157	0.2	-0.806***	0.4	-0.166	0.4	-0.174	0.3	-0.143
		(0.174)		(0.290)		(0.258)		(0.158)		(0.263)		(0.259)
Gifted child	13.1	1.164	5.9	0.783	21.4	2.276	10.8	0.587	4.9	1.541*	17.8	-0.154
		(1.098)		(1.162)		(1.724)		(0.883)		(0.917)		(1.463)
N		14,148		7,577		6,571		13,854		7,498		6,356

Note: This table presents estimates of school-entry cutoff on focal children disability diagnoses and gifted status classification. Sample is based on focal children from sibling-pairs used in estimation in Table 3. Disability and gifted status are defined as ever being assigned an appropriate IEP (individual education plan) in grades KG to 8. Behavioral disabilities include emotionally handicapped, specific learning disabled, severely emotionally disturbed, and autistic; cognitive disabilities include educable mentally handicapped, trainable mentally handicapped, language impaired, intellectual disability, profoundly mentally handicapped, and developmentally delayed; physical disabilities include orthopedically impaired, speech impaired, deaf or hard of hearing, visually impaired, hospital/homebound, dual sensory impaired, deaf, and traumatic brain injury; while other disabilities include having established condition and having other health impairment. Columns 1 to 6 present estimates for older focal children (pairs from panel A of Table 3) while columns 7 to 12 present estimates for younger focal children (pairs from panel B of Table 3). Columns 1, 2, 7, and 8 present results for all families; columns 3, 4, 9, and 10 present results for families where any of the children was ever on free or reduced price lunch; while columns 5, 6, 11, and 12 present results for families where children were never on free or reduced price lunch. Even-numbered columns present sample means while odd-numbered columns present point estimates and standard errors. Standard errors clustered at running variable at daily level.

Table A2: Discontinuities in background characteristics. Non-parametric estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
						Free or reduced price lunch	
	White	Female	Birth month	Birth year	Age	Ever	Annual
			Panel A: Older-to-younger				
Focal child born after school entry cutoff	0.040	-1.723	0.059	-0.310	3.596	3.531	0.686
	(2.012)	(2.304)	(0.133)	(0.224)	(24.398)	(2.449)	(1.965)
Mean of Y	79.1	49.7	6.5	1987.4	1025.4	53.5	35.5
Implied imbalance (%)	0.1	-3.5	0.9	0.0	0.4	6.6	1.9
Bandwidth	49	35	51	48	59	49	67
Observations			41,201				165,513
			Panel B: Younger-to-older				
Focal child born after school entry cutoff	-1.248	-0.412	-0.103	-0.263	-23.348	0.902	-1.719
	(1.591)	(1.356)	(0.152)	(0.255)	(23.542)	(2.227)	(1.852)
Mean of Y	79.2	48.6	6.6	1984.4	1022.4	53.6	37.129
Implied imbalance (%)	-1.6	-0.8	-1.6	0.0	-2.3	1.7	-4.6
Bandwidth	31	40	46	50	55	50	59
Observations			40,181				163,189

Note: This table is a non-parametric version of results presented in Table 2 estimated using method developed by Calonico et al. (2017). Standard errors clustered at running variable at daily level.

Table A3: Discontinuities in background characteristics by free and reduced price lunch status

	(1) White	(2) Female	(3) Birth month	(4) Birth year	(5) Age difference	(6) Free or reduced price lunch
	Panel A: Older-to-younger, ever on free or reduced price lunch					
Focal child born after school entry cutoff	-0.791	4.034	-0.224	-0.066	-10.257	-1.651
	(2.149)	(2.534)	(0.146)	(0.183)	(24.574)	(1.401)
Mean of Y	69.2	50.0	6.7	1987.5	980.3	64.1
Implied imbalance (%)	-1.1	8.1	-3.4	0.0	-1.0	-2.6
Observations			7,577			61,302
	Panel B: Older-to-younger, never on free or reduced price lunch					
Focal child born after school entry cutoff	0.145	-0.250	0.118	-0.484**	36.507	
	(1.169)	(2.608)	(0.140)	(0.241)	(26.415)	
Mean of Y	96.1	48.5	6.5	1987.0	1075.2	N/A
Implied imbalance (%)	0.2	-0.5	1.8	0.0	3.4	
Observations			6,571			
	Panel C: Younger-to-older, ever on free or reduced price lunch					
Focal child born after school entry cutoff	2.422	1.463	0.214	0.151	-17.002	-1.559
	(2.086)	(1.888)	(0.154)	(0.208)	(23.210)	(1.484)
Mean of Y	69.8	48.8	6.6	1984.6	973.9	63.9
Implied imbalance (%)	3.5	3.0	3.2	0.0	-1.7	-2.4
Observations			7,498			60,637
	Panel D: Younger-to-older, never on free or reduced price lunch					
Focal child born after school entry cutoff	0.709	-1.163	-0.169	-0.353	-6.517	
	(1.261)	(2.438)	(0.147)	(0.266)	(27.908)	
Mean of Y	96.0	47.7	6.5	1983.9	1073.3	N/A
Implied imbalance (%)	0.7	-2.4	-2.6	0.0	-0.6	
Observations			6,356			

Note: This table replicates balancing checks presented in Table 2 separately for families ever and never on FRPL. Standard errors clustered at running variable at daily level.

Table A4: Effects of older sibling school-entry discontinuity on subsequent fertility

	(1)	(2)	(3)	(4)	(5)	(6)
	Probability of younger sibling birth					
	Parity 1 to 2		Parity 2 to 3		Parity 3 to 4	
Older sibling born after school entry cutoff	1.120	1.155	0.119	0.390	-2.116	-2.219
	(0.696)	(0.702)	(1.033)	(1.025)	(2.465)	(2.554)
Implied % effect	4.8	4.9	0.7	2.3	-15.4	-16.1
Mean of Y	23.4		16.6		13.7	
Observations	65,632		15,431		2,604	
Demographic controls	No	Yes	No	Yes	No	Yes

Note: This table shows the effects of older sibling school-entry cutoff on the probability that there is a younger sibling in the family. Sample is based on children born on or after September 1, 1979 and on or before December 31, 2000, where we can fully observe birth order. All outcomes are indicator variables multiplied by 100. Columns 1 and 2 analyze parity one transition, columns 3 and 4 analyze parity two transition while columns 5 and 6 analyze parity three transition. All regressions include indicators for older sibling school-entry cohort (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)). Demographic controls further include: indicator for gender of older child, indicator for White family and indicator for whether family has ever been on free or reduced price lunch. Standard errors clustered at running variable at daily level.

Table A5: Proxies for socioeconomic status of an older child measured at younger child's school-entry cutoff, before and after the younger child turns 5

	(1) Free or reduced price lunch	(2) Median income	(3) % high school dropouts	(4) % college graduates	(5) % single mothers
	Panel A. Year before younger turns 5				
Focal child born after school entry cutoff	-2.264	981.240	-0.479	0.684	-0.570
	(2.010)	(690.564)	(0.518)	(0.649)	(0.652)
Implied imbalance (%)	-5.6	2.5	-2.8	2.4	-1.9
Mean of Y	40.6	39413	17.4	28.9	29.3
Observations	6,003	4,814			
	Panel B. First year when younger turns 5				
Focal child born after school entry cutoff	-3.091	445.750	-0.070	0.124	-0.101
	(2.120)	(744.573)	(0.504)	(0.688)	(0.689)
Implied imbalance (%)	-7.4	1.1	-0.4	0.4	-0.3
Mean of Y	41.8	39644	17.3	29.0	29.0
Observations	6,003	4,814			

Note: Local linear regression with bandwidth of +/- 60 days for reduced form estimates based on equation 7. Empirical sample based on spillovers from younger (treated) to older (outcome) children where the dependent variables are: an indicator variable for annually measured free or reduced price lunch status (column 1), median micro-neighborhood income (column 2), fraction of high school dropouts in micro-neighborhood (column 3), fraction of college graduates in micro-neighborhood (column 4), and fraction of households headed by single mothers in micro-neighborhood (column 5). Estimates in panel A are based on a school year before the younger child turns five years of age (school-entry cutoff) while estimate in panel B are based on school year in a year when younger child turns five years of age. Indicators and fractions are multiplied by 100. Standard errors clustered at running variable at daily level.

Table A6: Discontinuities in background characteristics: Neighborhood characteristics and annual FRPL

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Median income		% high school dropouts		% college graduates		% single mothers		Annual FRPL	
	Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older
Panel A. All families										
Focal child born after school entry cutoff	-548.734	18.305	0.202	-0.075	-0.361	-0.181	0.049	-0.039	0.544	-0.261
	(425.102)	(415.958)	(0.268)	(0.283)	(0.396)	(0.392)	(0.465)	(0.448)	(1.292)	(1.473)
Mean of Y	40648	40361	16.8	16.9	29.8	29.7	28.1	28.4	34.1	34.6
Implied imbalance (%)	-1.3	0.0	1.2	-0.4	-1.2	-0.6	0.2	-0.1	1.6	-0.8
Observations	106,558	102,132	106,558	102,132	106,558	102,132	106,554	102,132	106,554	102,132
Panel B. Families ever on free or reduced price lunch										
Focal child born after school entry cutoff	48.684	142.691	-0.301	-0.408	0.371	0.265	-0.219	-0.136	-1.834	-1.638
	(423.253)	(373.624)	(0.339)	(0.318)	(0.391)	(0.357)	(0.622)	(0.545)	(1.441)	(1.484)
Mean of Y	35874	35613	20.1	20.2	25.4	25.2	33.4	33.6	63.7	63.7
Implied imbalance (%)	0.1	0.4	-1.5	-2.0	1.5	1.0	-0.7	-0.4	-2.9	-2.6
Observations	56,986	55,545	56,986	55,545	56,986	55,545	56,986	55,545	56,986	55,545
C. Families never on free or reduced price lunch										
Focal child born after school entry cutoff	-720.099	54.413	0.419*	0.184	-0.725	-0.538	-0.232	-0.151	N/A	
	(566.047)	(573.124)	(0.242)	(0.268)	(0.470)	(0.562)	(0.374)	(0.414)		
Mean of Y	46136	46021	13.0	12.9	35.0	35.1	22.0	22.2		
Implied imbalance (%)	-1.6	0.1	3.2	1.4	-2.1	-1.5	-1.1	-0.7		
Observations	49,572	46,587	49,572	46,587	49,572	46,587	49,572	46,587		

Note: Local linear regression with bandwidth of +/- 60 days for reduced-form estimates based on equation 7. Samples in panel A are based on all families; samples in panel B are limited to families where any of the children was ever on free or reduced price lunch; while samples in panel C are limited to families where children were never on free or reduced price lunch. Odd-numbered columns present spillover estimated from older (treated) to younger (outcome) children while even-numbered columns present spillover estimates from younger (treated) to older (outcome) children. Outcomes in columns 1 to 8 are based on time varying characteristics of micro-neighborhoods of residence that are merged into the data at school year level. Dependent variables are median neighborhood income (columns 1 and 2), percentage of high school dropouts in neighborhood (columns 3 and 4), percentage of college graduates in neighborhood (columns 5 and 6), percentage of single mothers in neighborhood (columns 7 and 8). Columns 9 and 10 present results for individual level indicator (varying at annual level) for being on free or reduced price lunch. Columns 9 and 10 replicate analysis in column 7 of Table 2 for the sample where neighborhood information is available. Indicators and percentages are multiplied by 100. Standard errors clustered at running variable at daily level.

Table A7: Younger sibling compliance with school-entry rules by older focal child school-entry cutoff

	(1)	(2)	(3)	(4)	(5)	(6)
	Younger child from older-to-younger pair					
	All children		Ever on free or reduced price lunch		Never on free or reduced price lunch	
	Mean	Estimate	Mean	Estimate	Mean	Estimate
Held back at school entry	3.3	1.010	2.8	0.697	3.8	1.491
		(0.669)		(0.940)		(1.117)
Fast-tracked at school entry	0.6	-0.146	0.7	0.100	0.5	-0.424
		(0.331)		(0.526)		(0.387)
On-time school entry	96.1	-0.865	96.5	-0.796	95.6	-1.067
		(0.769)		(1.131)		(1.174)
N		7,858		4,373		3,485

Note: Local linear regression with bandwidth of +/- 60 days for reduced form estimates based on equation 7. Empirical sample based on spillovers from older (treated) to younger (outcome) children where the dependent variables are: an indicator variable for being held back at school entry (row 1), an indicator variable for being fast-tracked at school entry (row 2), and an indicator variable for on-time school entry (row 3). Odd-numbered columns present means of dependent variables while even-numbered columns present point estimates. Column 1 and 2 pool all families, columns 3 and 4 are limited to families where any of the children was ever on free or reduced price lunch, while columns 5 and 6 are limited to families where children were never on free or reduced price lunch. Indicators are multiplied by 100. Standard errors clustered at running variable at daily level.

Table A8: Main results: Estimates separately for reading and mathematics

	(1)	(2)	(3)	(4)	(5)	(6)
	Older-to-younger			Younger-to-older		
	All	Ever on free or reduced price lunch	Never on free or reduced price lunch	All	Ever on free or reduced price lunch	Never on free or reduced price lunch
	Panel A. Mathematics					
Focal child born after school entry cutoff	1.787**	4.501***	-0.467	-0.222	1.013	-1.110
	(0.888)	(1.059)	(1.251)	(0.839)	(1.325)	(1.012)
Mean/SD of Y	62.6/28.7	53.2/29.0	72.6/24.7	63.4/29.0	54.1/29.4	74.0/24.7
	Panel B: Reading					
Focal child born after school entry cutoff	1.561*	3.413***	0.267	-1.116	0.507	-2.511**
	(0.826)	(1.100)	(1.193)	(0.845)	(1.329)	(1.061)
Mean/SD of Y	57.5/28.3	48.4/28.0	67.4/25.2	60.6/28.6	51.1/28.6	71.4/24.4
Observations	56,116	29,142	26,974	55,203	29,491	25,712

Note: Local linear regressions with bandwidth of +/- 60 days for reduced-form estimates based on equation 7. Each panel presents point estimates, standard errors as well as means and SDs for mathematics test scores (panel A) and reading test scores (panel B). Columns 1 through 3 present spillovers from older (treated) to younger (outcome) children while columns 4 through 6 present spillovers from younger (treated) to older (outcome) children. Columns 1 and 4 pool all families, columns 2 and 5 are limited to families where any of the children was ever on free or reduced price lunch, while columns 3 and 6 are limited to families where children were never on free or reduced price lunch. Free or reduced price lunch is measured in school records between 1989/90 and 2004/05. All regressions include both own and sibling controls from Table 3. Standard errors clustered at running variable at daily level.

Table A9: Effects of younger sibling school-entry discontinuity on older child achievement: Estimates by timing of testing

	(1) Main effect	(2) Before first testing	(3) After first testing	(4) p-value difference (2) vs (3)
		Panel A. All		
Younger child born after school entry cutoff	-0.588	0.015	-1.827*	0.071
	(0.796)	(0.884)	(1.027)	
Mean/SD of Y	61.7/26.7	62.4/26.3	60.8/27.3	
Observations	55,813	33,454	22,359	
		Panel B. Ever on free or reduced price lunch		
Younger child born after school entry cutoff	0.809	1.362	-0.036	0.254
	(1.269)	(1.402)	(1.401)	
Mean/SD of Y	52.3/26.8	53.0/26.5	51.4/27.0	
Observations	29,981	17,452	12,529	
		Panel C. Never on free or reduced price lunch		
Younger child born after school entry cutoff	-1.773*	-1.097	-3.300***	0.099
	(0.969)	(1.129)	(1.231)	
Mean/SD of Y	72.7/22.2	72.6/22.0	72.8/22.6	
Observations	25,832	16,002	9,830	

Note: Local linear regression with bandwidth of +/- 60 days for reduced-form estimates based on equation 7 and specifications with controls from columns 3, 6 and 9 of panel B of Table 3. Control variables include: indicator for White student, indicator for female student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)), indicator for sibling being a female, and age difference between siblings in days. Sample in panel A presents estimates for full population, in panel B for families with a child ever on free or reduced price lunch, and in panel C for families with children never on free or reduced price lunch. Column 1 replicates the main effects from Table 3 while columns 2 and 3 divide the sample by timing of the testing of the younger child: column 2 presents spillovers into older child for scores before younger child is first tested (grade 3 or before) and column 3 presents spillovers into older child for scores after younger child is first tested (grades 4 and after). Column 4 presents p-value from statistical test that the estimates in columns 2 and 3 are equal. Outcome variable is averaged mathematics and reading scores represented as percentiles of nationally normed distribution. We present both means as well as standard deviations of this variable. Standard errors clustered at running variable at daily level.

Table A10: Main results: Sibling spillovers in test scores. Non-parametric estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		All			Ever on free or reduced price lunch			Never on free or reduced price lunch	
					Panel A. Older-to-younger				
Focal child born after school entry cutoff	1.627	1.986*	2.050*	3.920***	3.824***	3.971***	0.275	0.929	0.818
	(1.168)	(1.164)	(1.129)	(1.356)	(1.331)	(1.312)	(1.362)	(1.429)	(1.340)
Mean/SD of Y		60.1/26.4			50.8/26.2			70.2/22.6	
Bandwidth	54	53	54	48	52	52	72	61	67
Observations		165,513			86,202			79,311	
					Panel B. Younger-to-older				
Focal child born after school entry cutoff	-1.158	-1.007	-0.973	0.061	0.018	0.084	-2.387*	-1.702	-1.719
	(1.156)	(1.139)	(1.128)	(1.756)	(1.743)	(1.755)	(1.252)	(1.230)	(1.235)
Mean/SD of Y		62.1/26.7			52.4/26.6			73.1/22.0	
Bandwidth	56	48	47	64	55	54	50	49	48
Observations		163,189			86,504			76,685	
Own controls		X	X		X	X		X	X
Sibling controls			X			X			X

Note: This table is a non-parametric version of results presented in Table 3 estimated using method developed by Calonico et al. (2017). Standard errors clustered at running variable at daily level.

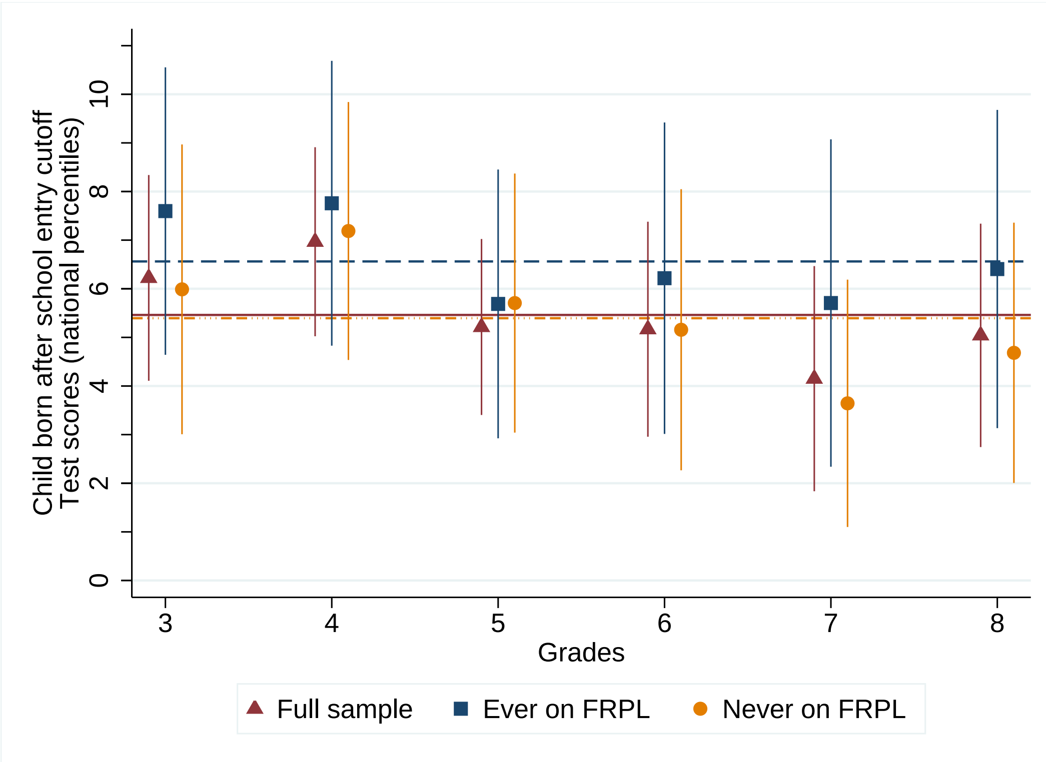
Table A11: Main results with 95% confidence intervals: Testing various standard errors assumptions

		(1)	(2)	(3)	(4)	(5)	(6)
		All		Ever on free or reduced price lunch		Never on free or reduced price lunch	
		Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older	Older-to-younger	Younger-to-older
Focal child born after school entry cutoff	OLS	1.347	-0.561	3.875	0.960	-0.387	-1.983
	robust	(0.479; 2.216)	(-1.449; 0.326)	(2.691; 5.059)	(-0.247; 2.167)	(-1.469; 0.695)	(-3.070; -0.896)
	cluster RV	(-0.296; 2.990)	(-2.308; 1.185)	(2.036; 5.713)	(-1.649; 3.569)	(-2.706; 1.933)	(-3.990; 0.024)
	cluster family	(-0.387; 3.082)	(-1.449; 0.326)	(2.691; 5.059)	(-0.247; 2.167)	(-1.469; 0.695)	(-3.070; -0.896)
	two-way	(-0.278; 2.973)	(-2.286; 1.163)	(2.065; 5.685)	(-1.625; 3.545)	(-2.696; 1.922)	(-3.964; -0.001)
Focal child born after school entry cutoff	RDrobust	1.627	-1.158	3.920	0.061	0.275	-2.387
	nn	(0.441; 2.813)	(-2.304; -0.012)	(2.217; 5.632)	(-1.448; 1.570)	(-1.005; 1.555)	(-3.904; -0.871)
	cluster	(-0.663; 3.917)	(-3.424; 1.108)	(1.262; 6.577)	(-3.380; 3.502)	(-2.395; 2.945)	(-4.841; 0.067)
Focal child born after school entry cutoff	RDHonest	1.079	-1.610	3.660	0.690	-0.206	-1.865
	uniform	(-0.272; 2.431)	(-3.105; -0.114)	(2.157; 5.162)	(-0.826; 2.205)	(-1.610; 1.198)	(-3.498; -0.232)
	optimal M	0.002	0.003	0.001	0.001	0.001	0.002

Note: This table tests various standard errors assumptions for estimates from Tables 3 and A10 without additional controls and provides 95% confidence intervals. Row (1) presents nonparametric local linear estimates while rows (2) to (5) provide confidence intervals under various assumptions on standard errors estimation. Row (2) uses Eicker-Huber-White heteroskedasticity robust standard errors; row (3) uses standard errors clustered by the running variable; row (4) uses standard errors clustered at the family level; and row (5) uses standard errors clustered two-way by running variable and at the family level. Row (6) presents estimates using the optimal bandwidth selection procedure proposed by Calonico et al. (2017) while rows (7) and (8) provide two sets of confidence intervals for these estimates. First (row (7)) uses standard errors based on heteroskedasticity-robust nearest neighbor variance estimator while second (row (8)) uses standard errors based on cluster-robust by the running variable variance estimator. Subsequent three rows present estimates, honest confidence intervals and optimal M based on procedure developed by Kolesár and Rothe (2018). These computations are based on RDHonest command with uniform kernel, FLCI optimal criterion and Hölder smoothness class to replicate panel B of Table 2 in Kolesár and Rothe (2018). Optimal M, smoothness constant, is computed for each sample separately using data-driven procedure from Armstrong and Kolesár (2019). Odd numbered columns present estimates for spillovers from older (treated) to younger (outcome) children while even numbered columns present estimates for spillovers from younger (treated) to older (outcome) children. Columns 1 and 2 pool all families, columns 3 and 4 are limited to families where any of the children was ever on free or reduced price lunch, while columns 5 and 6 are limited to families where children were never on free or reduced price lunch.

Appendix Figures

Figure A1: Persistance of school-entry cutoff advantage: Focal children effects over grades

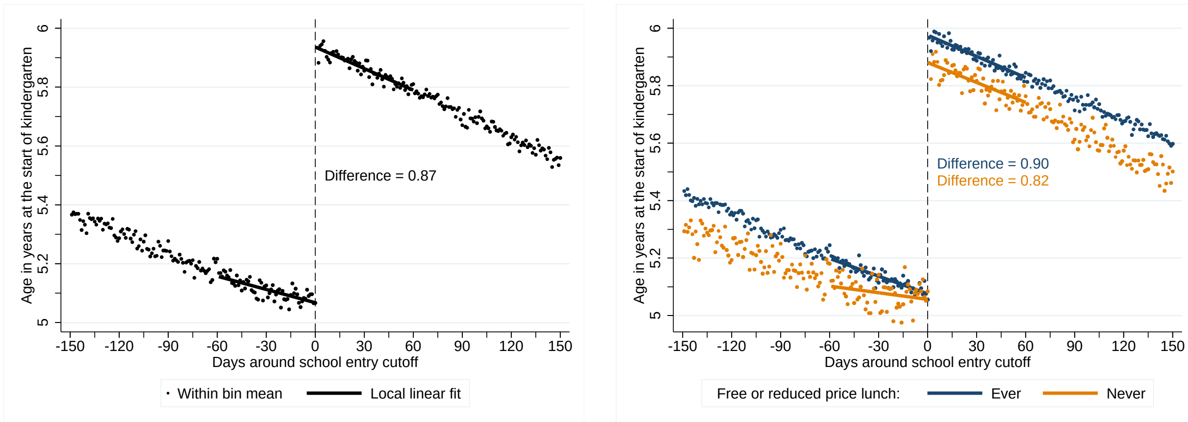


Note: This figure presents estimates of school-entry cutoff rule on focal children test scores (averaged mathematics and reading) over grades. Sample is based on focal children from sibling-pairs used in estimation in Table 3. Each marker represents point estimate from a separate regression with 95% confidence interval. Standard errors clustered at running variable at daily level. Maroon triangles depict estimates pooled across all families; navy squares depict estimates for families where any of the children was ever on free or reduced price lunch; while orange circles depict estimates for families where children were never on free or reduced price lunch. Lines present estimates pooled across all grades (3 to 8): maroon solid for all families; navy dashed for families where any of the children was ever on free or reduced price lunch; and orange dash-dotted for families where children were never on free or reduced price lunch.

Figure A2: Discontinuity in age at school-entry

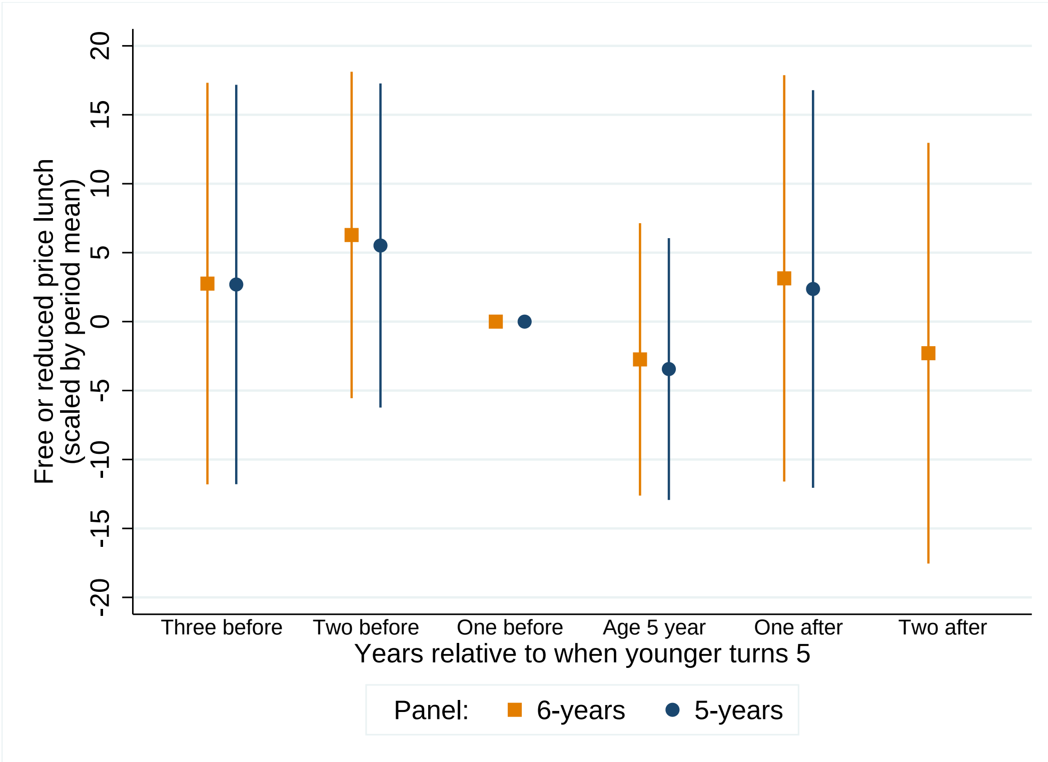
I. Full Sample

II. Breakdown by SES



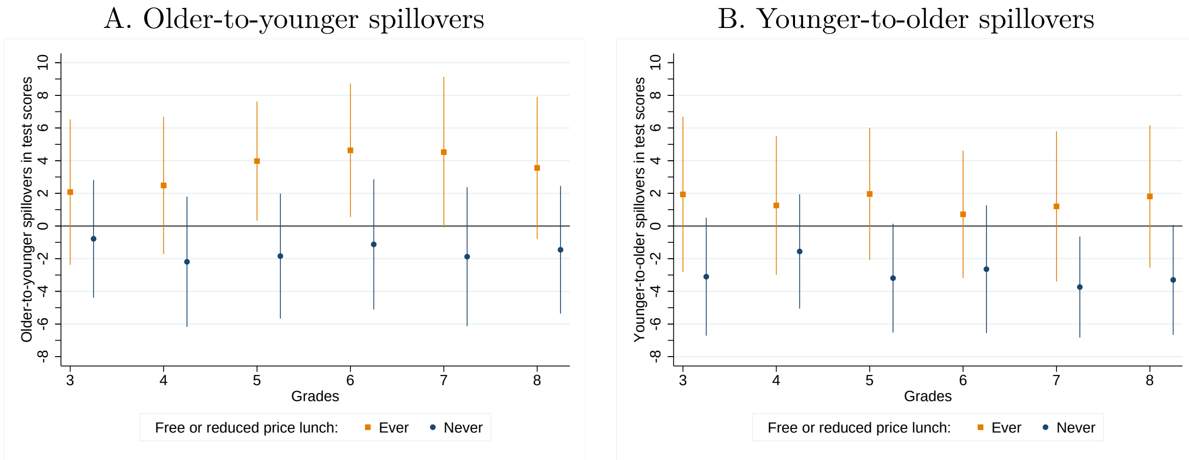
Note: Sample is based on all children first observed in kindergarten. Each figure presents daily means for +/- 150 days (dots) and linear fits for preferred bandwidth of +/- 60 days (lines) on either side of the school-entry cutoff. Panel A pools all families while panel B divides sample by free or reduced price lunch status of the family. Navy dots represent children from families ever on free or reduced price lunch while orange dots represent children from families never on free or reduced price lunch.

Figure A3: Free or reduced price lunch status of an older child as a function of younger child's school entry age



Note: This figure presents event-study estimates for annually recored indicator for free or reduced price lunch of an older sibling as a function of school-entry cutoff of their younger brother or sister. Each series represents a separate panel requirement for either six (orange) or five (navy) consecutive observations of older child's free or reduced price lunch status around younger child's 5th birthday (school-entry cutoff), and comes from a separate regression. Markers depict coefficients on interaction terms between an indicator variable of younger child's birth after September 1st and an indicator variable for relative timing to younger child's 5th birthday (in school years). Omitted variable is one year before age 5 when a child becomes eligible to start school. Spikes depict 95% confidence intervals with standard errors clustered at running variable at daily level. No additional demographic controls are included in regressions.

Figure A4: Sibling spillovers over grades

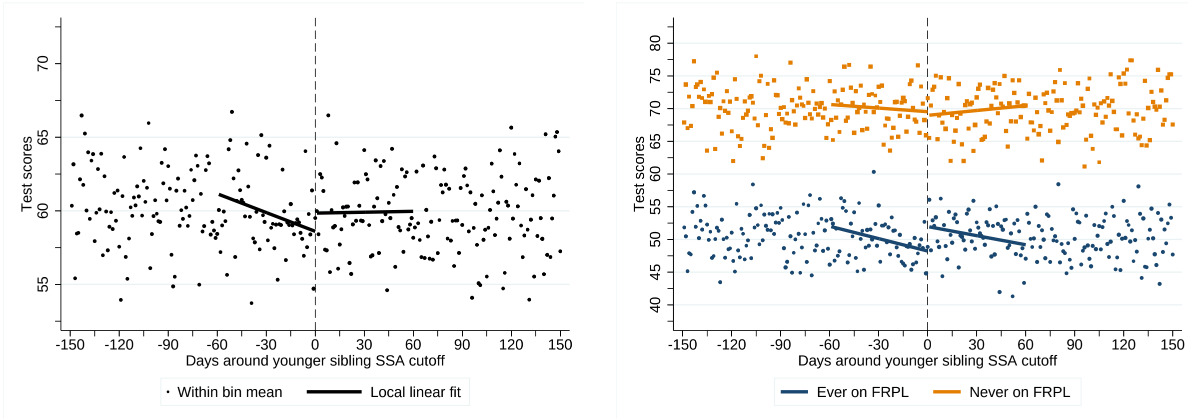


Note: Each graph presents spillover effects, older-to-younger in panel A and younger-to-older in panel B, over grades based on repeated cross-section and specifications from columns 6 and 9 in Table 3. Markers represent point estimates while spikes reflect 95% confidence intervals with standard errors clustered at running variable at daily level. Orange series presents estimates for families where any of the children was ever on free or reduced price lunch while navy series presents estimates for families where children were never on free or reduced price lunch.

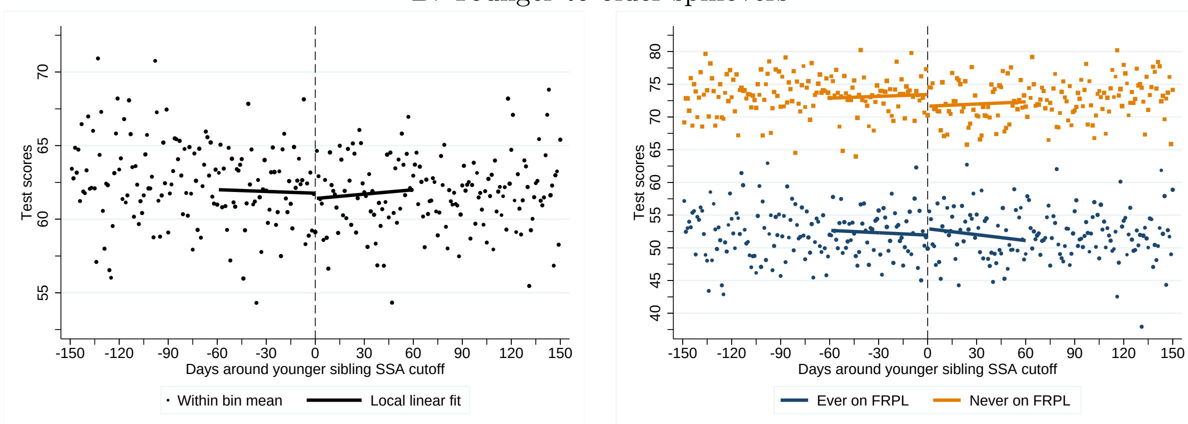
Figure A5: Discontinuity in test scores

I. Full Sample | II. Breakdown by SES

A. Older-to-younger spillovers

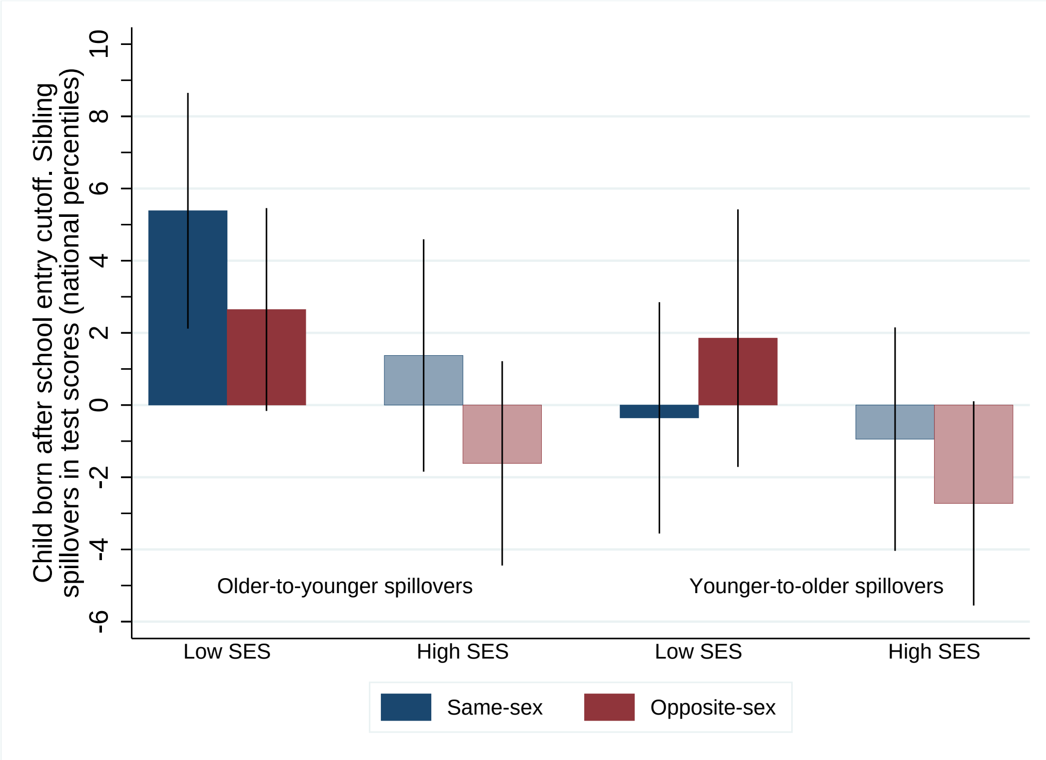


B. Younger-to-older spillovers



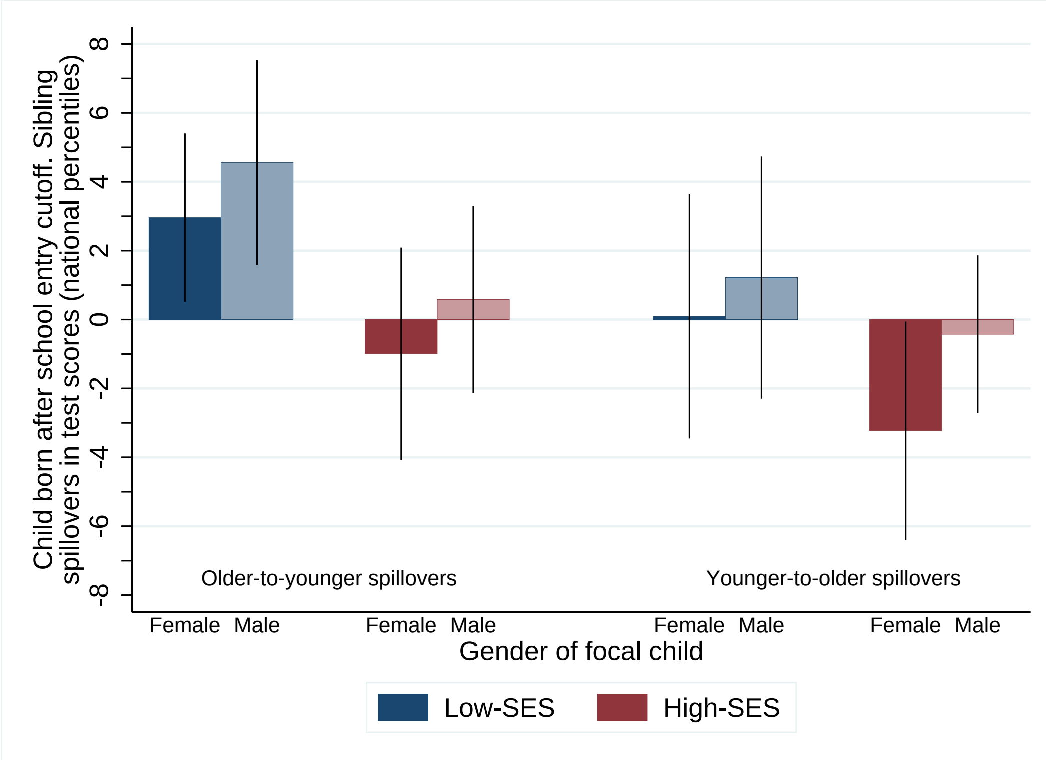
Note: Main results in graphical form with each figure presenting daily means for +/- 150 days (dots) and linear fits for preferred bandwidth of +/- 60 days (lines) on either side of the school-entry cutoff for older child in the sibling pair (panel A) and younger child in the sibling pair (panel B), respectively. Column (I) presents results for all children while column (II) separates the sample into those whose families ever experienced free or reduced price lunch (navy) and never experienced free or reduced price lunch (orange). Samples are based on those used in Table 3.

Figure A6: Heterogeneity analysis by gender composition of sibling pair



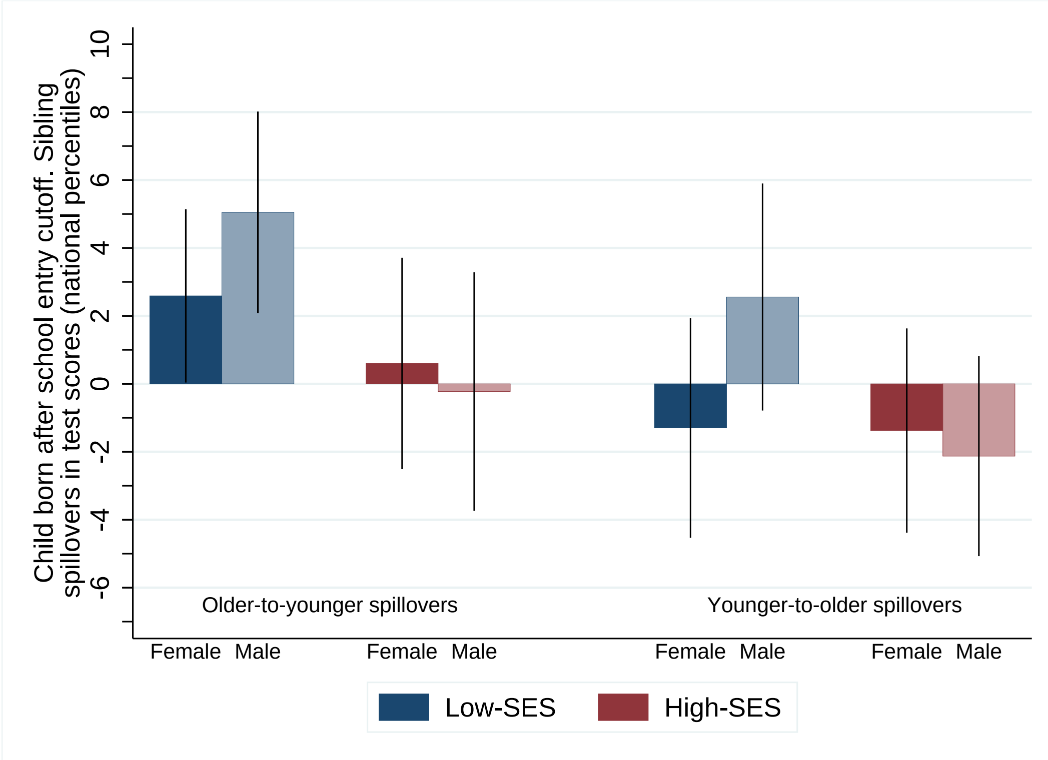
Note: Analysis is based on the samples used in columns 4 through 9 in Table 3; local linear regressions based on +/- 60 days bandwidth; control variables include: indicator for White student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)), and age difference between siblings in days. Each bar represents point estimate from a separate regression with 95% confidence interval. Standard errors clustered at running variable at daily level. Navy bars present estimates for same-sex sibling dyads while maroon bars present estimates for opposite-sex dyads; darker bars present estimates for families ever on free or reduced price lunch while lighter bars present estimates for families never on free or reduced price lunch.

Figure A7: Heterogeneity analysis by gender of the focal child



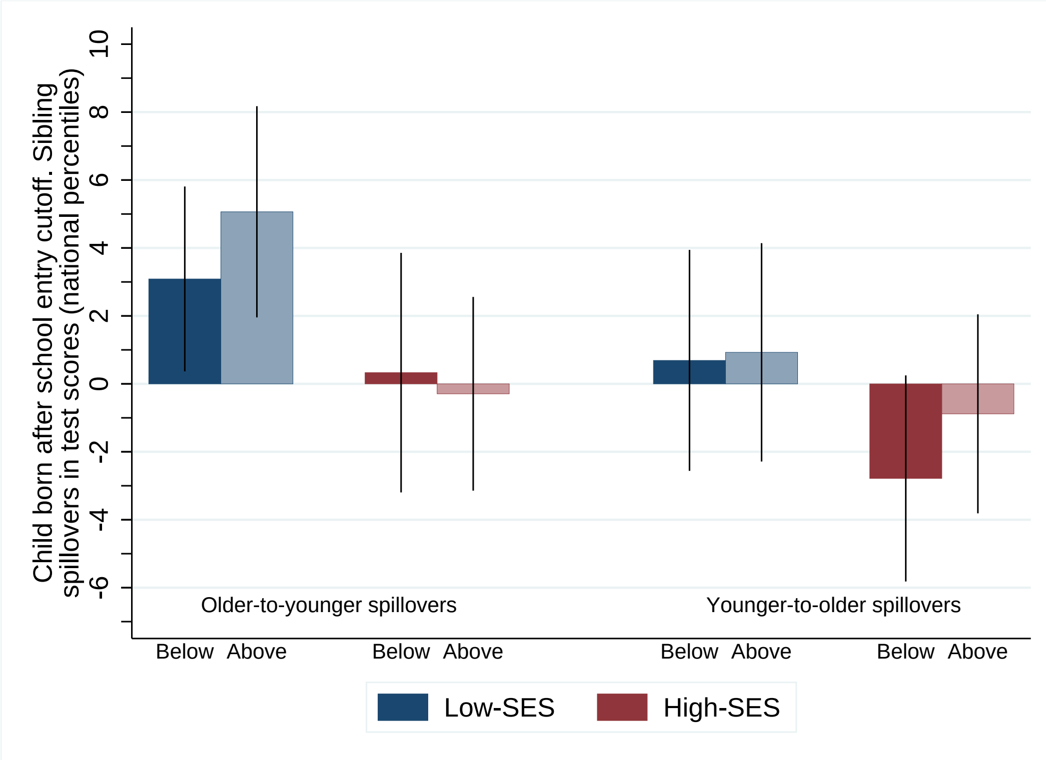
Note: Analysis is based on the samples used in columns 4 through 9 in Table 3; local linear regressions based on +/- 60 days bandwidth; control variables include: indicator for White student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)), and age difference between siblings in days. Each bar represents point estimate from a separate regression with 95% confidence interval. Standard errors clustered at running variable at daily level. Navy bars present estimates for families ever on free or reduced price lunch while maroon bars present estimates for families never on free or reduced price lunch; darker bars present estimates for female focal children while lighter bars present estimates for male focal children.

Figure A8: Heterogeneity analysis by gender of the sibling



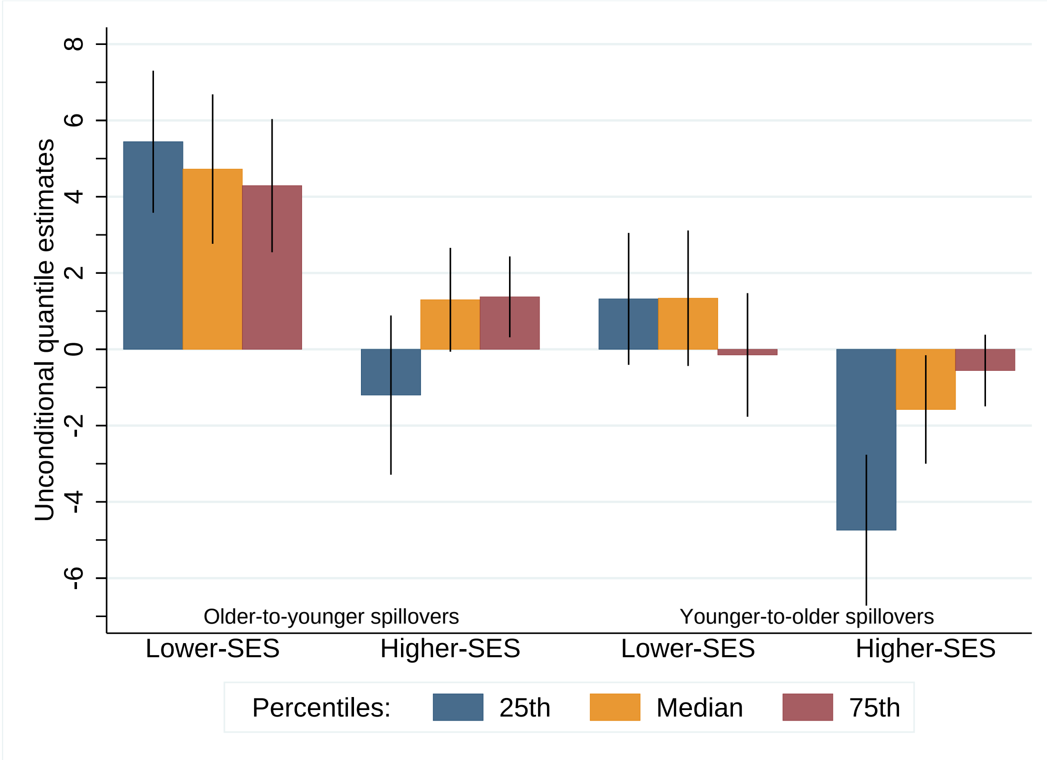
Note: Analysis is based on the samples used in columns 4 through 9 in Table 3; local linear regressions based on +/- 60 days bandwidth; control variables include: indicator for White student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts (cutoff-cohort level with a single indicator for both children born before and after cutoff in a given cohort (e.g., 1992)), and age difference between siblings in days. Each bar represents point estimate from a separate regression with 95% confidence interval. Standard errors clustered at running variable at daily level. Navy bars present estimates for families ever on free or reduced price lunch while maroon bars present estimates for families never on free or reduced price lunch; darker bars present estimates for female siblings while lighter bars present estimates for male siblings.

Figure A9: Heterogeneity analysis by median spacing between siblings



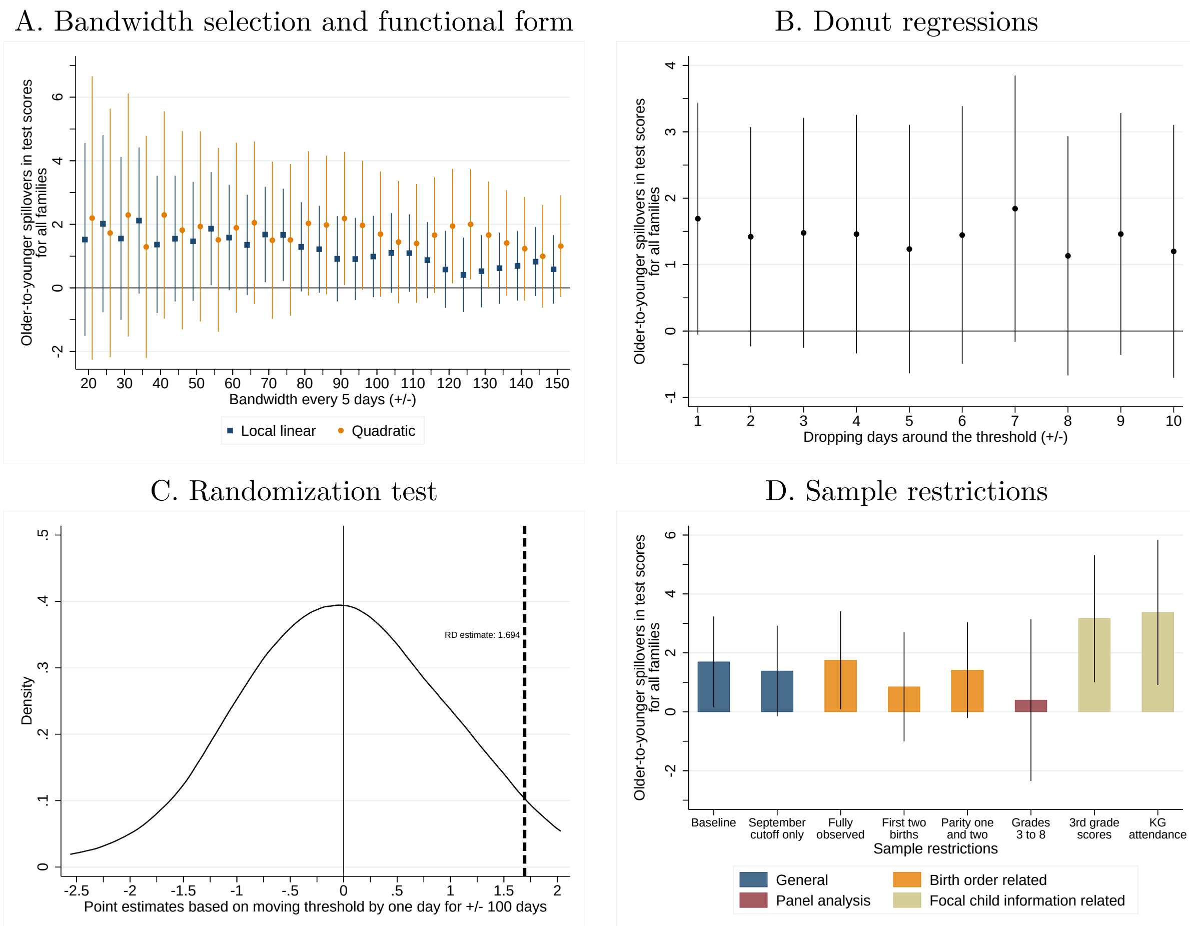
Note: Analysis is based on the samples used in columns 4 through 9 in Table 3; local linear regressions based on +/- 60 days bandwidth; control variables include: indicator for White student, indicator for female student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, and indicator for sibling being a female. Each bar represents point estimate from a separate regression with 95% confidence interval. Standard errors clustered at running variable at daily level. Navy bars present estimates for families ever on free or reduced price lunch while maroon bars present estimates for families never on free or reduced price lunch; darker bars represent spacing below medial and lighter bars represent spacing above median.

Figure A10: Unconditional quantile regression



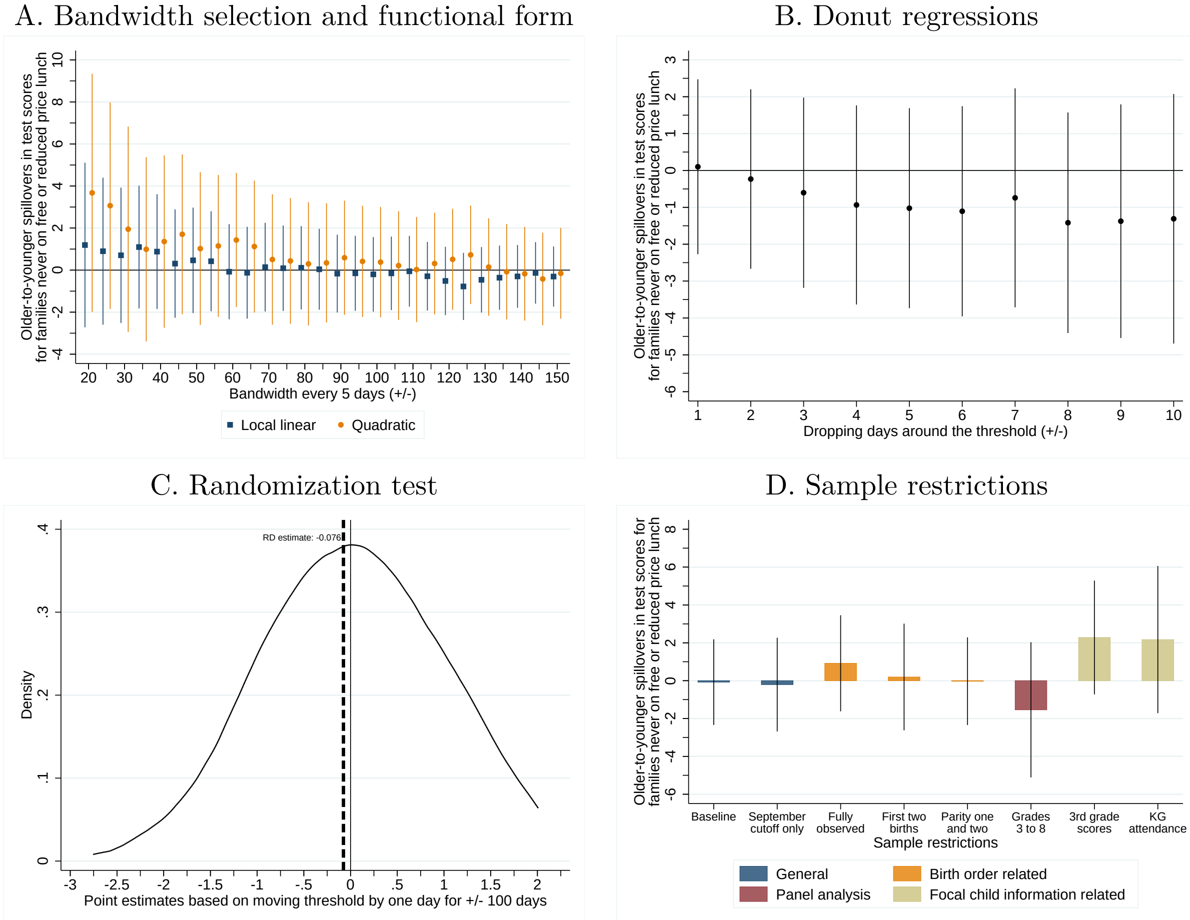
Note: Unconditional quantile regressions (Firpo et al. 2009). Navy bars present estimates for the 25th percentile, orange bars present estimates for the median, while maroon bars present estimates for the 75th percentile of test score distribution. Older-to-younger spillovers in the two left hand side sets of bars while younger-to-older spillovers in the two right hand side sets of bars. Families that were ever on free or reduced price lunch in first and third set of bars while families that never experienced free or reduced price lunch status in second and fourth set of bars. Bootstrapped standard errors with 200 replications.

Figure A11: Robustness of the main results: Older-to-younger spillovers in test scores among all families



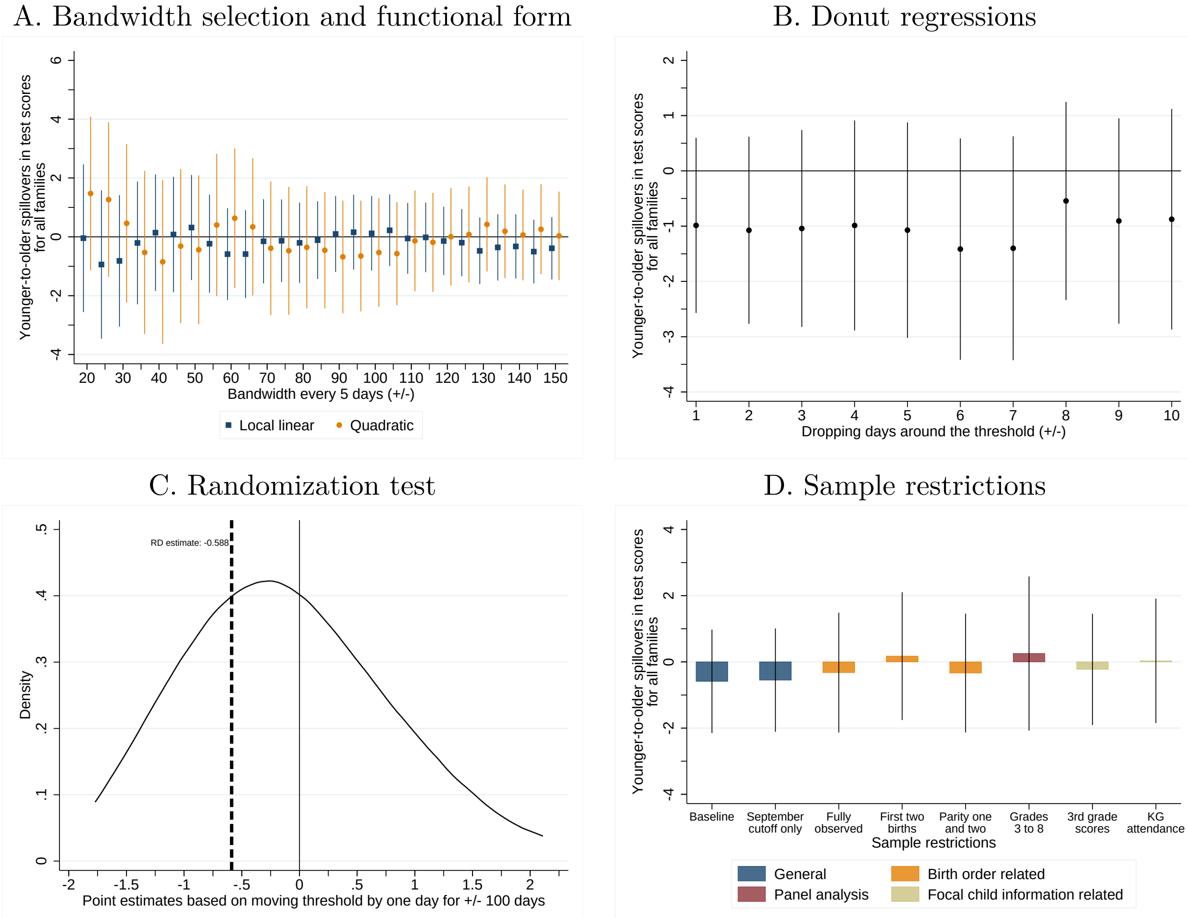
Note: This figure replicates analysis from Figure 2 for the specification from column (3) in panel A of Table 3.

Figure A12: Robustness of the main results: Older-to-younger spillovers in test scores among high SES families



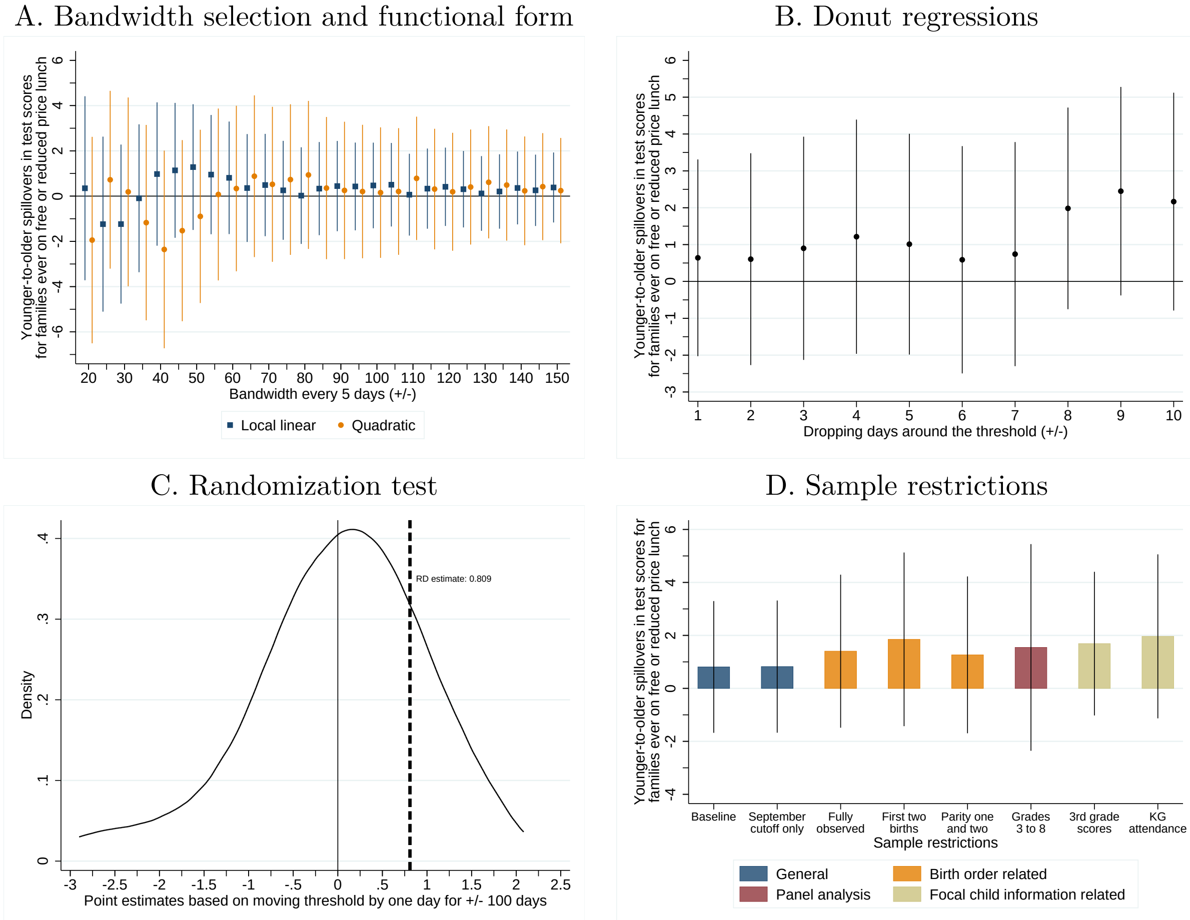
Note: This figure replicates analysis from Figure 2 for the specification from column (9) in panel A of Table 3.

Figure A13: Robustness of the main results: Younger-to-older spillovers in test scores among all families



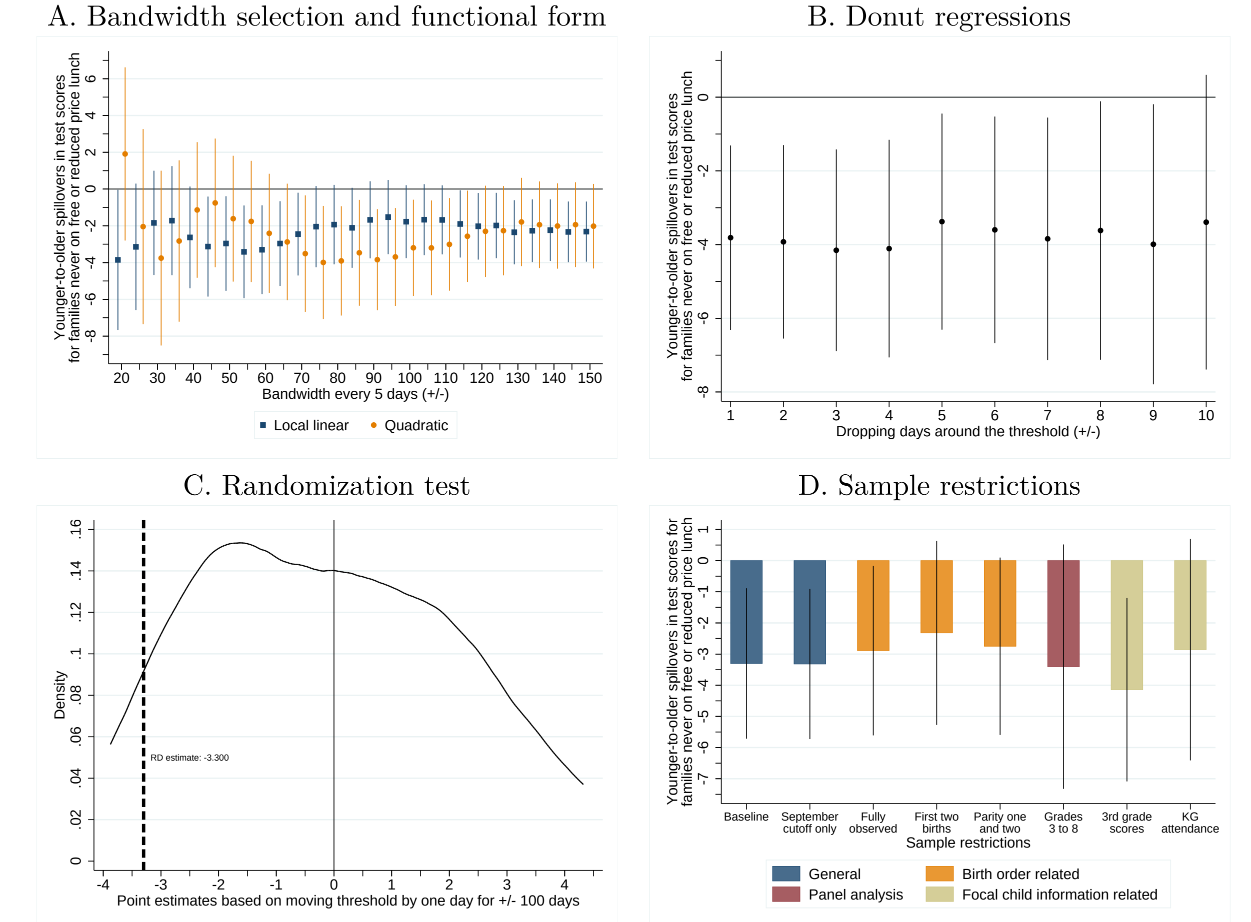
Note: This figure replicates analysis from Figure 2 for the specification from column (3) in panel B of Table 3.

Figure A14: Robustness of the main results: Younger-to-older spillovers in test scores among lower SES families



Note: This figure replicates analysis from Figure 2 for the specification from column (6) in panel B of Table 3.

Figure A15: Younger-to-older spillovers in test scores among higher SES families: Sample of older children after the commencement of testing of the younger focal child



Note: This figure replicates analysis from Figure 2 for the specification from column (9) in panel B of Table 3. Sample is restricted to observations column (3) in Panel C of Table A9.