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

February 19, 2019

Krzysztof Karbownik
Northwestern University

Umut Özek
American Institutes for Research

Abstract

We identify externalities in human capital production function arising from sibling spillovers. Using regression discontinuity design generated by school-entry cutoffs and school records from one district in Florida, 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 are consistent with direct spillovers dominating in economically disadvantaged families and with parental reinforcement in more affluent families.

*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, 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 2019). 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.

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 (Bedard and Dhuey 2006; Dhuey et al. 2019), and here we confirm these estimates using school records from a large, anonymous district in the state of Florida. 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.[1] Our data

[1]In this study, we are not concerned about the exact source of the developmental advantage of the focal child — child's age at school-entry, effect of their age at the time of outcome measurements, or the effect of their age relative to their peer group in school (Cascio and Whitmore Schanzenbach 2016). While we acknowledge that all these channels could be at play, our primary goal is to estimate reduced-form spillover effects across siblings conditional on the focal

further enable us to link all children to their place of residence, providing information on household characteristics, sibship composition, and educational outcomes of other children in the family. Using these family identifiers, we find Pearson correlations among two adjacent siblings of 0.5, 0.4, and 0.3 for test scores, gifted program enrollment, and disability diagnoses, respectively. Similarly, regressions of student test scores in elementary and middle school on their older siblings' test scores (or their younger siblings' test scores) yield coefficients of around 0.40 to 0.50 (p-value<0.001). Our main objective in this paper is to examine the extent to which these correlations reflect causality 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 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 compared to students whose older siblings were born right before the cutoff. We 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 are particularly pronounced when siblings are spaced closer together or when the older child is struggling academically. No support for younger-to-older spillovers is found in poorer households. Our findings are thus consistent with direct spillovers from older-to-younger children in less affluent households and suggestive of reinforcing behavior in more affluent households, especially when the older child performs poorly in school.[2]

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 (2014), Qureshi (2017), and Joensen and Nielsen (2018), all of whom examine sibling spillovers generated by educational shocks or policies in developed countries.[3] Using data from England and North Carolina, respectively, Nicoletti and Rabe (2014) and Qureshi (2017) employ fixed-effects strategies to estimate spillover effects of having a higher performing sibling in the household. In the former paper, the focal sibling's human capital is affected through his/her classroom peers, while the latter study examines

sibling's higher human capital irrespective of its exact cause.

[2]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.

[3]There is more extensive literature on spillovers generated by health shocks. Several 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. 2016). 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 2017), experiencing a disease before age 3 in China (Yi et al. 2015), or influenza pandemics in the U.S. (Parman 2015). Beyond studying a different input into human capital production function, in developed countries context only Breining et al. (2016) use an RD design, while the stronger identification-wise experimental papers are all based on data from developing countries. In settings outside of education or health, Bingley et al. (2017) study sibling spillovers from military conscription in Denmark, and Heissel (2018) studies sibling spillovers from teenage pregnancies in Florida.

the spillover effects of exposure to an experienced teacher. Both papers find evidence for positive spillovers and conclude that these operate primarily through direct exposure and interaction with a higher ability sibling rather than through parental resource reallocation.[4] In both cases, causality is established through saturated fixed-effects models that necessarily require stronger and potentially more problematic assumptions than natural experiments, such as the one explored in this paper. Furthermore, both studies rely on relatively small, and plausibly temporary effects on the focal sibling to study spillovers, and examine contemporaneous or short-run spillover effects. We complement these studies by (1) using a policy that has a large effect on focal siblings' achievement that persists through their schooling career, (2) examining spillover effects in the short and medium term; and (3) employing a more robust research 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. The study finds 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).

Landerso et al. (2019) is perhaps the closest in spirit to our study in that they utilize school starting age cutoff as a quasi-exogenous shock to study within-family spillovers using Danish data. The study shows that a child's advantage in school affects both parents and older siblings, and relates these findings to the reallocation of resources within the family. Older siblings of treated children, however, have improved test scores in ninth grade only if the spacing between children is large enough to affect the older ones exactly at the time around their ninth grade exit exam. Consistent with these results, we do not find any support for positive younger-to-older spillovers when the two siblings are closer in age, and in fact, in Florida the reverse appears to be true. Our study further diverges from Landerso et al. (2019) in that it provides a more comprehensive analysis of sibling spillover effects along two important dimensions. First, we are able to examine the older-to-younger sibling spillovers in an RD setting that is not feasible in the Danish context due to data limitations. Furthermore, while Landerso et al. (2019) can examine sibling spillover effects only on ninth grade test scores, our longitudinal data enable us to investigate spillover effects for outcomes measured as early as age 8.

This paper also naturally relates to the vast literature on school starting age (Deming and Dynarski 2008). A number of recent studies find that children who are relatively older than their classmates have improved cognitive outcomes (Bedard and Dhuey 2006; McEwan and Shapiro 2008;

[4] Another paper, Qureshi (2018), 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 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.

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; Landerso et al. 2017), and higher wages (Kawaguchi 2011; Fredriksson and Öckert 2014). The effects of school starting-age policies on policy-relevant, long-run outcomes, however, have been more mixed (Dobkin and Ferreira 2010; Black et al. 2011; Dhuey et al. 2019). We think about this extant literature as essentially providing the first-stage for a question that we are primarily interested in answering – namely whether sibling spillovers exist in human capital accumulation.[5]

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, the causal estimates imply that human capital shocks to the older child in a family with low socioeconomic status (SES) can trickle down to other children, while in more affluent families, parents can engage in reinforcing behavior even in a developed country setting. To our best knowledge this is the first paper that identifies reinforcement among high-SES families in developed country using quasi-experimental design. These findings in turn 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.

2 Theoretical considerations

To help interpret our empirical findings, in this section, we propose a simple model of human capital production that incorporates sibling spillovers arising through two possible channels as well as socioeconomic status of the family. 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 ever-FRPL-families) and unconstrained (corresponding to never-FRPL-families); (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 accurate proxies for children's human capital accumulation.

Human capital production function is of the following form:

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

[5] It is worth noting that thus far only three papers (Black et al. 2011; Dhuey et al. 2019; Landerso et al. 2019) 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.

where i denotes parental investments; μ_j is a shock to older child j and μ_k is a shock to younger child k.[6] Parents value human capital of their children as well as their own current consumption. Thus, their utility function can be written as:

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

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:

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

Let's now consider how human capital of the sibling (older or younger) is affected by a shock (μ) to 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 shock to 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 school entry shocks. That is, based on the evidence presented in the prior literature (reviewed in Section 1) and our findings in Table A9, 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 across all demographic groups irrespective of birth order.

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). Because

[6] 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 both the school-entry shock and parental investment decisions. Human capital of a particular child in this model depends on both his/her own shock and his/her sibling's shock, but we only consider one degree of contagion (i.e., a shocked child can affect their brother or sister, through, for example, mentoring, but this direct effect does not feed back to shocked child's human capital). An intuitive interpretation of our school-entry shock is to think about it as a change in a child's perceived ability that is an input into his/her human capital production.

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: $(i)\ \frac{\partial \Psi_k}{\partial \mu_j} > \frac{\partial \Psi_j}{\partial \mu_k} = 0$.[7] 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 to a larger degree. It is plausible to expect that families which 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.[8] 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 policy shock (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 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 a positive human capital shock to 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- 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 can be 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

[7] 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.

[8] 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 for complete family with two children (i.e., household size of four) of 31,543 USD. For the same year Lino (2000) reports that a 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 depending on a child's age, 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 (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 working 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).

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). More generally, 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 compensating 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. 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 the shock. 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 reinforce the shock and positive effects if they compensate for it.

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. It 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.[9] As such, we restrict the sample in our main analysis to adjacent sibling pairs born between 1976 and 1996, both of whom were tested at least once in these grades during our

[9]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. Our main older-to-younger spillover results among students in families ever on FRPL are somewhat larger in the CTBS sample (coefficient of 4.68 with SE of 1.35) than in the FCAT sample (coefficient of 3.31 with SE of 1.39), and in both cases they are statistically significant at conventional levels. We do no have enough power to detect statistically significant subgroup estimates among families never on FRPL.

sampling frame.[10] We drop (1) sibling pairs whose 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. Schooling records further provide us with additional information about student health, ability, and progress through schooling. We use this information to create indicators for whether a child has ever been diagnosed with either a cognitive/behavioral disability or a physical disability by grade 8; whether he/she has ever been enrolled in the gifted program by grade 8; and whether the child has ever repeated a grade between kindergarten and grade 8.

Columns 1 and 2 of Table 1 present the descriptive statistics for all students born between 1976 and 1996 and students in these birth cohorts who were tested at least once between the third and eight grades. 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.[11] Comparison of these numbers to 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. We observe 263,811 children born in these cohorts in our school records, but only 193,094 children have at least one test score in grades 3 through 8. There are three main reasons that 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. During that time frame, it was easier for students to avoid, through exceptions, being tested.[12] Children for whom we observe at least one test score are 4 percentage points more likely to ever experience FRPL in their family compared to all children in the data; however, this may be driven solely by the fact that we also observe them for a longer period of time. On the other hand, the racial composition is highly similar across columns 1 and 2 of Table 1.

Subsequent columns of Table 1 present characteristics of some of our estimation samples. First, columns 3 and 4 document the composition of our preferred sample that maximizes the number of

[10] 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.

[11] 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 during 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).

[12] 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.

children we can observe with test scores. Column 3 describes the younger sibling characteristics in the older-to-younger sibling spillover sample, while column 4 describes older sibling characteristics in the younger-to-older sibling spillovers sample. Comparing the characteristics of these spillover samples with all tested children in column 2 reveals that our preferred estimation samples are somewhat positively selected. These children are slightly more likely to be White, to have higher achievement levels, and to have increased propensity of being enrolled in a gifted program, however, the poverty rates are very similar in the two samples. We use these larger samples as our preferred specification to maximize statistical power; however, as columns 5 and 6 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 school-entry cutoff.[13] 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 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):

[13] 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 sibling 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.

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

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, and $A_j \equiv 1(S_j > 0)$ is an indicator for having a focal child who was born after the school-entry 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. 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). 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 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 six 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 12 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. Appendix Table A1 replicates this analysis using the non-parametric method (Calonico et al. 2017), and reaches a similar conclusion. Furthermore, Appendix Table A2 shows that balance holds within samples stratified by SES, a finding important for our main heterogeneity analysis.

Further evidence in support of this claim can be found by examining the distribution of the running variable in our samples (McCrary 2008). We present graphical evidence in Figure 1 for the distribution of sibling births around focal child school-entry cutoff, 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 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.[14] 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 A3, in which we show the likelihood of younger sibling birth 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 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) sibling being eligible to start school a year earlier on the outcomes of the younger (older) child 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 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

[14] It can be argued that our running variable is discrete rather than continuous and that we should implement a test that corrects for this data feature (Frandsen 2018). 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 present results wherein we exclude data points near the cutoff.

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 A of Appendix Figure A1).[15] These statistics suggest that for the cohorts we study, the school-entry policies were binding and had sizable effects on school-entry age.

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, 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.

First, we find significant positive sibling 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 starting 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

[15] It is important to note that if there was no redshirting/fast-tracking, this discontinuity would be almost exactly 1 year. Panel B of Appendix Figure A1 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. 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. Our results are also very similar in a sample of families whose children start school on-time; we return to this issue in Section 7, in which we discuss IV estimates. 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 this section.

siblings fall on 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 14.7 percent of the standard deviation) better on standardized tests in elementary and middle school 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.

To assess the magnitude of this positive spillover effect in less affluent families, it is helpful to compare it 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 over 50 percent larger than the size of birth order gap in reading scores (Breining et al. 2018). 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). In terms of the test score gaps between different student groups in our data, the spillover effect 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. Furthermore, this effect 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 2001). Thus, we view these estimates as not only plausible but also comparable to magnitudes found in other studies of the determinants of cognitive development.

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 households in which the younger child is endowed with the school-entry advantage, perform worse than older children in households in which the younger child is relatively more disadvantaged due to their school-entry eligibility. These negative estimates are about 55 percent smaller (in absolute terms) than our main positive findings from panel A, but they nonetheless suggest that families who do not face binding budget constraints can engage in reinforcing behavior in a developed country setting. To our best knowledge this is the first paper finding support for such household behavior using quasi-experimental variation. Finally, 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 estimated coefficients remain stable with the addition of own and sibling characteristics, lending further credibility to the identifying assumptions discussed in Section 4.2.[16]

Our empirical findings could be explained by several mechanisms described in Section 2. First, our results suggest 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 of reinforcing behavior among high-SES households, which offsets the positive direct effect in the older-to-younger setting, and leads to negative sibling spillover effects in the younger-to-older setting. Finally, the results are also consistent with negligible indirect effects among 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.[17]

To fully understand the variation in our data, Figure A2 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. Thus, our conclusions of direct spillovers among less affluent and modest reinforcing behavior among more affluent families remain unchanged.

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 observed abilities of their children.[18] Figure A3 presents

[16] We also examine the spillovers separately for mathematics and reading; these results are presented in Table A4. Positive estimates in older-to-younger sample with families ever on FRPL are similar across the two subjects, whereas in our younger-to-older analysis among never-FRPL families statistically significant effects are driven by reading only. This makes sense if home production, and thus resource reallocation, contributes more to reading than mathematics performance. In further examining the stability of our results across grades, we restrict the sample to individuals for whom we observe six consecutive assessments starting from grade 3. These estimates are comparable across grades, but standard errors in this reduced sample are large enough that we mostly lose statistical significance at conventional levels. 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.

[17] It is important to note 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, our results are not quite 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 Section 6 we show that older siblings' test scores are affected only once the younger child starts being tested and their differential relative ability is revealed. Nonetheless, a caveat remains: we do not have information on actual investments in children, and, therefore, with the help of theoretical predictions, we equate observed differential outcomes with particular pattern of parental behavior.

[18] It can be argued that another comparison that highlights distance and comparability between children 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. 2018). Nonetheless, for completeness, in Figures A4 and

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 mentor/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 sibling 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 A6 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 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 stronger reinforcing behavior 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 (driven by the school-entry policy), and allocate more resources toward this child because a higher return on the investment is more likely (Becker and Tomes 1976).

Finally, in Table A5, we examine spillover effects on other academic and health-related outcomes. Although the results are not equally striking, the general pattern of the findings in these tables mimics results presented in Table 3. Namely, in the older-to-younger analysis, we find larger and – for some outcomes – statistically significant spillovers in the less affluent sample, while we find no consistent spillover effects for sibling pairs in more affluent families. Reductions in cognitive/behavioral disabilities and grade repetition are the strongest results in this analysis. Similarly, we find inconsistently signed and never statistically significant coefficients in the younger-to-older

A5 we also examine differential effects by sex composition of the sibling pair as well as by gender of child for whom we observe outcomes. We find larger positive spillover effects from the older to the younger child in low-SES families when children are of the same-sex, which makes sense if these children are more likely to interact with each other. Likewise, estimates in this sample are larger for males who are more sensitive to inputs. On the other hand, we find negative effects from the younger to the older child in high-SES families with opposite-sex children and there are no striking gender differences in this sample. Our estimates in the remaining two samples are not consistently signed. Overall, we view this gender analysis as inconclusive and choose not to highlight it due to interpretational complexity.

analysis. We view these additional results as less informative compared to our primary test score analyses because it is not clear whether reduced disability labeling is a positive or a negative outcome. On one hand, this could reflect better health; but it could also be a sign of lower parental involvement if parents are less likely to seek the additional resources their children may need.

6 Robustness

This section presents a series of analyses to gauge the robustness of our main test score findings presented in column (6) in panel A and column (9) in panel B of Table 3.[19] 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.

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.[20] This is akin to the randomization inference idea, 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 similar to the p-value of 0.070 reported in the main regression table. Overall, we view these

[19] 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 A7 to A10.

[20] For example, if the threshold is September 1, 1985, then in the two most extreme cases, we are comparing siblings 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 over 201 coefficients.

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.

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 A6, and are in line with our parametric estimates.[21] This procedure, however, produces larger standard errors than our polynomial approach; therefore, our reinforcing behavior result is no longer statistically significant. 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, the point estimates for test scores among higher SES families in younger-to-older spillovers are very similar across the two methods.

Sixth, we check the robustness of our findings to alternative measures of family SES. In Table A7, we repeat our main analysis using neighborhood income to define family affluence. This measure, by construction, is noisier compared to individual FRPL eligibility, and the correlation between the two variables is only 0.44.[22] Nonetheless, columns (1) and (2) of Table A7 divide the sample by median neighborhood income; columns (3) to (5) into terciles of this variable. In panel A, akin to results from Table 3, we find positive spillovers among families living in relatively more impoverished neighborhoods and no effects for households living in relatively more affluent locations. Conversely, in panel B we find zero to small positive estimates in lower-SES and negative coefficients in higher-SES neighborhoods. The latter set of results is not statistically significant at conventional levels (p-values of 0.14 and 0.18 in columns (2) and (5), respectively) but largely confirms our findings from Table 3, and the point estimates are comparable. Overall, these findings suggest that our main

[21]For seven outcomes and 30 specifications/samples under investigation across the balancing check and the main results, the optimal bandwidths range from 27 to 56 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).

[22]These neighborhood units reflect division at a 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 sourced from the American Community Survey.

conclusions are robust to different measures of SES, even if derived at the neighborhood rather than the individual level.

Finally, in the sample where younger children are affected by the school-entry policy, we can examine their older siblings' test scores before and after the commencement of focal child testing. We use this information to conduct a falsification exercise: if the observed spillover effects are indeed driven by the school-entry policy, one would expect to find no significant discontinuity in the older sibling test scores before the younger child is tested for the first time. We present the results of this analysis in Table A8, 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 in two out of the three samples considered.[23]

7 Interpreting the results through IV

The analysis thus far has focused solely on reduced-form effects. While this approach is useful to investigate the existence and direction of sibling spillovers, another interesting and policy-relevant question is the causal effect of a student's achievement on the educational outcomes of his/her sibling. Such reframing of the relationship puts both inputs and outputs on the same scale and enables simple quantification of the magnitude of estimated externality. A complicating factor in assessing this effect in our context, in addition to monotonicity concerns outlined earlier, is that school-entry policies have been shown to affect not only the test scores of the focal students but also their non-cognitive traits, such as leadership skills (Dhuey and Lipscomb 2008) or the likelihood of engaging in risky behavior (Depew and Eren 2016; Cook and Kang 2016; Landerso et al. 2017). This, in turn, raises the possibility that the exclusion restriction in an IV framework in which we instrument for the focal sibling's test scores using his/her birth date around the school-entry cutoff is violated, thus possibly leading to biased estimates. The direction of this bias would depend on how these alternative channels affect the sibling's educational outcomes independently of the test score effects of the policy on the focal child. For example, conditional on test scores, if better leadership skills of the focal student impose positive externalities on siblings, then the IV estimates will overestimate the true causal effect of the focal student's test scores on their siblings' test scores.

With these caveats in mind, invoking exclusion restriction and monotonicity, Table A9 reports the results of IV analysis. In panel A, we present the OLS coefficients on the third through eight grade average test score of the focal sibling in regressions where the dependent variable is the

[23]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. 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.

pooled third through eight grade test scores of their younger sibling (columns 1 to 3), or his/her older sibling (columns 4 to 6), for all sibling pairs and broken down by family affluence. These are the correlations mentioned in the introduction and the causality of which we attempt to assess in this paper. In panel B, we repeat the same analysis, restricting the sample to focal siblings born 60 days before or after their school-entry cutoff. In both cases, we find inter-sibling associations in the range of 0.42 to 0.51 percentile points that are highly statistically significant. In panel C, we provide the reduced-form estimates that are identical, by design, to those presented in columns 3, 6, and 9 of Table 3.

We then instrument for the average test scores of the focal student using his/her school-entry cutoff. In the first row of panel D, we present the first-stage results of this analysis, which reveal significant differences in average test scores between students born right before and right after the cutoff. For example, focal children in the older-to-younger analysis who were born right after the cutoff score 4.3 to 6.0 percentiles (or roughly 16 percent to 23 percent of the standard deviation) better than focal children who were born right before the cutoff. These own-effects are comparable across older-to-younger and younger-to-older samples, and by family affluence. Furthermore, effective F-statistics suggest that we have strong instruments (Montiel Olea and Pflueger 2013). The subsequent row of panel D presents the second-stage results. For all cases except the older-to-younger effects in less affluent families, the IV estimates suggest that the observed inter-sibling associations obtained using OLS overestimate the effect of one sibling's academic performance on that of another. Pooling all families together (column 1), the estimate implies that a 1 percentile increase in older child achievement increases his/her younger sibling achievement by 0.39 percentiles – an externality of 39 percent. These average effects are larger than those documented for peer (Nicoletti and Rabe 2014) or teacher (Qureshi 2017) effects but are much smaller than what was found in Denmark (Landerso et al. 2019). This makes sense if peer and teacher effects create more transitory shocks than school-entry policies; on the other hand, the large positive estimate for Denmark stems from a policy affecting the household at a crucial time for the older child's. Our estimates are larger for poor families with an almost 80 percent spillover rate, and consistent with our main findings, we do not find sizable or statistically significant spillovers from older-to-younger children among more affluent families. Reinforcing spillover in high-SES families in column 6 implies an externality of about 37 percent, but this estimate, is marginally statistically insignificant at conventional levels (p-value of 0.12).

Estimates in Table A9 use averaged test scores in grades 3 through 8 as the endogenous variable. As such, they limit the scope for school-entry age fade-out effects (Elder and Lubotsky 2009) and represent treatment more closely resembling differential human capital rather than just relative age advantage. Another question of interest, however, is what the IV scaling implies when we use third grade tests scores as an endogenous variable – the first time that the focal child advantage is measured and directly communicated to the parents. As expected, our first-stage estimates in this setting increase to about 7 to 8 percentile points. The older-to-younger positive spillover among low-SES families decreases to 0.56 percentiles, while younger-to-older negative spillover among high-SES families decreases to 0.30 percentiles and is now statistically significant at 10 percent level. Taken together, we estimate that positive externalities generated by school-entry policies affecting the older child could range in magnitude from 56 percent to 77 percent among low-SES households but unlike the effects presented in some previous research, they never exceed own effects (Breining et al. 2016). Likewise, in contrast to earlier work, we detect negative reinforcement among high-SES households of up to 37 percent.

8 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 sibling 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.

Using student-level administrative data that enable us to identify siblings, and 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 sibling 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 that are larger for low-performing older children, families in which siblings are spaced closer together, and 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/tutoring 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. The latter set of findings has, to our best knowledge, not been documented to date in developed country context when using quasi-experimental variation.

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. Thus, policymakers should be cautious when designing and implementing such interventions, since their target population could differ from the full population of affected individuals. Finally, our findings suggest that the same educational intervention 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.

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," 2017.

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*, 2018, *forthcoming.*

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.

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.

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," 2017.

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 230062, 2017.

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

_ **, Joseph Doyle, David Figlio, Krzysztof Karbownik, and Jeffrey Roth**, "Birth order and delinquency: Evidence from Denmark and Florida," *Journal of Labor Economics*, 2018, *forthcoming.*

__ , Meltem Daysal, Marianne Simonsen, and Mircea Trandafir, "Spillover Effects of Early Life Medical Interventions," IZA Discussion Paper 9086, 2016.

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

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

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, *forthcoming.*

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.

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*, 2018, *forthcoming.*

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

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

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.

Heissel, Jennifer, "Spillover effects within families: Evidence from teenage motherhood and sibling academic performance," 2018.

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.

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.

Landerso, Rasmus, Helena Nielsen Skyt, and Marianne Simonsen, "School starting age and the crime-age profile," *Economic Journal*, 2017, *forthcoming.*

_ , _ , and _ , "How going to school affcts the family," *Journal of Human Resources*, 2019, *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*, 2019, *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," IZA Discussion Paper 8615, 2014.

Olea, Jose Montiel and Carolin Pflueger, "A robust test for weak instruments," *Journal and Business and Economic Statistics*, 2013, *31* (3), 358–369.

Ozier, Owen, "Exploiting Externalities to Estimate the Long-term Effects of Early Childhood Deworming," *American Economic Journal: Applied Economics*, 2017, *forthcoming.*

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

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, "Siblings, Teachers and Spillovers on Academic Achievement," *Journal of Human Resources*, 2017, *forthcoming.*

_ , "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," *Economic Journal*, 2018, *forthcoming.*

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, *forthcoming.*

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.

Tables

Table 1: Descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All 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	74.6	78.8	79.1	79.3	79.6
Ever on free or reduced price lunch	50.2	54.2	53.5	53.6	53.4	53.4
Female	48.3	49.2	49.7	48.5	49.7	48.3
Birth year	1986.5	1986.3	1987.3	1984.4	1988.6	1985.6
			Panel B: Grades 3 to 8 outcomes			
Test score	-	58.5	60.0	62.1	60.7	62.9
Gifted	-	9.9	10.8	12.8	11.2	13.2
Disabled	-	13.1	13.0	10.6	13.2	10.4
N	263,811	193,094	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 1975 and 1996; column 2 additionally requires at least one test score observation in grades 3 to 8; column 3 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 4 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 5 and 6 present subsamples of the previous two columns for the sample of the first two births in family. Running variable in columns 3 through 6 is restricted to +/- 150 days.

Table 2: Discontinuities in background characteristics

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

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. Outcome variables are: indicator for White student, indicator for female student, birth month, birth year, age difference between siblings, and indicator for family ever being on free or reduced price lunch. Indicators are multiplies by 100. Standard errors clustered at running variable at daily level.

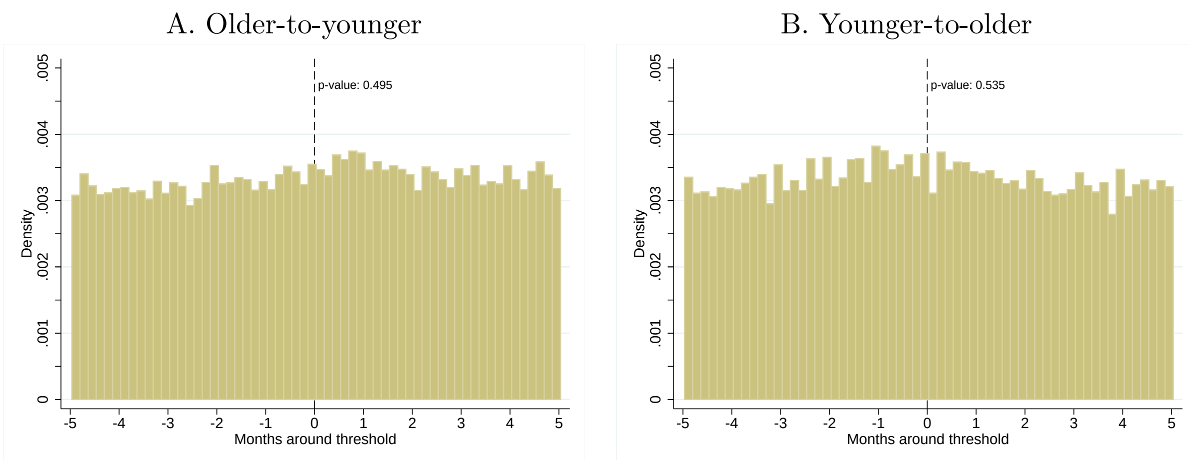
Table 3: Main results: sibling spillovers in test scores

	(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 cutoff	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 cutoff	-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			25,981	
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. 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 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, indicator for sibling being a female, and age difference between siblings in days. Standard errors clustered at running variable at daily level.

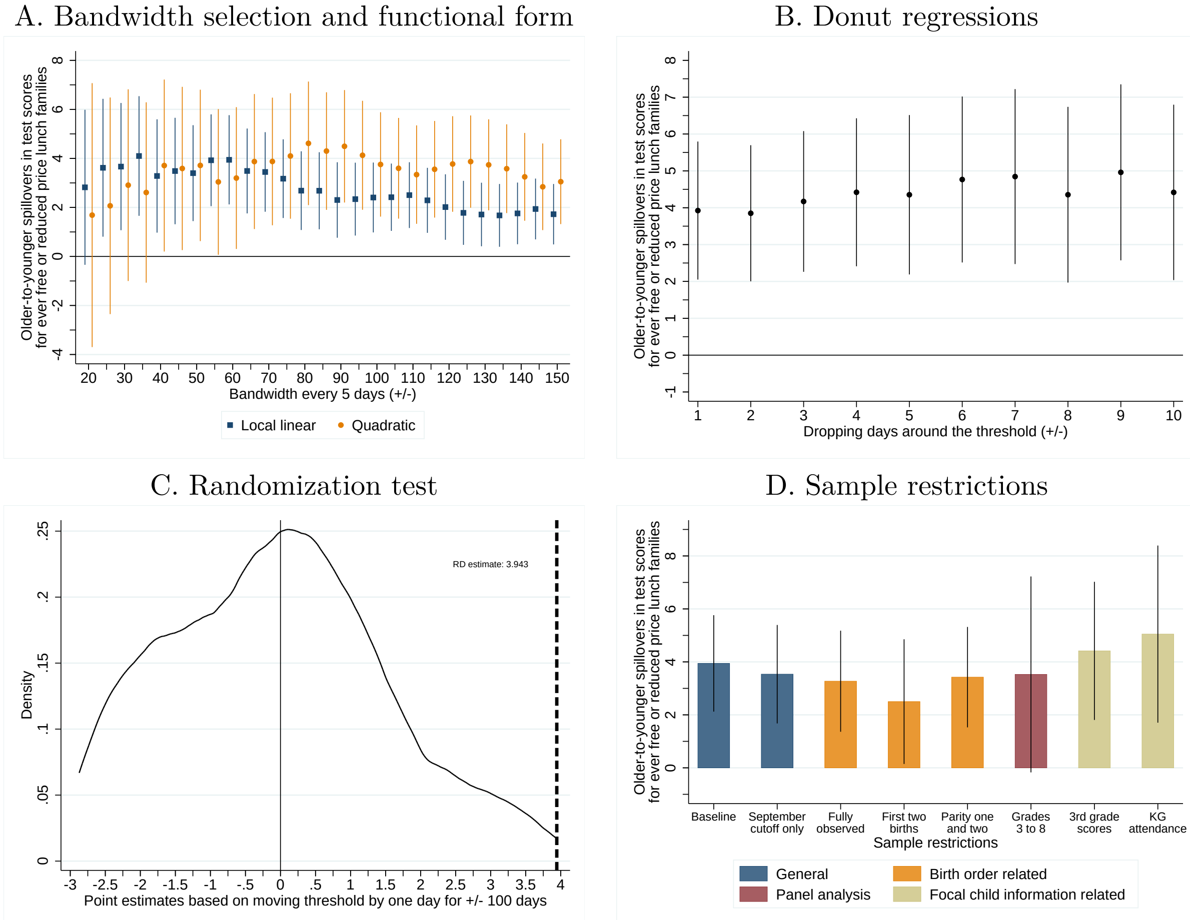
Figures

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



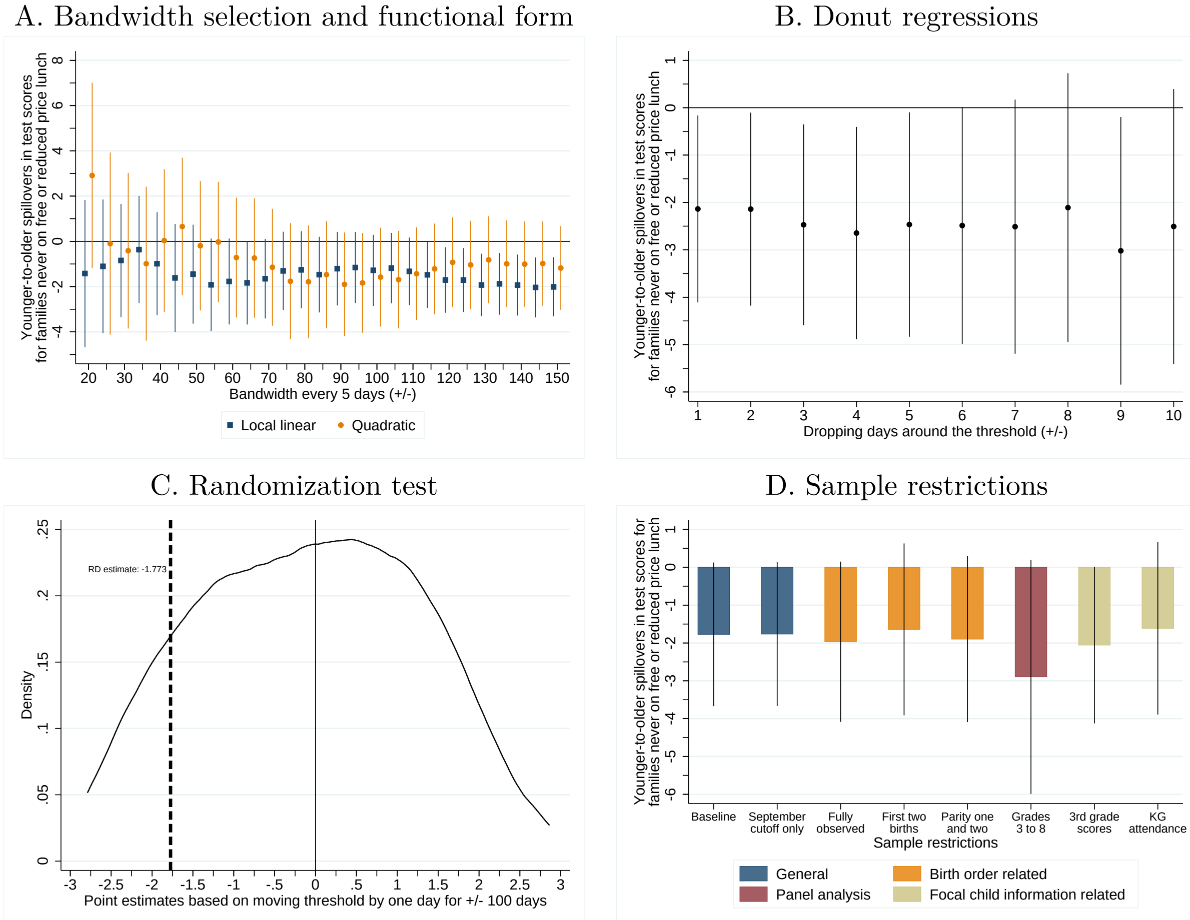
Note: Figures present histogram of density of children born around their sibling's school-entry cutoff (+/- 150 days). Bin width is set to 5 days. Sample in panel A is based on births of younger siblings around the older sibling cutoff while sample in panel B is based on births of older sibling around 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 result: 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 regression 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; the x-axis presents the range of estimates, and in each regression we use local linear specification with +/- 60 days bandwidth. 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 result: 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.

Appendix Tables

Table A1: Discontinuities in background characteristics. Non-parametric estimates

	(1) White	(2) Female	(3) Birth month	(4) Birth year	(5) Age difference	(6) Ever on free or reduced price lunch
			Panel A: Older-to-younger			
Focal child born after school entry cutoff	-0.297	-2.235	0.092	-0.294	-10.593	4.174
	(2.380)	(2.343)	(0.138)	(0.240)	(30.055)	(2.585)
Mean of Y	78.8	49.7	6.6	1987.3	1024.5	53.5
Implied imbalance (%)	-0.4	-4.5	1.4	0.0	-1.0	7.8
Bandwidth	38	31	45	41	38	43
Observations			34,331			
			Panel B: Younger-to-older			
Focal child born after school entry cutoff	-1.139	-0.893	-0.154	-0.235	-36.287	0.888
	(1.657)	(1.425)	(0.168)	(0.307)	(29.493)	(2.220)
Mean of Y	79.1	48.5	6.6	1984.4	1020.5	53.6
Implied imbalance (%)	-1.4	-1.8	-2.3	0.0	-3.6	1.7
Bandwidth	27	31	35	40	38	51
Observations			33,290			

Note: This table is a non-parametric version of results presented in Table 2 estimated using method developed by Calonico et al. (2017). Number of observations and means reflect sample based on +/- 150 days around school-entry cutoff. Standard errors clustered at running variable at daily level.

Table A2: Discontinuities in background characteristics by free and reduced price lunch status

	(1) White	(2) Female	(3) Birth month	(4) Birth year	(5) Age difference
			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
	(2.149)	(2.534)	(0.146)	(0.183)	(24.574)
Mean of Y	69.2	50.0	6.7	1987.5	980.3
Implied imbalance (%)	-1.1	8.1	-3.4	0.0	-1.0
Observations			7,577		
			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
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
	(2.086)	(1.888)	(0.154)	(0.208)	(23.210)
Mean of Y	69.8	48.8	6.6	1984.6	973.9
Implied imbalance (%)	3.5	3.0	3.2	0.0	-1.7
Observations			7,498		
			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
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 A3: Effects of older sibling discontinuity on 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. 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 A4: 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 and reading test scores. 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 A5: Estimates for non-test score outcomes.

	(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: Cognitive or behavioral disability			
Focal child born after school entry cutoff	-2.518*	-3.941**	-0.759	0.091	0.147	-0.719
	(1.480)	(1.921)	(1.912)	(1.104)	(1.640)	(1.149)
Mean of Y	16.5	21.6	10.5	13.5	18.5	7.8
Observations	12,863	6,906	5,957	13,059	7,023	6,036
			Panel B: Physical disability			
Focal child born after school entry cutoff	0.381	0.771	0.034	-0.450	-0.324	-0.903
	(1.214)	(1.744)	(1.682)	(0.893)	(1.418)	(1.055)
Mean of Y	13.7	15.7	11.7	9.1	11.5	6.4
Observations	12,457	6,418	6,039	12,420	6,470	5,950
			Panel C: Ever gifted			
Focal child born after school entry cutoff	0.270	1.714*	-0.989	-1.712	-0.623	-2.846
	(1.036)	(0.906)	(1.917)	(1.082)	(0.977)	(1.883)
Mean of Y	10.9	4.8	18.0	12.7	6.0	20.6
			Panel D: Ever repeating grade			
Focal child born after school entry cutoff	-1.829**	-3.578**	-0.107	-0.298	-0.953	0.082
	(0.864)	(1.527)	(0.849)	(0.721)	(1.133)	(0.696)
Mean of Y	8.1	11.8	3.8	5.4	7.9	2.5
Observations	14,148	7,577	6,571	13,854	7,498	6,356

Note: Local linear regressions with bandwidth of +/- 60 days for reduced-form estimates based on equation 7. Each panel presents point estimates, standard errors, sample sizes and means for indicator for ever being diagnosed with cognitive or behavioral disability (panel A), indicator for ever being diagnosed with physical disability (panel B), indicator for ever being in a gifted program (panel C), and indicator for ever repeating a grade (panel D). Indicator variables are multiplied by 100. 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 with the exception of grade indicators and test type (FCAT). Standard errors clustered at running variable at daily level.

Table A6: 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	1.465	2.086*	2.071*	3.751**	3.678**	3.861***	1.013	1.416	1.489
	(1.289)	(1.116)	(1.120)	(1.485)	(1.457)	(1.404)	(1.913)	(1.865)	(1.854)
Mean/SD of Y		60.0/26.4			50.7/26.2			70.1/22.6	
Bandwidth	47	56	55	43	44	47	36	36	35
Observations		137,958			71,887			66,071	
					Panel B. Younger-to-older				
Focal child born after school entry cutoff	-1.199	-1.078	-1.100	-0.511	-0.825	-0.962	-2.218*	-1.576	-1.703
	(1.220)	(1.159)	(1.162)	(2.058)	(1.913)	(1.936)	(1.363)	(1.350)	(1.373)
Mean/SD of Y		62.1/26.6			52.3/26.6			73.1/22.0	
Bandwidth	51	44	43	44	42	41	40	34	35
Observations		135,205			71,766			63,439	
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). Number of observations and means reflect sample based on +/- 150 days around school-entry cutoff. Standard errors clustered at running variable at daily level.

Table A7: Sibling spillovers in test scores: estimates by neighborhood affluence

	(1)	(2)	(3)	(4)	(5)
	By median of neighborhood income		By tercile of neighborhood income		
	Below	Above	Bottom	Middle	Top
			Panel A. Older-to-younger		
Focal child born after school entry cutoff	3.709***	-0.276	4.060***	1.540	-0.236
	(0.922)	(0.988)	(1.233)	(1.088)	(1.360)
Mean/SD of Y	52.0/26.7	67.5/23.8	48.9/26.7	60.9/25.2	69.4/23.2
Observations	27,978	28,697	18,955	17,953	19,767
			Panel B. Younger-to-older		
Focal child born after school entry cutoff	0.801	-1.514	-0.005	0.751	-1.994
	(1.271)	(1.011)	(1.459)	(1.367)	(1.491)
Mean/SD of Y	53.6/27.1	70.0/23.7	50.0/27.0	63.0/25.7	72.3/22.5
Observations	28,173	27,608	19,063	17,572	19,146

Note: Local linear regression with bandwidth of +/- 60 days for reduced-form estimates based on equation 7 and specification with controls from columns 3, 6 and 9 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, indicator for sibling being a female, and age difference between siblings in days. 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 and 2 divide the sample by median of neighborhood income in which family resides while columns 3 to 5 divide the sample by terciles of this measure. Neighborhood income is based on minimum average neighborhood income experienced by the family as defined at busing zone. Standard errors clustered at running variable at daily level.

Table A8: Effects of younger sibling 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 (0.796)	0.015 (0.884)	-1.827* (1.027)	0.071
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 (1.269)	1.362 (1.402)	-0.036 (1.401)	0.254
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* (0.969)	-1.097 (1.129)	-3.300*** (1.231)	0.099
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 specification 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, 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 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.

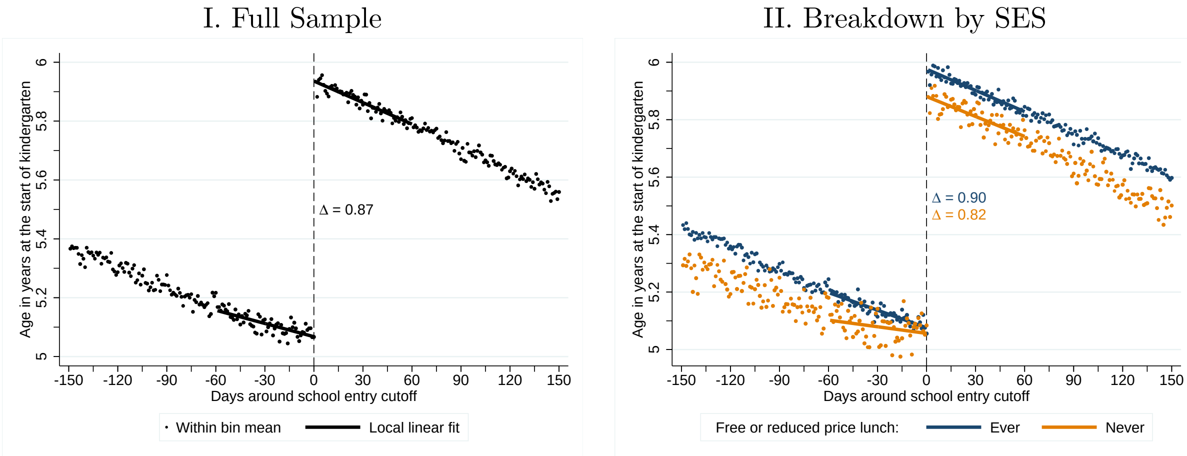
Table A9: Sibling spillovers: Instrumental variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: pooled third to eight grade test scores					
	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. Full sample OLS					
Focal child third to eight grade test scores	0.491***	0.434***	0.432***	0.497***	0.444***	0.417***
	(0.004)	(0.006)	(0.008)	(0.005)	(0.006)	(0.007)
Mean/SD of Y	60.1/26.4	50.8/26.2	70.2/22.6	62.1/26.7	52.4/26.6	73.1/22.0
Observations	165,513	86,202	79,311	163,189	86,504	76,685
	B. Bandwidth restricted OLS					
Focal child third to eight grade test scores	0.491***	0.439***	0.427***	0.508***	0.464***	0.420***
	(0.007)	(0.010)	(0.014)	(0.008)	(0.010)	(0.012)
	C. Reduced-form					
Focal child born after school entry cutoff	1.694**	3.943***	-0.076	-0.588	0.809	-1.773*
	(0.785)	(0.929)	(1.155)	(0.796)	(1.269)	(0.969)
	D. 2 SLS (instrumenting focal child averaged third to eight grade test scores with their school entry cutoff)					
Focal child own effect (first-stage)	4.347***	5.122***	4.327***	5.274***	6.032***	4.760***
	(0.872)	(1.188)	(1.100)	(0.867)	(1.173)	(1.132)
Effective F-stat	18.8	18.9	14.5	51.8	30.1	25.4
Focal child third to eight grade test scores	0.390**	0.770***	-0.018	-0.111	0.134	-0.372
	(0.155)	(0.195)	(0.268)	(0.158)	(0.200)	(0.241)
Mean/SD of Y	59.9/26.4	50.6/26.3	70.0/22.6	61.7/26.7	52.3/26.8	72.7/22.2
Observations	56,701	29,582	27,119	55,813	29,981	25,832

Note: Columns 1 through 3 investigate spillovers from older-to-younger child while columns 4 through 6 investigate spillovers from younger-to-older child. 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. Panel A presents regressions of pooled 3rd- through 8th-grade test scores of younger (older) sibling on averaged 3rd- through 8th-grade test scores of older (younger) child in the full sample of children. Panel B restricts OLS regression from panel A to focal children born within +/- 60 days of their school-entry cutoff, our preferred bandwidth. Panel C recreates reduced-form effects from columns 3, 6 and 9 of Table 3. Panel D presents 2SLS estimates where we instrument focal child's averaged 3rd- through 8th-grade tests scores with their school-entry cutoff in a local linear regression with bandwidth of +/- 60 days. First set of coefficients in this panel presents first-stage relationship together with effective F-statistic (Montiel Olea and Pflueger 2013). Second set of coefficients in this panel presents 2SLS estimates. All regressions include full set of controls akin to columns 3, 6 and 9 of Table 3. Standard errors clustered at running variable at daily level.

Appendix Figures

Figure A1: Discontinuity in age at school-entry

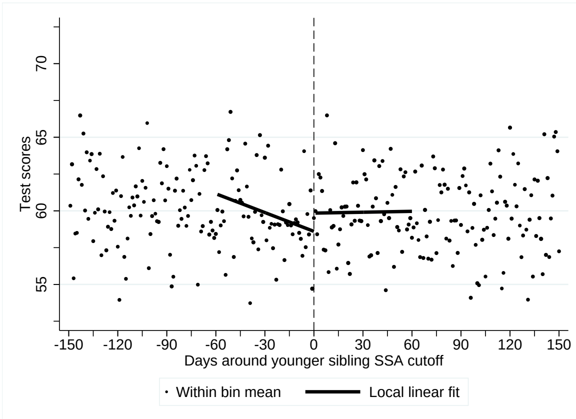


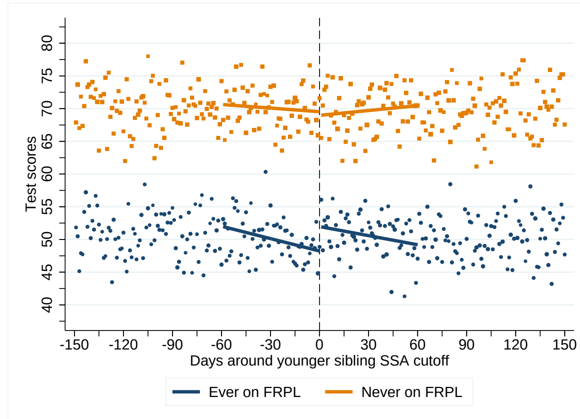
Note: Sample is based on all children first observed in kindergarten. Each figure presenting daily means for +/- 150 days (dots) and linear fits for preferred bandwidth of +/- 60 days (lines) on either side of the school-entry cutoff. Panel A pools all families while panel B divides sample by free or 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 A2: Discontinuity in test scores

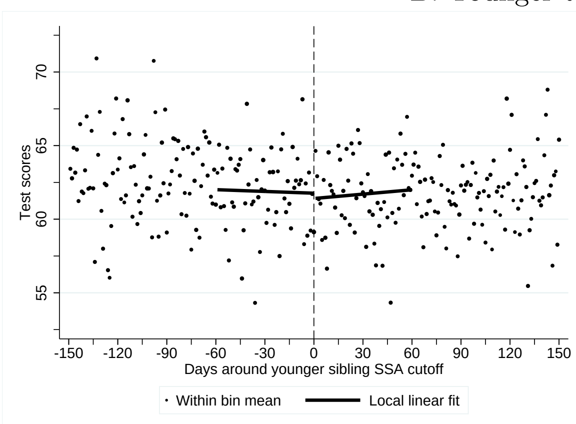
I. Full Sample II. Breakdown by SES

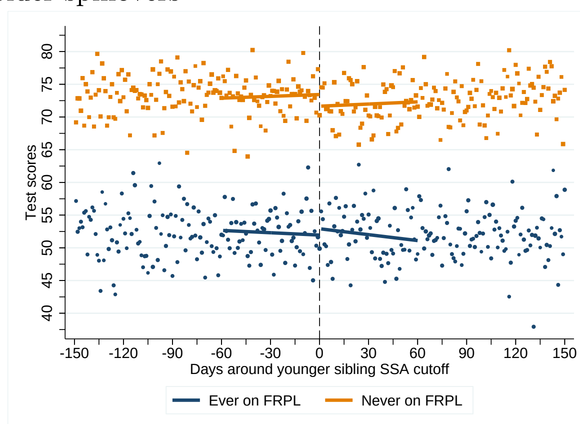
A. Older-to-younger spillovers





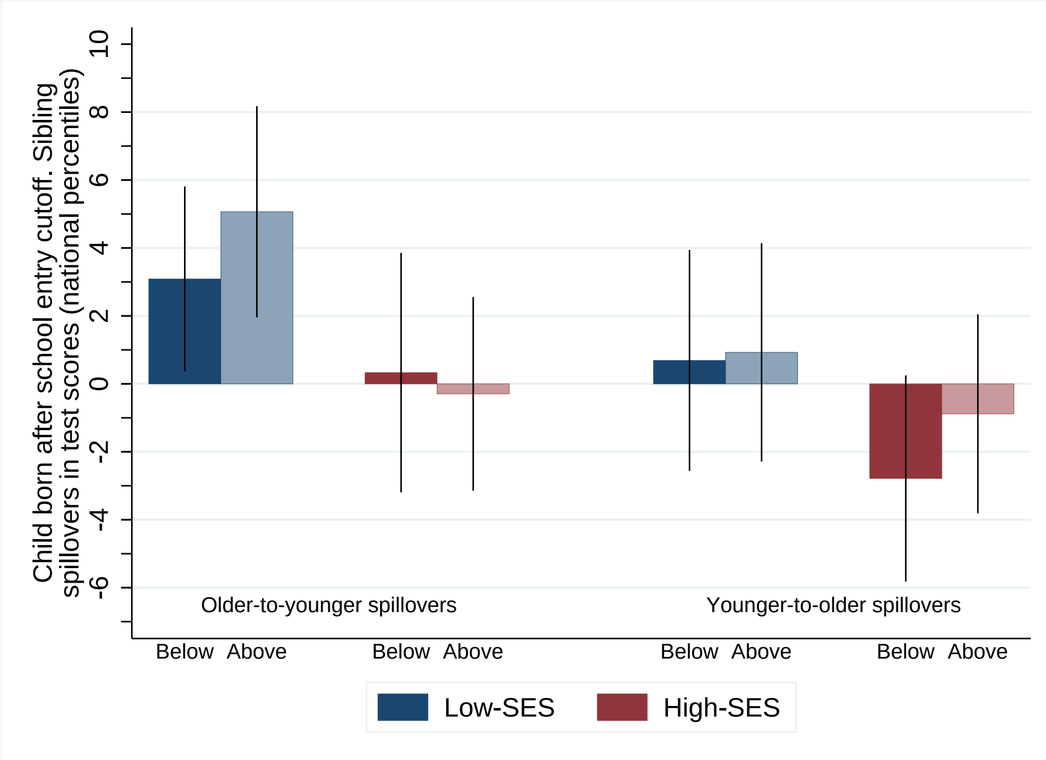
B. Younger-to-older spillovers





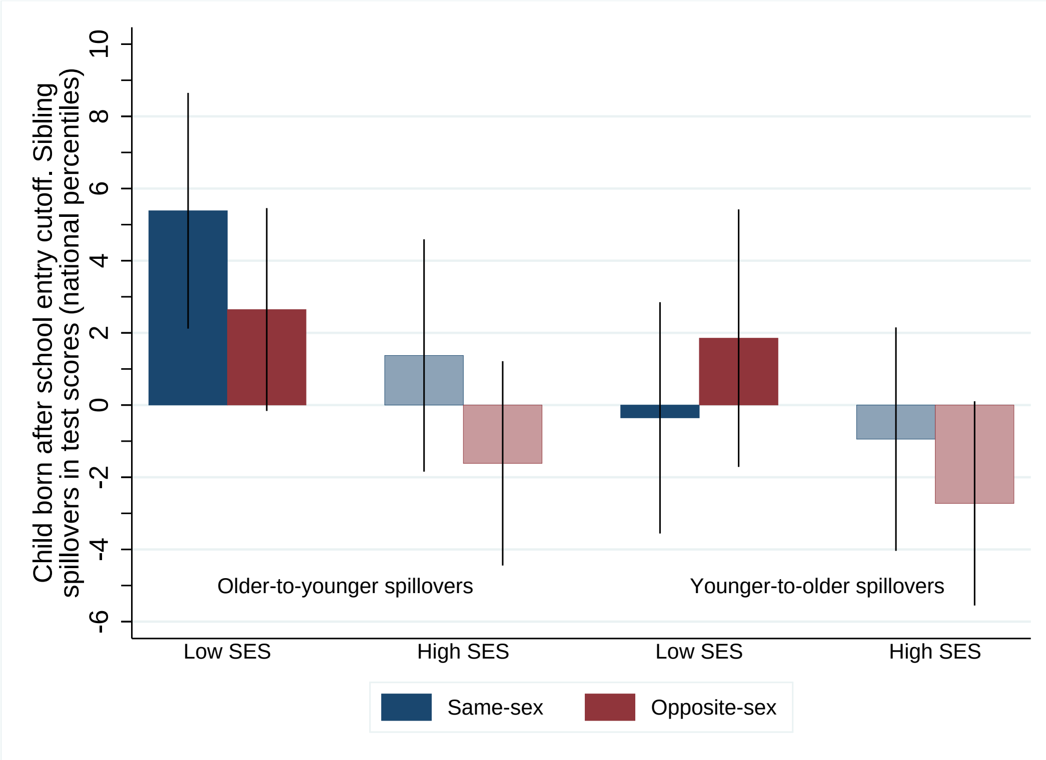
Note: Main results in graphical form with each figure presenting daily means for +/- 150 days (dots) and linear fits for preferred bandwidth of +/- 60 days (lines) on either side of the school-entry cutoff for older child in the sibling pair (panel A) and younger child in the sibling pair (panel B), 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 A3: Heterogeneity analysis by median spacing between siblings



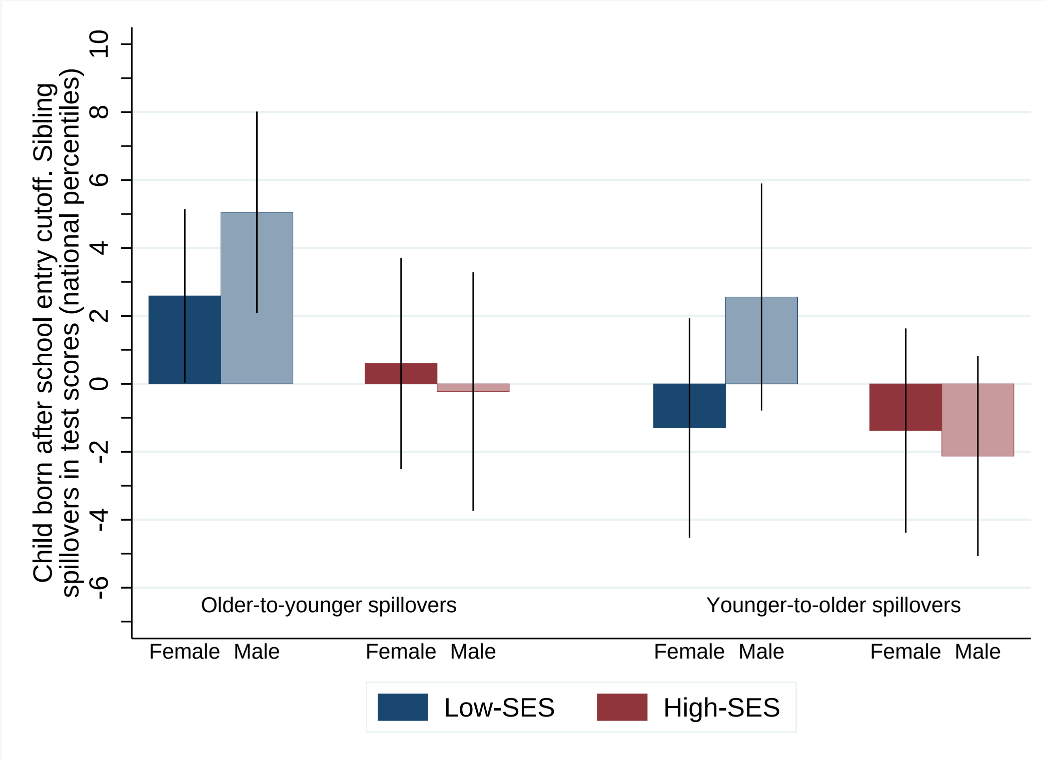
Note: Analysis is based on the sample from columns 4 through 6 in panel A of Table 3; local linear regressions based on +/- 60 bandwidth; control variables include: indicator for White student, indicator for female student, indicators for grades (3 to 8), indicators for birth year, indicators for birth 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 A4: Heterogeneity analysis by gender composition of sibling pair



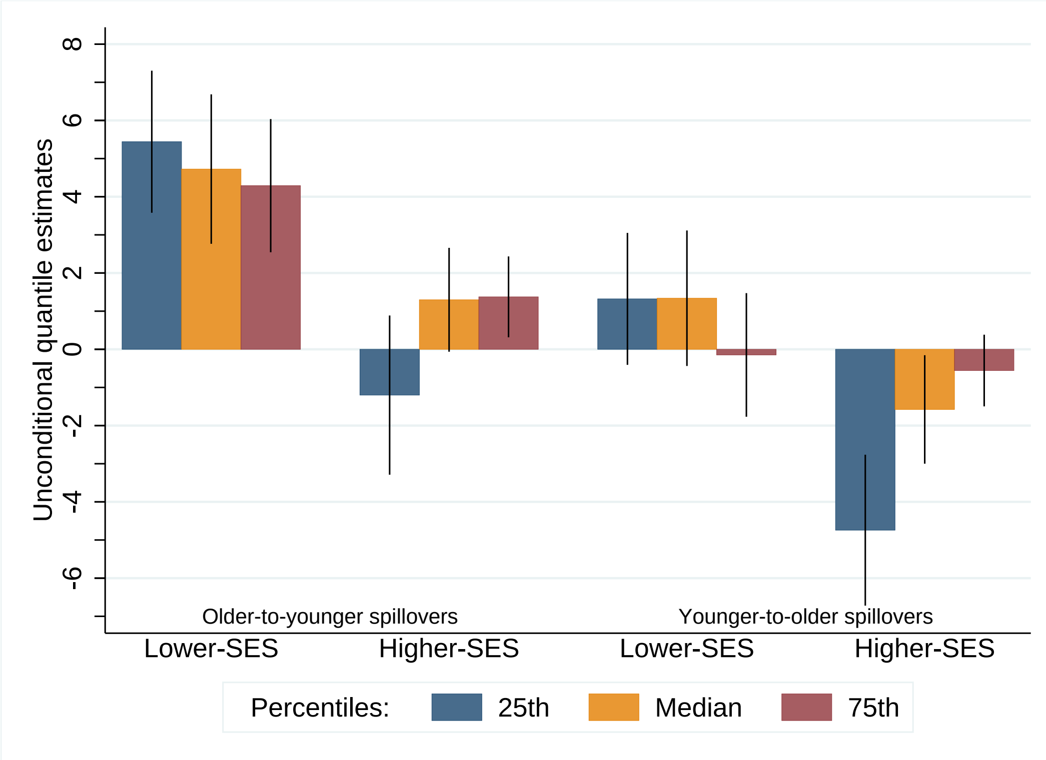
Note: Analysis is based on the sample from columns 4 through 6 in panel A of Table 3; local linear regressions based on +/- 60 bandwidth; control variables include: indicator for White student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts, 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 A5: Heterogeneity analysis by gender of sibling



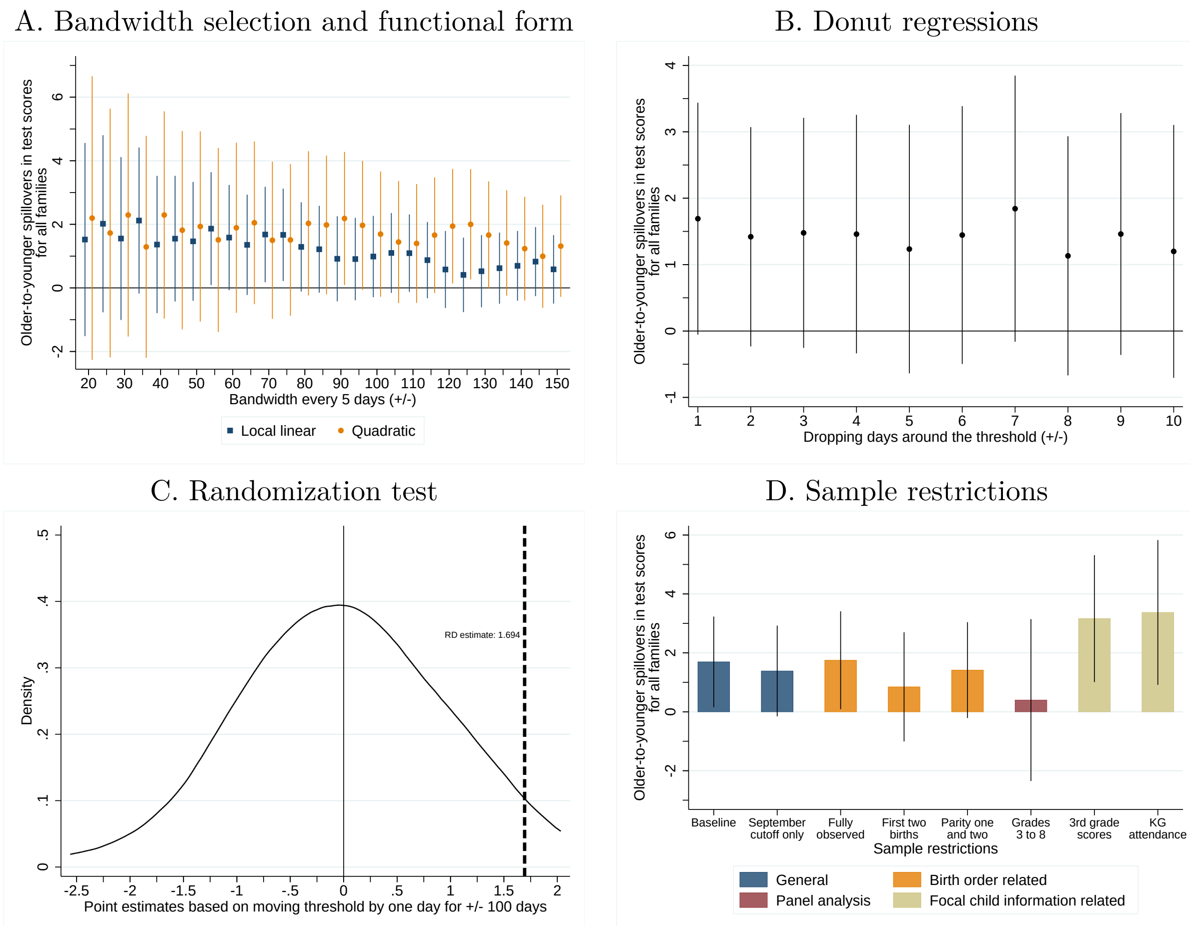
Note: Analysis is based on the sample from columns 4 through 6 in panel A of Table 3; local linear regressions based on +/- 60 bandwidth; control variables include: indicator for White student, indicators for grades (3 to 8), indicators for birth year, indicators for birth month and indicator for FCAT test, indicators for sibling school starting cohorts, 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 A6: Unconditional quantile regression: Test scores



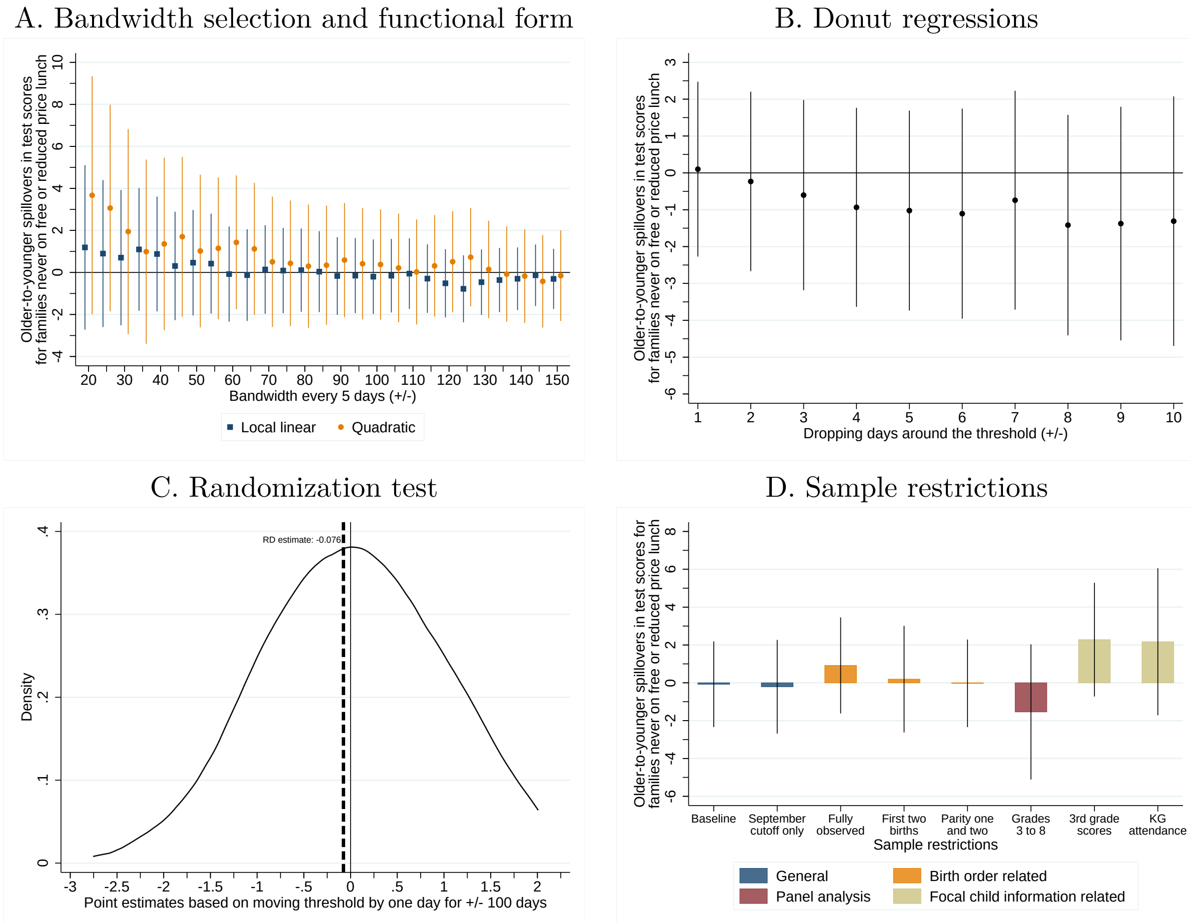
Note: Unconditional quantile regressions (Firpo et al. 2009). Navy bars present estimates for 25th percentile, orange bars present estimates for median, while maroon bars present estimates for 75th percentile of test score distribution. Older-to-younger spillovers in the two left hand side sets of bars 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 in second and fourth set of bars. Bootstrapped standard errors with 200 replications.

Figure A7: Robustness of the main result: Older-to-younger spillovers in test scores among all families



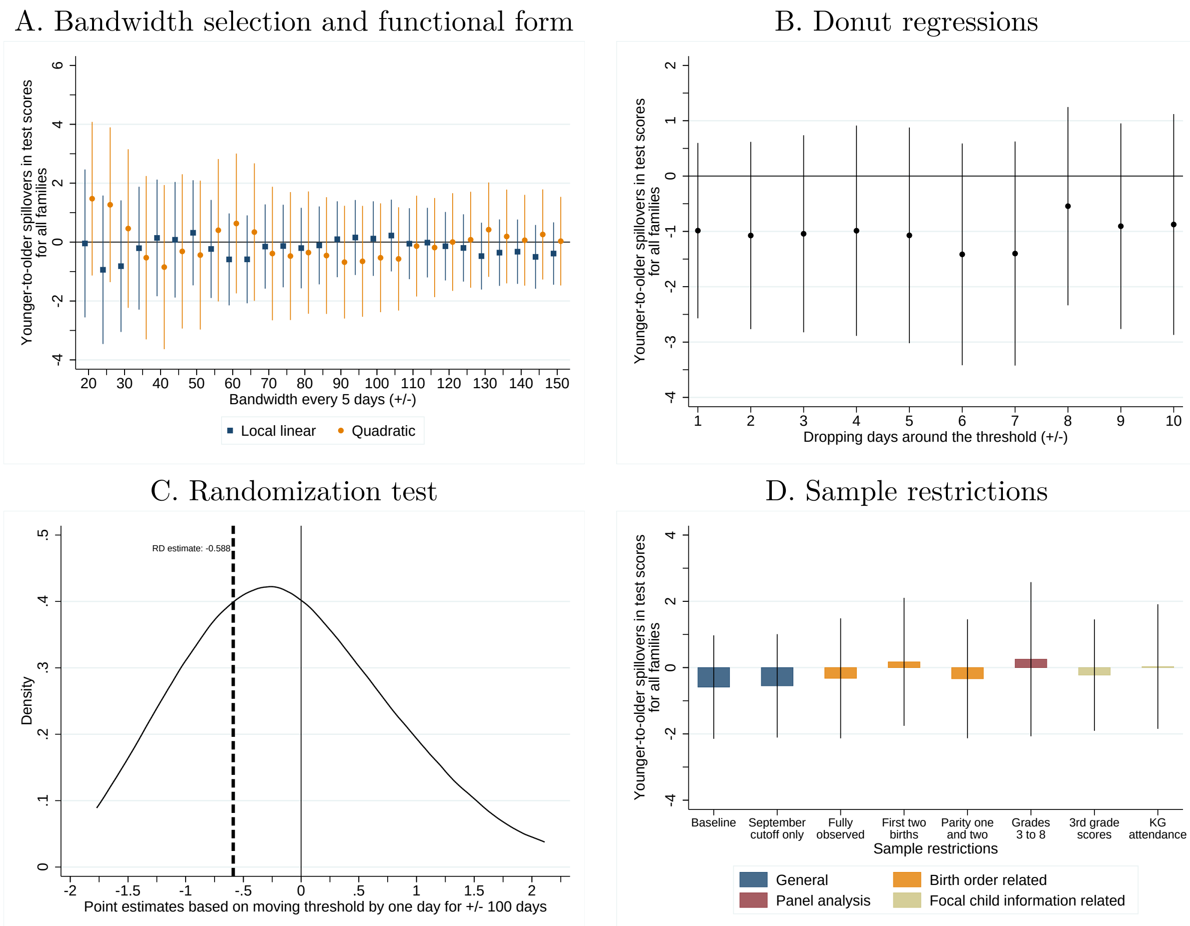
Note: This figure replicates analysis from Figure 2 for the specification from column (3) in panel A of Table 3.

Figure A8: Robustness of the main result: Older-to-younger spillovers in test scores among high SES families



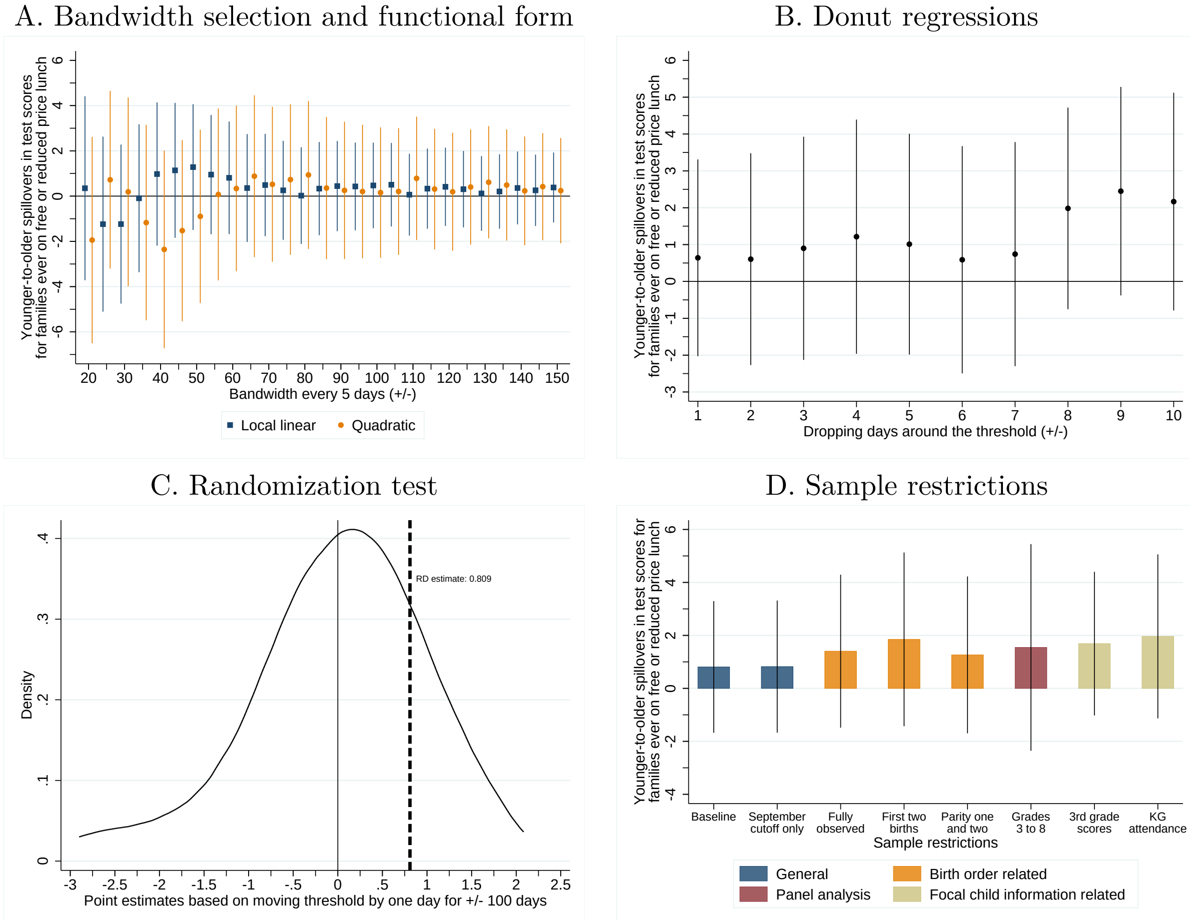
Note: This figure replicates analysis from Figure 2 for the specification from column (9) in panel A of Table 3.

Figure A9: Robustness of the main result: 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 A10: Robustness of the main result: 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.