Long-term effects of human capital spillovers within families[*]

Facundo Albornoz
University of Nottingham
facundo.albornoz@nottingham.ac.uk

Antonio Cabrales
Universidad Carlos III Madrid
antonio.cabrales@uc3m.es

David Contreras
Universidad Adolfo Ibáñez
david.contreras.g@uai.cl

Richard Upward[†]
University of Nottingham
richard.upward@nottingham.ac.uk

Revision for *Journal of Human Resources*
November 2025

[*]We thank Abhijeet Singh, Gianni De Fraja, Pamela Searle, Sebastián Gallegos, and participants at the 21st Nordic Conference in Development Economics, XXV Applied Economics Meeting, Oxford Development Economics Workshop, 9th Warwick Economics PhD Conference, SECHI 2024 meetings, seminar participants at the Universidad Adolfo Ibáñez and Universität Tübingen, two anonymous referees and the editor for helpful comments and suggestions. We also thank the DEMRE and Ministry of Education (MINEDUC) of the government of Chile for giving us access to the administrative data we use in this paper.

[†]Corresponding author

Abstract

We provide new evidence on the long-term effects of human capital externalities within families. Using comprehensive national administrative data for Chile, covering 15 cohorts of school-leavers, we show that initial differences in human capital of one child affect the educational outcomes of their siblings at age 18. To identify these effects we exploit exogenous variation in school starting age caused by school entry cut-off dates. We find robust evidence that spillovers are greater from younger-to-older siblings, and that these spillovers persist at least until the end of compulsory schooling. Our results suggest that compensating parental investment is an important mechanism for human capital spillovers, and that parents compensate more from the younger to the older child.

JEL Classifications: C21, D13, I21.

Keywords: Early-life shocks, family spillovers, siblings, regression discontinuity

1 Introduction

Externalities in the human capital production function are a potentially important source of inequality, and also provide an opportunity to leverage human capital investment. An obvious place to look for these externalities is within the family, and in particular in the relationship between siblings. Siblings have frequent and long-lasting interactions, and they share family resources. Recent empirical work has provided plausible evidence of spillovers of human capital between siblings, often using exogenous variation in outcomes of one child to estimate causal effects on their sibling. But almost all of this work focuses on spillovers from the older to the younger sibling. And older-to-younger spillovers could plausibly be the result of both role-model effects as well as parental choices and investment decisions. In addition, there is little evidence on whether sibling spillovers persist or diminish over time.

In this paper, we use comprehensive national administrative data for Chile, covering 15 cohorts of school-leavers, to examine how initial human capital advantages of one child affect the long-term educational outcomes of their siblings. We are able to trace out how the shock affects school starting age (SSA) and the child's own human capital development, and then how this affects their sibling throughout their school career. We explore how the spillover effect varies with birth order, gender, the age gap between siblings, the resources which families are able to invest in their children's education, and the time since the school start date. By doing this, we provide new evidence on the way in which externalities affect the human capital production function.

To identify these effects, we exploit exogenous variation in SSA caused by a school entry cut-off date. We link data on college admission test scores to children's exact date of birth and SSA. We complement this dataset with information on school achievement over each child's school career (from grade 4 to 12), and survey data on parental investment.

We find significant, positive and long-lasting spillovers from younger to older siblings which persist at least up to the end of the older siblings' school career. But we find no evidence of spillovers from older to younger siblings. This result is consistent across the timing of the outcome measure and the economic status of the household.

Our results suggest that parental investment responses to children's human capital are an important source of sibling spillovers because explanations based on direct sibling interactions typically imply stronger older-to-younger spillovers. More specifically, our results are consistent with compensatory parental investment: parents allocate more time and resources to the child who is disadvantaged by their school starting age. This hypothesis is supported by the direct survey evidence on parental investment. The fact that parental compensation is greater from younger-to-older further suggests that the marginal return on investment is greater for the older child and that household time constraints evolve with children's age.

This paper connects to three strands of the literature. First, it is related to the recent literature on sibling spillovers and the role of the family on educational outcomes. Most of this literature focuses on spillovers from older to younger siblings, and considers spillovers in educational choices. Joensen and Nielsen (2018), Dustan (2018), Altmejd, Barrios-Fernández, Drlje, Goodman, Hurwitz, Kovac, Mulhern, Neilson and Smith (2021), Aguirre and Matta (2021), Dahl, Rooth and Stenberg (2020) and Barrios-Fernández (2022) all show that older siblings' choices and school assignments increase the probability of their younger siblings making similar choices. Each of these papers exploits some exogenous variation in older sibling choice, such as admission thresholds.

A smaller literature considers spillovers on academic performance, which is the focus of our paper. Table 1 summarises all the existing papers we have identified which estimate sibling spillover effects on academic performance. Only three papers use the same identification strategy (age cut-offs), and only two of those consider younger to older spillovers. Qureshi (2018*a*), Nicoletti and Rabe (2019) and Zang, Tan and Cook (2023) all find evidence of positive older-to-younger spillover effects: better academic

Table 1: Previous evidence of sibling spillovers on academic performance

Authors	Direction of spillover	Setting	Treatment	Outcome	Method	Result
Oettinger (2000)	Both directions	US high school graduates	Sibling graduation outcome	Graduation outcome	2SLS (IV = family background)	Positive older-to-younger effects; zero younger-to-older effects
Qureshi (2018*a*)	Older sisters to younger brothers	India	Older sibling years of education	Younger brother's years of schooling, literacy and numeracy	2SLS (IV = distance to nearest girls school)	Positive older (sister) to younger (brother) effects
Qureshi (2018*b*)	Both directions	United States (NC)	Older sibling allocated to more experienced teacher	Reading test score grades 4–8	School-year fixed effects and lagged test scores	Positive older-to-younger effects; zero younger-to-older effects
Nicoletti and Rabe (2019)	Older to younger	UK	Older sibling test scores at age 16	Test scores at age 16	FE + 2SLS (IV = school test score)	Positive older-to-younger effects
Black et al. (2020)	Younger to older	United States (FL) and Denmark	Health shock to third child	Test scores in elementary and middle school	Within family differences in exposure	Positive younger-to-older effects
Landersø et al. (2020)	Younger to older	Denmark	Sibling SSA	Test scores grade 9	RD using date of birth	Positive younger-to-older effects
Barrios-Fernández (2022)	Older to younger	Chile	Eligibility for student loan	College entrance test score	RD for eligibility	Insignificant older-to-younger effects
Karbownik and Özek (2023)	Both directions	United States (FL)	Sibling SSA	Test scores grades 3–8	RD using date of birth	Positive older-to-younger effects in less affluent families; negative younger-to-older effects in more affluent families
Zang et al. (2023)	Older to younger	United States (NC)	Sibling SSA	Test scores grades 3–8	RD using date of birth	Positive older-to-younger effects, stronger in disadvantaged families

outcomes for the older sibling lead to better academic outcomes for the younger sibling. Barrios-Fernández (2022) is of particular interest because he uses a plausibly exogenous discontinuity in older siblings' college entrance test scores to estimate spillovers on younger siblings' test scores. He finds no significant effect, which we will see is consistent with our own estimates in this paper, using a different exogenous shock.[1]

Younger-to-older spillovers are of interest because they are far more likely to be the result of parental investment decisions. Oettinger (2000) and Qureshi (2018*b*) find no evidence of younger-to-older spillovers, but neither of these papers uses a clearly exogenous shock. In contrast, we can show that the sibling date of birth is plausibly exogenous to potential outcomes.

Black et al. (2020) consider younger-to-older spillovers which occur as a result of a health shock to the third child using data from Florida and Denmark. Exploiting the difference in exposure to the shock which occurs between the first and second child they find evidence of positive younger-to-older spillovers: a negative health shock to the third child reduces test scores for the second child. They suggest that this effect occurs through the mechanism of parental investment: parents have less time and financial resources to invest in the second child if the third child experiences a health shock.

Landersø et al. (2020) use Danish data and exploit the discontinuity in SSA which arises because of minimum school-entry age requirements. Their focus is primarily on the effects of SSA on parental outcomes, but they do also consider sibling spillovers both from older-to-younger and younger-to-older. They find evidence of positive younger to older spillovers. A later SSA for children improves their older siblings' academic outcomes at the end of grade 9. They argue that a higher SSA of younger siblings improves the school performance of the older siblings because parental resources are freed up. This suggests that parents compensate rather than reinforce the advantage of their younger sibling starting school older by investing more in the older sibling.

[1] See Barrios-Fernández (2022, Table 8). He does however find evidence of a positive spillover on the likelihood of taking the college entrance exam.

However, these effects are limited to siblings with large age gaps and are imprecisely estimated and in some cases implausibly large.

Karbownik and Özek (2023) use administrative data on student test scores for 3rd-8th graders for a single school district in Florida. Using the same discontinuity device as Landersø et al., they find positive spillovers from older-to-younger siblings: a child who starts school at an older age has a positive effect on their younger sibling. However, these gains are entirely driven by students from low income backgrounds. They argue that these positive spillovers are the result of direct "role-model" effects, since in low income households parental investment is less likely to compensate or reinforce human capital advantages which arise in one child. In contrast, Karbownik and Özek find small negative spillovers from younger to older siblings for children from higher income backgrounds. They argue that these arise as a result of parental investment reinforcing initial advantages in human capital which arise as a result of variation in SSA.

Our study complements and extends this work in several dimensions. We provide a systematic comparison of older and younger siblings and focus on the effect of SSA on longer-run educational outcomes. We provide the first evidence that SSA has a *smaller* effect on older siblings, and that spillover effects from younger to older siblings are actually *larger* than from older to younger siblings. The size of the own- and spillover-effects are both consistent with the hypothesis that parental investment is a channel by which spillover occur. Second, we explore behavioural channels that might explain the results. We provide evidence that parents do indeed respond to the timing of child's school start, as proposed by Landersø et al. (2020). The differences between our findings and those of Karbownik and Özek (2023) suggest that the size and existence of spillovers within families may be context specific. In that light, we provide the first estimates of spillovers for a middle-income country.

Second, our paper relates to the literature on intra-household resource allocation and its interaction with early life shocks to human capital. A long-running literature considers how parents allocate resources among their offspring to reinforce or compensate initial endowments. In their classic theoretical model of intra-household allocation decisions, Becker and Tomes (1976) argue that parents invest more human capital in better-endowed children, and therefore reinforce endowment differentials among their children. On the other hand, Behrman, Pollak and Taubman (1982) posit that parents face an equity-productivity trade off. If equity is weighted more heavily, parents will compensate for endowment differences. The empirical evidence is mixed. Some studies find evidence that parents reinforce (e.g. Rosenzweig and Zhang, 2009; Datar, Kilburn and Loughran, 2010; Rosales-Rueda, 2014; Frijters, Johnston, Shah and Shields, 2013), while others find that parents compensate (e.g. Bono, Ermisch and Francesconi, 2012; Yi, Heckman, Zhang and Conti, 2015; Bharadwaj, Eberhard and Neilson, 2018).[2] These differences may be explained by differences in the measures of children's endowments, or by differences in the empirical methods. Previous studies have extensively used birth weight as a measure of child endowments (e.g. Royer, 2009; Bono et al., 2012; Yi et al., 2015; Restrepo, 2016; Bharadwaj et al., 2018); whereas measures of other endowments (e.g. cognitive skills) have been less used (e.g. Aizer and Cunha, 2012; Frijters et al., 2013; Sanz-de Galdeano and Terskaya, 2019). Moreover, parental responses might change according to the dimensions of human capital. In that regard, Yi et al. (2015) show that parents simultaneously compensate or reinforce in different dimensions of human capital in response to early health shocks. In particular, they find that in response to early health shocks, parents make compensating investments in child health; whereas they make reinforcing investments in education. Berniell and Estrada (2020) show that college-educated parents make compensating time investments in their children who are younger at school entry. We will see that this result is very consistent with our own findings, but in addition we can show that this compensating behaviour has consequences for the other sibling.

[2] There are also studies that find that parents are neutral, and do not compensate or reinforce endowment differences (e.g. Royer, 2009; Currie and Almond, 2011).

The estimation of parental responses to children's endowments poses several econometric challenges. The main threat to identification is selection bias caused by unobserved characteristics that might drive both parental investments and children's endowments. To deal with this problem, studies have used different strategies, such as family fixed effects (e.g. Rosales-Rueda, 2014), natural experiments (e.g. Aizer and Cunha, 2012) and instrumental variables (e.g. Yi et al., 2015). Relative to these papers, we contribute by providing evidence consistent with parents acting to compensate the well-being of worse-endowed children.[3] For identification, we exploit an arguably exogenous variation caused by school entry laws, where children are randomly assigned to start school based on their date of birth.

Third, our paper also contributes to the extensive literature that examines the relationship between SSA and academic performance. Previous research has shown that students who enter school at an older age score higher on in-school tests (Bedard and Dhuey, 2006; Puhani and Weber, 2007; McEwan and Shapiro, 2008; Nam, 2014; Lubotsky and Kaestner, 2016; Attar and Cohen-Zada, 2018; Dhuey, Figlio, Karbownik and Roth, 2019) and are more likely to attend college (Bedard and Dhuey, 2006; Celhay and Gallegos, 2022).[4] It is well known that SSA may be endogenous, and is likely to correlate with student and family characteristics, parental preferences and child maturity (Attar and Cohen-Zada, 2018; Landersø et al., 2020). To address these concerns some studies exploit discontinuities in SSA due to birth date cut-off rules, comparing children who born just before the enrolment cut-off with those who born just after (McEwan and Shapiro, 2008; Black et al., 2011; Peña, 2017; Celhay and Gallegos, 2022). We

[3]These results reflect those of Bharadwaj et al. (2018) which also use data from Chile. They examine the relationship between health at birth – measured by birth weight – and school outcomes, finding that birth weight increases outcomes in maths and literacy throughout all the school years. They show that parents invest more in children with lower birth weight, which suggests that parental investments are compensatory.

[4]Recent studies have also documented that higher SSA decreases the incidence of crime in youth (Cook and Kang, 2016; Depew and Eren, 2016; Landersø, Nielsen and Simonsen, 2017) and lowers the probability of teenage pregnancy (Black, Devereux and Salvanes, 2011). The literature on job market outcomes is less consistent. Some authors find that students who enter school at an older age have higher wages (Fredriksson and Öckert, 2013; Peña, 2017), while others find no effect of the SSA on earnings (Dobkin and Ferreira, 2010; Black et al., 2011).

provide the first evidence that SSA effects vary by birth order. On the other hand, and consistent with the literature of birth order effects, we show that first-borns do better on several measures of educational attainment than later-borns (Black, Devereux and Salvanes, 2005; Barclay, 2018).

The paper is organised as follows. We describe our data in Section 2, explain how we identify siblings and show that older siblings have systematically better academic outcomes across a range of measures. Section 3 explains our empirical method, which relies upon the fact that differences in dates of birth around school-entry cut-offs create large differences in SSA. We show that compliance with school-entry cut-offs is non-random with respect to birth order and household income. Older siblings and children from higher income households exhibit less compliance, consistent with the hypothesis that parental responses to date of birth begin before children start school, and that these responses are larger for older siblings and children from higher income households. Our main results are described in Section 4. We consistently find significant younger-to-older spillovers but precisely-estimated zero effects for older-to-younger spillovers. Section 5 explains how these findings can be motivated by parents reallocating resources towards the child with the highest marginal return to investment and then provides some direct evidence that parental investment responses are consistent with this proposed mechanism. Section 6 concludes, focusing on a comparison of our results with the limited existing evidence on sibling spillovers and human capital externalities.

2 Data

Our main analysis uses three different datasets. The first contains the enrolment records of the entire student population in Chile who complete their schooling between 2004 and 2019. The data come from administrative records provided by the Ministry of Education, and contain information on students' exact date of birth, the year in which they start school and the municipality where they live.

The second dataset contains results from a national standardised test (SIMCE: *Sistema de Medición de la Calidad de la Educación*) administered by the Ministry of Education to all students in certain grades and years. The test is taken at three different times: 4th grade (9-10 years old), 8th grade (13-14 years old) and 10th grade (15-16 years old). SIMCE is the main tool to measure the quality of education in Chile. We use standardised test scores spanning 2000-2017 in two subjects: Mathematics and Spanish.[5] The SIMCE test also contains information from surveys of parents and students which we can use to directly test whether parental investment is affected by differences in SSA. In 2012-2015 there was a survey of parents of 2nd grade children which asks how much time parents spend on activities related to their children's reading. In 2009, 2011, 2013 and 2014 there was a survey of 8th grade children which asks about parental involvement and parental monitoring.

The third dataset contains a record of students who register for the PSU (*Prueba de Selección Universitaria*) test after graduating from high school. The PSU test is the college entrance exam, and is administered by the Department of Evaluation, Assessment and Educational Records (DEMRE), which is the agency responsible for admission to higher education in Chile. Applicants take two mandatory tests in Mathematics and Literacy, and at least one of two optional tests in Social Science and Natural Science. College admission is based on a weighted average of their PSU test scores, high school grades and GPA ranking.[6] We use PSU exam data for students who graduated from high school between 2004 and 2019. The PSU dataset also provides information on student characteristics — including gender, monthly family income, parents' schooling, family size, parents' work status and health coverage — and high school GPA. Finally, we also observe college enrolment for those who take the PSU test, which is available from 2006 onwards.

[5]We focus on these two subjects because of the availability of SIMCE test score information. The national examination also included social sciences and natural sciences, but these are taken for certain grade-year combinations.

[6]GPA ranking was introduced in 2013 as one of the components of the weighted average for college admissions.

All three sets of test scores (SIMCE, PSU and high school GPA) are standardised to have zero mean and unit standard deviation. SIMCE and PSU test scores are standardised by academic year, while high school GPA is standardised by school and academic year.

The three datasets are linked using a unique student identification number. Our main analysis focuses on the cohorts of students that graduated from high school between 2004 and 2019, and immediately took the PSU test.[7] Our main focus is on PSU test scores because we are interested in longer-run outcomes, and also because this data offers the largest sample size.[8]

We identify siblings using surnames provided by the Ministry of Education. Traditionally, Chile follows the Spanish naming system, in which children have two surnames, the first being the father's first surname followed by the mother's first surname. We define two students as siblings if they share the same pair of surnames in the same order, live in the same municipality and attend the same school at some point during the sample period.[9] For our main analysis, we keep only families with exactly two identified siblings i.e. we drop those who have no identified siblings, or who have multiple identified siblings.[10] Using this strategy, we identify 969,406 sibling pairs (1,938,812 students) who both graduate from high school between 2004 and 2019.

There are two potential sources of error in the identification of sibling pairs. False positives occur if non-siblings share the same surname, school and municipality. False negatives occur if siblings do not share the same surname, school or municipality. To measure the error rate, we use parental ID numbers which are available to us from 2019 onwards. We find that 95% of siblings identified by common surnames within

[7]Between 2004 and 2019, 86% of the students who graduated from high school registered for the PSU immediately after graduation, and 91% of them took PSU test.

[8]Table A1 shows the cohorts and years for which we have information on test scores and parental investment.

[9]We therefore do not lose observations on sibling pairs when the older sibling moves to a new school.

[10]Including families with more than two siblings introduces multiple possible spillover effects. We show in Section 4 that extending our sample to include families with three children does not alter our main findings.

the same school and municipality also share the same pair of parental ID numbers. False positives are therefore unlikely to affect our results or threaten external validity. Figure 1 shows the distribution of age differences between identified siblings pairs. The very small number of pairs with a date of birth difference of less than 9 months (0.9% of pairs) confirms our assumption that surname pairs in the same school and the same municipality of residence are siblings.

FIGURE 1 HERE

However, false negatives are more common. Using the same parental ID data we identify 875,570 siblings. Of these, 57% are also identified as siblings based on the surname-school-municipality match. In Table A2 we compare the characteristics of siblings who have the same parental ID and who are identified as siblings by a surname-school-municipality match with those who are not identified as siblings by a surname-school-municipality match. Differences in characteristics between these two samples are statistically significant, but in most cases extremely small. Siblings who are linked by a surname-school-municipality match are not obviously more advantaged than siblings who are not linked. In fact, they have slightly lower levels of parental education, are slightly less likely to live in the capital and are slightly less likely to attend private school.

For our analysis, we drop pairs (i) which have the same date of birth, and are therefore presumably twins (30,568 pairs), (ii) pairs born less than nine months apart (9,097 pairs), and (iii) pairs whose age difference is greater than 12 years (421 pairs). Our largest estimation sample comprises 260,393 sibling pairs (520,786 students) who both have valid PSU test scores and complete information on socioeconomic characteristics.

Our research design will exploit differences in birth order, and therefore in Table 2 we presents summary statistics separately by older and younger siblings. By definition, siblings share some characteristics, and therefore most of the characteristics in Panel

(a) are identical for older and younger siblings. However, time-varying characteristics such as parents' work status can vary between siblings because it reflects the conditions at the end of grade 12 when students register for the PSU test. For example, fathers of older siblings are more likely to be working than fathers of younger siblings, while the reverse is true for mothers. Panel (b) shows that older siblings achieve significantly better academic outcomes than younger siblings, in terms of school grades, PSU test scores, college enrolment and SIMCE test scores.

TABLE 2 HERE

3 Empirical strategy

The estimation of spillover effects between siblings is challenging because siblings typically share the same household, have common unobserved traits and may experience common shocks. This problem is shared with many other settings in which one wants to estimate the causal effect of social interactions (Manski, 1993; Blume, Brock, Durlauf and Ioannides, 2011; Angrist, 2014). We therefore exploit a potentially exogenous difference in the human capital of one child in each sibling pair (who we call the focal child) and trace out the effects of that difference on their sibling's human capital development. In common with several other papers in the spillover literature, the difference arises because of the application of SSA policies. A substantial literature, briefly referred to in our introduction, has shown that SSA has large effects on children's development and human capital formation. Using focal children born just before and after the cut-off date, we compare test scores of their siblings using a Regression Discontinuity (RD) design.

The Ministry of Education establishes the school entry cut-off date as 1 July. Children who turn 6 before 1 July are required to start school in that year, while those who turn 6 on or after 1 July are required to start school in the following year. However, schools are free to implement other cut-offs before 1 July. Figure 2 plots the mean SSA in grade 1 against date of birth in days relative to 1 July. The jump in SSA on 1 July is approximately 0.5 years, but there are also jumps on 1 June, 1 May and 1 April corresponding to children who were in schools which adopted earlier cut-offs when they started grade 1.[11]

[FIGURE 2 HERE]

Figure 2 suggests that the effect of the cut-offs on SSA can be accurately estimated by a piecewise-linear model:

$$\mathrm{SSA}_i = \alpha_0 + \sum_{k=1}^{4} \alpha_k \mathbf{1}(B_i \geq c_k) + \theta_0 B_i + \sum_{k=1}^{4} \theta_k B_i \cdot \mathbf{1}(B_i \geq c_k) + \nu_i \tag{1}$$

where SSA_i is the SSA of child i; B_i is the day of the year of birth of child i relative to 1 July (so $B_i = 0$ on 1 July); and c_k are the four cut-off points at 1 April ($c_1 = -91$), 1 May ($c_2 = -61$), 1 June ($c_3 = -30$) and 1 July ($c_4 = 0$). The coefficients on α_1–α_4 indicate the relative importance of each of the four cut-offs in determining SSA. Table 3 presents the estimates of Eqn. (1). The results confirm that there is a much larger effect on SSA at 1 July than at the other cut-off points: students born at the end of June start school 0.45 years younger than students born at the start of July. The inclusion of controls, shown in column (2), makes no substantive difference to this finding.

[TABLE 3 HERE]

When we split the sample into younger and older siblings, there is a large and significant difference ($p < 0.01$) in the size of the SSA effect at $B_i = 0$, showing that there is more compliance with the 1 July threshold for younger children.[12] In the final two

[11] It is not possible to accurately identify a single cut-off which applied to each child at the time they started school, because schools could change their cut-off over time. For this reason in our main results we focus on the most commonly used cut-off of 1 July.

[12] Differences in the SSA effect at other values of B_i are insignificant.

columns we show that the SSA effect is almost identical for a subsample of children who are born in June and July. Our preferred and simplest specification is one which just compares June- and July-born children, so it is reassuring to see that the effect on SSA in this subsample is almost identical to that in the full sample. The significant difference in the SSA effect of the discontinuity is our first evidence of differences in parental decision-making with respect to birth order. If it is costly to change children's school start date in response to their date of birth (for example by choosing to send them to a school with a different cut-off date), then reductions in compliance may be interpreted as evidence of additional parental investment.[13]

Given these large and significant discontinuity in SSA at the 1 July cut-off, we can estimate the "own effect" of the cut-off on the focal child using the following reduced form model:

$$Y_i = \beta_0 + \beta_1 \mathbf{1}(B_i \geq 0) + f_i(B_i) + \mathbf{1}(B_i \geq 0) \times g_i(B_i) + \varepsilon_i, \tag{2}$$

Each focal child i in our sample has one sibling labelled j. The "spillover" effect of the cut-off on the sibling can be estimating using:

$$Y_j = \gamma_0 + \gamma_1 \mathbf{1}(B_i \geq 0) + f_i(B_i) + \mathbf{1}(B_i \geq 0) \times g_i(B_i) + \eta_i. \tag{3}$$

Y_i and Y_j are the outcomes of the focal child and their sibling, such as test scores. B_i is the running variable, the day of the year of birth of the focal child normalised to 1 July. The parameters β_1 and γ_1 correspond to the own and spillover effect of starting school at an older age. We focus on these reduced form regressions rather than instrumenting SSA with the school entry cut-off because the assumptions needed to support a local average treatment effect (LATE) are unlikely to hold. In particular, the monotonicity assumption

[13] Figure B1 shows that the probability of starting school in the year in which a student turns 7 jumps at the cut-off. In addition, Table A3 shows that redshirting is more common for high income families, for families with higher levels of parental education and for children in private school. We take this as additional evidence of redshirting as an investment decision by parents to compensate for being born before the cut-off.

is violated because being born in June reduces SSA for some children, but increases it for others. This arises because June-born children can be held back and start school in the following school year ("redshirted"), as discussed in Barua and Lang (2016). Celhay and Gallegos (2022, Fig C4) show a very similar pattern in school starting age using the same data.[14]

We concentrate on the effect of the 1 July cut-off because, as shown above, the effect on SSA is largest at this point. In our preferred, and simplest, specification, we restrict B_i to lie within 30 days of the 1 July cut-off. In this case, the dummy variable $\mathbf{1}(B_i \geq 0)$ is simply a dummy for July born. To estimate Eqn. (2) we restrict the sample to those children born in June or July, and to estimate Eqn. (3) we restrict the sample to those children whose siblings are born in June or July. In this specification we treat $f(B)$ and $g(B)$ as linear functions whose slope can vary either side of the cut-off. Both Eqns. (2) and (3) are estimated separately for older and younger siblings, giving us estimates β_1^O and β_1^Y for the own effects and γ_1^O (younger-to-older) and γ_1^Y (older-to-younger) for the spillover effects.

For robustness, we also consider three alternative approaches. First, we use the whole range of B_i and treat f and g as piecewise linear functions (as in Eqn. (1)). Second, we implement the methodology proposed by Calonico, Cattaneo and Titiunik (2014) in which the bandwidth is data-driven and f and g are estimated non-parametrically. Third, we follow Cattaneo, Idrobo and Titiunik (2024) and treat the running variable as discrete rather than continuous. To do this, we collapse the data down to one observation per mass-point (day of birth) and again estimate f and g non-parametrically with a data-driven bandwidth.

The RD design will produce unbiased estimates of a causal effect only if there is no manipulation of the running variable around the threshold. To empirically test the validity of the RD design, we consider two falsification tests. First, we examine whether

[14] In our data, "redshirting" is common amongst June-born children, as shown in Table A3.

students are selected into treatment via manipulation of their date of birth, causing a jump in the density after the cut-off date. Figure 3 shows a kernel density estimate of the date of birth for younger (Panel a) and older siblings (Panel b). The graphical evidence suggests that the empirical density is continuous across the threshold. We also implement a formal test suggested by Cattaneo, Jansson and Ma (2018), with the null hypothesis that the there is no jump in the density in 1 July. In the case of younger siblings the p-value is 0.40, whereas in the case of older siblings the p-value is 0.19. Therefore there is no evidence of manipulation of the running variable at the cut-off.[15]

[FIGURE 3 HERE]

Second, we test whether the observable characteristics of children are balanced on the two sides of the school-entry cut-off. If there is a non-random sorting around assignment cut-off, pre-determined characteristics of children on one side of the threshold may differ from those on the other side. In order to implement the analysis, we estimate reduced form regressions like Eqn. (2) and Eqn. (3) where Y_i and Y_j are replaced with a series of child and household characteristics.[16] We report the results in Table 4.

[TABLE 4 HERE]

In the first two rows of each panel we estimate whether the July born dummy and age differ at the discontinuity. The sibling spillover effect (γ_1) for both variables is a precisely estimated zero. This means that any spillover effect we estimate is not an age at test effect, nor the effect of differences in physical strength or mental characteristics.[17] In addition, this means that the cut-off is not correlated across siblings, and hence B_j is not an omitted variable in Eqn. (3). There are also no significant differences across the school-entry cut-off for any of the other characteristics.

[15]This is consistent with Dickert-Conlin and Elder (2010), who find no evidence that school cut-off dates influence timing of births.

[16]Strictly speaking, only birth date, sex and parents' schooling are determined before the outcome because the other characteristics are time-varying and are recorded at the end of grade 12.

[17]In contrast, the own effect (β_1) may be the result of differences in age at testing, physical maturity etc.

4 Results

Before reporting estimates of the spillover effect, we first demonstrate that the discontinuity has the expected own effect on the focal child. Figure 4 illustrates the discontinuity for the earliest test score outcome available to us (grade 4).

[FIGURE 4 HERE]

As expected, there is a significant jump in test scores at the cut-off. The effect is larger for younger siblings (left hand panel) compared to older siblings (right hand panel). As noted in our introduction, this finding is consistent with an extensive literature on the relationship between SSA and academic performance. The jump in test scores may also reflect differences in age at test between June and July-born children. For our purposes, however, we are agnostic as to the mechanism which causes the jump in the focal child's test scores at the discontinuity. To measure spillovers, we simply require variation in the human capital of the focal child which is exogenous to the potential outcomes of their sibling.

In Table 5 we show that the own effect of July birth on test scores persists, but declines with age. The decline in the size of the own effect across grades is strikingly similar to that reported by Dhuey et al. (2019, Figure 1), and seems likely to arise from the fact that the age-at-test effect declines with age: the relative difference in age between June and July-born children declines with age. In each grade, the effect of being born after the cut-off is larger for younger siblings.[18]

[TABLE 5 HERE]

Having established that there is a significant own effect of date of birth on academic outcomes, we now turn to our main research question: whether these differences in the

[18]The difference between older and younger siblings across all grades is -0.47σ with a standard deviation of 0.017.

human capital of the focal child i have a long-run spillover effect on their sibling j. We have shown that children whose siblings are born either side of the school-year cut-off are observably identical (Table 4), so our claim is that these estimates represent the causal effect of differences in sibling human capital on academic outcomes. Figure 5 illustrates the discontinuity for PSU test scores, measured in grade 12 and for which we have the largest sample.[19] There is a significant jump in test scores for older siblings when their younger sibling is born in July, but no effect on test scores for younger siblings when their older sibling is born in July.

[FIGURE 5 HERE]

We explore the robustness of the spillover effect on PSU test scores in Table 6. The addition of controls (column 2), and the use of a piecewise linear model (column 3) which uses all the cut-offs make very little difference to the result. We also implement the robust data-driven RD methodology proposed by Calonico et al. (2014) which which estimates $f(B_i)$ as a local polynomial (column 4).[20] The optimal bandwidth either side of the 1 July cut-off is somewhat less than 30 days for younger siblings, and slightly more for older siblings, but the estimated effect size is quite similar. In column 5 we treat the running variable as discrete (one observation per day of birth) as recommended by Cattaneo et al. (2024). The estimated spillover effect of July birth on test scores is again very similar.[21] Younger-to-older spillovers are consistently positive and about 0.06σ, while older-to-younger spillovers are mostly negative, much smaller, and insignificantly

[19] In Table A4 we investigate whether sibling human capital might also affect selection into taking the PSU test, either through high school completion or by opting out of the PSU test. Results are shown in Table A4. There is no effect of sibling July birth on the probability of high school completion or taking the PSU for older siblings. There is a very small positive effect of in the older-to-younger regression. Our test score results are therefore not primarily the result of sample selection.

[20] The model includes indicator variables for the other cut-offs. We do not use the multiple cut-off model proposed by Cattaneo, Titiunik and Vazquez-Bare (2020) because we do not know which cut-off applies to each child — schools did not use a consistent cut-off over time, and it was therefore not possible to establish a reliable school-based rule.

[21] We have also explored whether the estimated standard errors are sensitive to the use of clustering, as suggested by Kolesár and Rothe (2018). Table A5 shows that standard errors are very similar regardless of the clustering chosen. In Figure B2 we re-estimate the spillover effect using a "donut-RD" model which removes observations within a small radius of the 30 June-1 July cut-off. Our estimates of both younger-to-older and older-to-younger spillovers are very stable as we increase the size of the radius.

different from zero. A positive shock to the human capital of the younger sibling has a positive effect on the older sibling, and the size of the spillover is a significant fraction of the size of the own effect shown in Table 5. But, a positive shock to the human capital of the older sibling does not have a positive effect on the younger sibling.

[TABLE 6 HERE]

In order to have a sense for the magnitude of the spillover effect, note that according to Correa, Gutiérrez, Lorca, Morales and Parro (2019), a one standard deviation (1 SD) increase in PSU test scores is associated with a 0.342 SD increase in the monthly income of students during their first four years on the labour market. In their sample, the average monthly wage is USD 822, with a standard deviation of USD 600. Using these results, our younger-to-older effect of 0.056 SD translates into a 0.019 SD increase in monthly income, equivalent to a 1.4% increase relative to the average income.

As explained in Section 2, our sample uses families with exactly two siblings, which simplifies the analysis of older-to-younger and younger-to-older siblings. In Table A7 we also report estimates of own and spillover effects when we include families with two or three children. All of the possible older-to-younger spillover effects are close to zero and insignificantly different from zero, and all of the younger-to-older spillovers are positive and significant with the exception of spillovers from third-to-second oldest. The pooled spillover effect from older-to-younger siblings is 0.007σ (0.012) and not statistically significant, while the pooled spillover effect from younger-to-older siblings is 0.052σ (0.011) and significant at the 1% level. Overall, extending our sample to include families with three children does not alter our main findings that spillover effects from younger to older children are significant and positive.

We now consider how the estimated spillover effect changes over the course of the

We also conduct three placebo tests in Table A6, comparing children born in July/August (panel a), August/September (panel b) and September/October (panel c). All estimates are insignificant with one exception in panel (b). There is a marginally significant placebo effect for older-to-younger spillovers when comparing July and August (p-value $= 0.095$).

sibling's school career. Table 7 reports estimates of γ_1 for test scores from grade 4 up to grade 12. We find positive younger-to-older spillover effects on test scores from 8th to 12th grade, but no positive older-to-younger spillover effects for any grade. We do not find a younger to older spillover effect for high school GPA.[22] Importantly, the younger-to-older spillover effect does not appear to decline with age in the same way as the own effect shown in Table 5. This is presumably because, although the proportional age effect of the cut-off on the focal child varies reduces with age, in the spillover model the age of child j does not vary systematically between those whose sibling i (the focal child) is born in June and those whose sibling is born in July.

[TABLE 7 HERE]

We can also examine whether spillovers persist up to college enrolment and graduation outcomes. College enrolment data is available from the Ministry of Education from 2007 to 2024, allowing us to track enrolment for students who completed high school in 2006 and onward. We also identify a group of selective institutions: the 30 universities that comprise the Council of Rectors of Chilean Universities (*Consejo de Rectores de las Universidades Chilenas*, CRUCH). These institutions are among the most established and traditionally prestigious in the country. They receive direct public funding, are research-oriented, and participate in the centralized admissions system. Moreover, they tend to enrol students with the highest admission scores. We measure both immediate enrolment and enrolment within five years of high school completion. Graduation data is available from 2007 to 2023. We measure graduation from any post-secondary institution up to six years after high school completion for 14 cohorts of 12th-grade students, spanning from 2004 to 2017. In Chile, a university degree typically takes 4 to 6 years to complete, depending on the field of study, with programs like engineering and medicine often requiring 6 to 7 years. We also measure 10-year graduation. Due to

[22] However, note that the high school GPA is not a national anonymised test score, and therefore may be affected by teacher grading effects. Contreras (2024) provides evidence on teacher grading biases using similar data.

data availability, this outcome is measured for only 10 cohorts of high school graduates, spanning from 2004 to 2013.

Table 8 shows the estimated effects on higher education outcomes. Spillover effects from older to younger are always very small and insignificantly different from zero. Spillover effects from younger to older are generally larger. Unsurprisingly, because PSU test scores are higher we also find positive effects on College enrolment and (to some extent) on graduation rates. However, we do not find any effect on selective enrolment.

[TABLE 8 HERE]

We have also investigated how these spillover effects differ between siblings of the same and different sex, by the age gap between siblings, and by family income. We explore these questions in Table 9.

[TABLE 9 HERE]

Our pooled estimate for the younger-to-older spillover effect (column 1, row 1) is 0.056 and significantly different from zero. This does not vary significantly between siblings of the same or different sex (rows 2 and 3). When we further divide by sex as well as sex parity (rows 4–7) we cannot reject the null that the spillover effect is equal. Thus, the younger-to-older spillover is quite consistent by sex and sex parity.

The overall older-to-younger spillover effect reported in Table 9 (column 2, row 1) is small and insignificantly different from zero. However, this overall effect disguises a significant difference between siblings of the same sex (row 2) and siblings of the different sex (row 3). When we further divide by sex as well as sex parity, the positive older-to-younger spillover is found only for siblings where the older is male and the younger is female (row 7).

In rows (8) and (9) we examine whether spillovers are stronger when the siblings are closer in age. Older to younger spillovers are close to zero regardless of the age gap. Younger to older spillovers are larger for siblings who are close in age, but the difference between the two groups is not itself significant.

Finally, in rows (10) and (11) we compare spillover effects between high and low income families. Karbownik and Özek (2023) find evidence of positive spillovers from older-to-younger in less affluent households, and *negative* spillovers from younger-to-older in more affluent households. We do not find evidence to support this from the results in Table 9. Older to younger spillovers are zero in both higher and lower income families, and higher income households have larger positive spillovers from younger to older (although the difference is not itself statistically significant).

A possible mechanical explanation for our finding that younger-to-older spillovers are larger (and significant) is that the time from the school start date of the focal child to the test score outcome of their sibling is smaller when the focal child is the younger sibling.[23] If parental investment compensates for differences in human capital which arise because the focal child experiences a human capital advantage, then we might expect that spillovers will diminish with time. We can examine this by stacking all SIMCE test scores in grades 4, 8, 10 and the PSU test in grade 12 into a panel which includes a measure of the number of years since the focal child's school start date. This panel contains variation in the age gap between the focal child and their sibling, and variation in the timing of test scores, so it allows us to explore the relationship between the time between the focal child school start date and their sibling's outcome by estimating Eqn. (3) separately by the distance between the focal child school start date and the test score.

The results of this exercise are shown in Figure 6. This shows that the difference in

[23]This is illustrated in Figure B3. Suppose that the older sibling starts school in year t. Then the younger sibling takes the PSU test in year $t+a+12$ where a is the sibling age gap. Now suppose that the younger sibling starts school in year t, then the older sibling takes the PSU test in year $t-a+12$.

the spillover effect between older and younger siblings is not due to the fact that older sibling's test outcomes occur nearer to the focal child's school start date. In panel (a), we see that the older-to-younger spillover effect is always insignificantly different from zero regardless of the distance from the focal child school start date to the sibling's test score outcome. And in panel (b), we see that the younger-to-older spillover effect does not diminish with the distance from the focal child school start date. Panel (b) also provides a placebo test, because it includes some test score outcomes which occur *before* the focal child school start date. There is no estimated spillover effect from younger-to-older before the focal child starts school, and the spillover effect jumps as soon as the focal child starts school.

[FIGURE 6 HERE]

However, the bands in Figure 6 may still disguise significant differences in time since focal child school start date between older and younger siblings.[24] In Table 10 we therefore report estimates of the spillover effect after controlling for a full set of "time since focal child school start date" dummies. In columns (1) and (2) we report estimates of the average spillover effect for all SIMCE and PSU test scores, which shows a small and insignificant older-to-younger effect, and a much larger and significant younger-to-older effect. Exactly the same results can be achieved by interacting the July dummy with a dummy for whether the sibling is older (identifying younger-to-older spillovers), giving the equivalent result in column (3). In column (4) we add a complete set of dummies for time since the focal child's school start date. This has very little effect on the size or significance of the additional younger-to-older spillover effect, confirming that the larger spillover effect from younger to older is not driven by the timing effect.

[TABLE 10 HERE]

[24] Figure B4 shows that the distribution of the time since sibling school start date varies significantly by birth order.

To conclude our results, we consistently find that younger-to-older spillovers are positive while older-to-younger spillovers are close to zero. Younger-to-older spillovers appear to persist at least until the end of compulsory schooling, are consistently positive across sex, age gap and income, and are not the mechanical result of timing differences. This finding is surprising, because we expect role-model effects to be more important for older-to-younger sibling relationships. Indeed, as noted in the introduction, the literature is dominated by studies of older-to-younger spillovers. Landersø et al. (2020, Table 3) provide some evidence for positive academic spillovers from younger to older, but their results are imprecise and mostly insignificant. In contrast to our findings, Karbownik and Özek (2023) find evidence of positive spillovers from older-to-younger in less affluent households, and *negative* spillovers from younger-to-older in more affluent households. In Section 5, we suggest that parental investment is a mechanism to explain this result and provide some direct evidence that parental investment does vary in ways which are consistent with this mechanism.

5 Potential mechanisms: Sibling Spillovers and Parental Investment Responses

We start with the general observation that sibling spillovers may arise through direct interactions, through parental investment responses, or through a combination of both mechanisms (as noted by Karbownik and Özek, 2023). Direct interactions may include role-model effects, motivation, information sharing, and study support. The prevailing view on the literature is that these direct effects are typically positive and stronger from older to younger siblings. Indeed, younger-to-older spillovers are often used as a placebo test in empirical work on sibling educational preferences.

Our main finding challenges this emphasis on direct interactions. By documenting younger-to-older spillovers, we argue that the most plausible mechanism involves parental investment responses rather than direct sibling effects.[25] Explanations based on role-model effects or motivation imply stronger older-to-younger spillovers, as younger siblings are unlikely to influence their older siblings through direct interaction — they are neither salient role models nor key sources of academic information. Hence, explaining our findings requires a form of indirect interaction, which we posit operates primarily through parental involvement (Doepke and Zilibotti, 2017; Albornoz, Berlinski and Cabrales, 2018).

More specifically, our results are consistent with compensatory parental investment: parents appear to allocate more time and resources to the child who is temporarily disadvantaged. A child born in July — typically among the strongest in her class — may therefore free up parental time to be redirected toward her siblings. To substantiate this mechanism, we examine direct measures of parental investment. The survey of parents of 2nd-grade children includes questions about how frequently they engage in specific educational activities with their childreading to or with them, discussing their reading, giving books, and visiting libraries. Each activity is coded on a 1–5 scale (1 = "never," 5 = "very often"). The 8th-grade survey captures parental involvement and monitoring through questions such as: "How often do your parents congratulate you for your grades?"; "Do your parents know your grades?"; and "Are your parents willing to help you when you struggle in school?" Responses are coded from 1 ("always") to 4 ("never"). We construct an aggregate parental-investment index by summing these items and standardising all measures to have mean zero and unit standard deviation.[26]

Table 11 shows the RD estimates of the effect of July birth on these parental investments — reported by parents in grade 2 and by students in grade 8.[27] At this point,

[25]The spillover effect from younger to older could be the result of parental school choice decisions (which could itself be regarded as a kind of parental investment). Table A8 in the Appendix shows there is a small positive effect of having a July-born younger sibling on the overall school transition rate, but no evidence that families switch to private schools. We do not therefore consider it likely that school choice is a primary explanation for the human capital spillover we observe.

[26]Table A9 reports descriptive statistics by birth order and household income.

[27]We cannot report PSU test scores for these samples because, at the time of writing, these children have not reached that point in their schooling.

we are asking a lot of the data partly because the available sample sizes on parental investments are far smaller than the available sample sizes on test scores. Nevertheless, across all direct measures, parental investment is *smaller for children born in July*. All eight estimates in panel (a) are negative, and three are significantly negative. If wealthier parents face lower marginal costs of helping their children (due to greater educational experience) or derive higher marginal utility (because they better internalise the returns to education), the compensatory-investment effect should be *stronger in richer households*.[28] Consistent with this, Table 11 shows that the negative effect on parental investment is indeed larger among high-income families. The larger spillover effect on test scores (Table 9) for richer families also points to stronger compensatory behaviour. In addition, Table A3 shows that richer families are more likely to delay school entry. We interpret this as further evidence of greater parental compensatory investment for June-born children in high income families. However, we do not find strong evidence that this compensatory parental investment translates into positive investment spillovers for the sibling, shown in panel (b) of Table 11. We do not have the power to reject the null of zero spillovers, but nor can we reject our hypothesis that the falls in investment for the focal child cause increases in investment for the sibling. In particular, we are hampered by the fact that younger-to-older spillovers rely on the small sample of families who have a younger child at school and an older child in grade 2.

[TABLE 11 HERE]

These patterns suggest that parental investment is a plausible mechanism, but they do not yet explain the *direction* of the effect by birth order. Why does a shock to the younger sibling translate into greater investment in the older one, rather than the reverse? We propose two complementary explanations.

First, the marginal return to parental investment may differ by birth order. Older children may yield higher returns — either because parents place a higher value on their

[28]The descriptive statistics on parental investment shown in Table A9 shows that in general richer families have more parental investment.

progress or because helping them has a lower opportunity cost when the younger sibling is still tackling simpler material and is farther from high-stakes exams.[29] Under this logic, a positive shock to the younger child releases parental time that is reallocated toward the older sibling, for whom, given birth order, the marginal returns to parental time are relatively higher than for younger siblings. Whereas a positive shock to the older child releases less parental time because diverting time to the younger yields smaller payoffs.

Second, household time constraints may evolve with schooling stages. When the first child starts school, parents can focus all attention on them. Once the younger child enters school, parents must divide their attention between both. A positive shock to the younger child at this stage relaxes parental time constraints, enabling a visible reallocation toward the older child. Landersø et al. (2020) provides related evidence that delaying school entry for a younger sibling allows parents to redirect resources toward older children approaching exams. Additional evidence that parental time constraints may play a role comes from the results in Table 9 which show that younger to older spillovers are larger for siblings who are closer in age. If household time constraints are tighter for siblings who are close in age, a shock to the younger child has a larger spillover effect on their older sibling.

Taken together, these mechanisms imply that parents reallocate resources toward the child with the highest marginal return to investment. Because older siblings are closer to critical educational outcomes and less resource-intensive overall, they become the main beneficiaries when a younger sibling receives a positive shock. This reasoning explains the asymmetry we observe: spillovers flow from younger to older siblings, but not the other way around.

[29] Jensen and Jorgensen-Wells (2025), in a meta-analysis of 30 studies covering roughly 19,000 participants, find systematic favouritism toward firstborns. Similarly, Price (2008) shows that firstborns receive about 20–30 minutes more "quality time" per day than second-borns of the same age. Our own descriptive evidence of parental investment (Table A9) shows that older siblings receive more investment than younger siblings at the same age.

6 Conclusions

We provide the first evidence that initial advantage in children's human capital, identified by exogenous differences in SSA, can have long-run consequences for their siblings. However, the spillover in human capital is asymmetric. Younger siblings born just after the school enrollment cut-off have a 0.1σ advantage in age 18 test scores, and this transmits to a 0.06σ advantage in their older sibling in the same test scores. In contrast, older siblings born just after the cut-off have a smaller 0.04σ advantage and there is no spillover effect on their younger sibling. We rule out the possibility that the asymmetry arises because the older sibling's outcome is measured closer to the focal child's school entry date, and we verify that the spillover effect is consistent across a range of test scores taken at different points in time.

The asymmetry in the spillover provides evidence that parental investment is a key channel by which human capital spillovers occur within families. Since we expect role-model (or sibling interaction) effects to be greater for older-to-younger human capital transmission, our finding of larger younger-to-older spillovers implies a parental investment response which differs by birth order. Our results support the idea that families compensate rather than reinforce initial differences in human capital. The size of the spillover varies by birth order in a direction which is consistent with the hypothesis that parents reallocate resources toward the child with the highest marginal return to investment. Because older siblings are closer to critical educational outcomes and less resource-intensive overall, they become the main beneficiaries when a younger sibling receives a positive shock.

We are also able to exploit rich information from surveys of parents and students conducted at two points in time during the siblings' school career. These surveys provide a direct measure of parental investment. Positive differences in initial human capital have a consistent negative effect on investment in that child, and a consistent positive (albeit imprecisely estimated) effect on investment in the sibling. These results again provide evidence in favour of a compensation strategy across siblings.

An important question for future research is to understand why human capital spillovers vary across different settings. Our results are consistent with the suggestive evidence provided by Landersø et al. (2020) for Denmark, but contrast with Karbownik and Özek (2023) for a school district in the US, who find that parental investment is reinforcing rather than compensating for initial human capital advantages, and only in higher income families, and that role-model effects matter more in low income families. Our paper provides the first longer-run evidence for a middle income country and the results suggest that compensation occurs across the income distribution and that parental investment outweighs role-model effects.

References

Aguirre, J. and Matta, J. (2021), 'Walking in your footsteps: Sibling spillovers in higher education choices', *Economics of Education Review* **80**, 102062.

Aizer, A. and Cunha, F. (2012), The production of human capital: Endowments, investments and fertility, Working Paper 18429, National Bureau of Economic Research.

Albornoz, F., Berlinski, S. and Cabrales, A. (2018), 'Motivation, resources, and the organization of the school system', *Journal of the European Economic Association* **16**(1), 199–231.

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., Mulhern, C., Neilson, C. and Smith, J. (2021), 'O brother, where start thou? Sibling spillovers on college and major choice in four countries', *The Quarterly Journal of Economics* .

Angrist, J. D. (2014), 'The perils of peer effects', *Labour Economics* **30**, 98–108.

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

Barclay, K. J. (2018), 'The birth order paradox: Sibling differences in educational attainment', *Research in Social Stratification and Mobility* **54**, 56–65.

Barrios-Fernández, A. (2022), 'Neighbors' effects on university enrollment', *American Economic Journal: Applied Economics* **14**(3), 30–60.

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

Becker, G. S. and Tomes, N. (1976), 'Child endowments and the quantity and quality of children', *Journal of Political Economy* **84**(4), S143–S162.

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

Behrman, J. R., Pollak, R. A. and Taubman, P. (1982), 'Parental preferences and provision for progeny', *Journal of Political Economy* **90**(1), 52–73.

Berniell, I. and Estrada, R. (2020), 'Poor little children: The socioeconomic gap in parental responses to school disadvantage', *Labour Economics* **66**, 101879.

Bharadwaj, P., Eberhard, J. P. and Neilson, C. A. (2018), 'Health at birth, parental investments, and academic outcomes', *Journal of Labor Economics* **36**(2), 349–394.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., Roth, J. and Simonsen, M. (2020), 'Sibling spillovers', *The Economic Journal* .

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005), 'The more the merrier? The effect of family size and birth order on children's education', *The Quarterly Journal of Economics* **120**(2), 669–700.

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

Blume, L. E., Brock, W. A., Durlauf, S. N. and Ioannides, Y. M. (2011), Identification of social interactions, *in* J. Benhabib, A. Bisin and M. O. Jackson, eds, 'Handbook of Social Economics', Vol. 1, North-Holland, chapter 18, pp. 853–964.

Bono, E. D., Ermisch, J. and Francesconi, M. (2012), 'Intrafamily resource allocations: A dynamic structural model of birth weight', *Journal of Labor Economics* **30**(3), 65–706.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2014), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2024), *A practical introduction to regression discontinuity designs: Extensions*, Cambridge University Press.

Cattaneo, M. D., Jansson, M. and Ma, X. (2018), 'Manipulation testing based on density discontinuity', *The Stata Journal* **18**(1), 234–261.

Cattaneo, M. D., Titiunik, R. and Vazquez-Bare, G. (2020), 'Analysis of regression-discontinuity designs with multiple cutoffs or multiple scores', *The Stata Journal* **20**(4), 866–891.

Celhay, P. and Gallegos, S. (2022), 'Early skill effects on parental beliefs, investments and children long-run outcomes', *Journal of Human Resources* .

Contreras, D. (2024), 'Gender differences in grading: teacher bias or student behaviour?', *Education Economics* **32**(6), 762–785.

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

Correa, J. A., Gutiérrez, P., Lorca, M., Morales, R. and Parro, F. (2019), 'The persistent effect of socioeconomic status on education and labor market outcomes', *Applied Economic Analysis* **27**(79), 62–90.

Currie, J. and Almond, D. (2011), Human capital development before age five, *in* D. Card and O. Ashenfelter, eds, 'Handbook of Labor Economics', Vol. 4, Elsevier, pp. 1315–1486.

Dahl, G. B., Rooth, D.-O. and Stenberg, A. (2020), Family spillovers in field of study, Working Paper 27618, National Bureau of Economic Research.

Datar, A., Kilburn, M. and Loughran, D. (2010), 'Endowments and parental investments in infancy and early childhood', *Demography* **47**(1), 145–162.

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

Dhuey, E., Figlio, D., Karbownik, K. and Roth, J. (2019), 'School starting age and cognitive development', *Journal of Policy Analysis and Management* **38**(3), 538–578.

Dickert-Conlin, S. and Elder, T. (2010), 'Suburban legend: School cutoff dates and the timing of births', *Economics of Education Review* **29**(5), 826–841.

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

Doepke, M. and Zilibotti, F. (2017), 'Parenting with style: Altruism and paternalism in intergenerational preference transmission', *Econometrica* **85**(5), 1331–1371.

Dustan, A. (2018), 'Family networks and school choice', *Journal of Development Economics* **134**, 372 – 391.

Fredriksson, P. and Öckert, B. (2013), 'Life-cycle effects of age at school start', *The Economic Journal* **124**(579), 977–1004.

Frijters, P., Johnston, D. W., Shah, M. and Shields, M. A. (2013), 'Intrahousehold resource allocation: Do parents reduce or reinforce child ability gaps?', *Demography* **50**(6), 2187–2208.

Jensen, A. C. and Jorgensen-Wells, M. A. (2025), 'Parents favor daughters: A meta-analysis of gender and other predictors of parental differential treatment', *Psychological Bulletin* **151**(1), 33–47.

Joensen, J. S. and Nielsen, H. S. (2018), 'Spillovers in education choice', *Journal of Public Economics* **157**, 158 – 183.

Karbownik, K. and Özek, U. (2023), 'Setting a good example? Examining sibling spillovers in educational achievement using a regression discontinuity design', *Journal of Human Resources* **58**(5), 1567–1607.

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

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

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

Lubotsky, D. and Kaestner, R. (2016), 'Do 'skills beget skills'? evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood', *Economics of Education Review* **53**, 194–206.

Manski, C. (1993), ‘Identification of endogenous social effects: The reflection problem’, *Review of Economic Studies* **60**(3), 531–542.

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

Nam, K. (2014), ‘Until when does the effect of age on academic achievement persist? evidence from Korean data’, *Economics of Education Review* **40**, 106–122.

Nicoletti, C. and Rabe, B. (2019), ‘Sibling spillover effects in school achievement’, *Journal of Applied Econometrics* **34**(4), 482–501.

Oettinger, G. S. (2000), ‘Sibling similarity in high school graduation outcomes: Causal interdependency or unobserved heterogeneity?’, *Southern Economic Journal* **66**(3), 631–648.

Peña, P. A. (2017), ‘Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood’, *Economics of Education Review* **56**, 152–176.

Price, J. (2008), ‘Parent–child quality time: Does birth order matter?’, *Journal of Human Resources* **43**(1), 240–265.

Puhani, P. and Weber, A. (2007), ‘Does the early bird catch the worm?’, *Empirical Economics* **32**, 359–386.

Qureshi, J. A. (2018*a*), ‘Additional returns to investing in girls’ education: Impact on younger sibling human capital’, *The Economic Journal* **128**(616), 3285–3319.

Qureshi, J. A. (2018*b*), ‘Siblings, teachers, and spillovers on academic achievement’, *Journal of Human Resources* **53**(1), 272–297.

Restrepo, B. J. (2016), 'Parental investment responses to a low birth weight outcome: who compensates and who reinforces?', *Journal of Population Economics* **29**(4), 969–989.

Rosales-Rueda, M. F. (2014), 'Family investment responses to childhood health conditions: Intrafamily allocation of resources', *Journal of Health Economics* **37**, 41–57.

Rosenzweig, M. R. and Zhang, J. (2009), 'Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy', *The Review of Economic Studies* **76**(3), 1149–1174.

Royer, H. (2009), 'Separated at girth: US twin estimates of the effects of birth weight', *American Economic Journal: Applied Economics* **1**(1), 49–85.

Sanz-de Galdeano, A. and Terskaya, A. (2019), Sibling differences in educational polygenic scores: How do parents react?, IZA Discussion Papers 12375, Institute of Labor Economics (IZA).

Yi, J., Heckman, J. J., Zhang, J. and Conti, G. (2015), 'Early health shocks, intra-household resource allocation and child outcomes', *The Economic Journal* **125**(588), F347–F371.

Zang, E., Tan, P. L. and Cook, P. J. (2023), 'Sibling spillovers: Having an academically successful older sibling may be more important for children in disadvantaged families', *American Journal of Sociology* **128**(5), 1529–1571.

Tables

Table 2: Sample means by birth order

	Younger siblings		Older siblings		Difference	Std. err.
	Mean	Obs.	Mean	Obs.		
(a) Student characteristics						
1=Female	0.52	260,393	0.52	260,393	0.002	(0.001)
Age difference	3.58	260,393	3.58	260,393	0.000	(0.006)
1=Same-sex sibling	0.54	260,393	0.54	260,393	0.000	(0.001)
1=Lives in the capital	0.42	260,393	0.42	260,393	0.000	(0.001)
Monthly family income	563.05	260,393	563.05	260,393	0.000	(1.412)
Mother's schooling	12.44	260,393	12.44	260,393	0.000	(0.009)
Father's schooling	12.51	260,393	12.51	260,393	0.000	(0.010)
Family size	4.53	260,393	4.53	260,393	0.000	(0.003)
1=Father employed	0.82	166,174	0.85	227,281	-0.031^{***}	(0.001)
1=Mother employed	0.47	166,174	0.45	227,281	0.017^{***}	(0.002)
1=Public health coverage	0.65	166,174	0.64	227,281	0.014^{***}	(0.002)
(b) Academic performance						
High school GPA	0.08	260,393	0.13	260,393	-0.046^{***}	(0.003)
PSU: Math score	0.15	260,393	0.17	260,393	-0.028^{***}	(0.003)
PSU: Language score	0.07	260,393	0.13	260,393	-0.058^{***}	(0.003)
1=College enrolment	0.44	258,274	0.47	230,397	-0.026^{***}	(0.001)
SIMCE 4th grade	0.45	161,177	0.51	54,145	-0.062^{***}	(0.005)
SIMCE 8th grade	0.43	129,958	0.53	94,954	-0.100^{***}	(0.004)
SIMCE 10th grade	0.54	161,859	0.59	126,450	-0.051^{***}	(0.004)

Notes: The sample comprises siblings for whom their own and their siblings PSU test score is available (260,393 sibling pairs). This sample creates imbalance in the number of observations for other characteristics and outcomes because those characteristics are not available in every year. The information on whether parents are employed (*Father employed* and *Mother employed*) and health coverage status (*Public health coverage*) is only available for 393,794 students (76% of the sample). All academic performance measures apart from college enrolment are standardised.

Table 3: Discontinuities in SSA at different birth date cut-offs.

	Full sample		Full sample		Born in June or July	
	No controls	Controls	Younger sibling	Older sibling	Younger sibling	Older sibling
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}(B_i \geq -91)$	0.089***	0.087***	0.080***	0.095***		
	(0.006)	(0.005)	(0.005)	(0.010)		
$\mathbf{1}(B_i \geq -61)$	0.071***	0.071***	0.072***	0.069***		
	(0.007)	(0.006)	(0.009)	(0.008)		
$\mathbf{1}(B_i \geq -30)$	0.113***	0.111***	0.112***	0.111***		
	(0.006)	(0.006)	(0.007)	(0.009)		
$\mathbf{1}(B_i \geq 0)$	0.454***	0.444***	0.476***	0.411***	0.465***	0.417***
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
Birth-year dummies		Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes	Yes	Yes
R-squared	0.213	0.299	0.329	0.271	0.305	0.235
Observations	520,786	520,786	260,393	260,393	40,470	43,437

Notes: Table reports estimates of α_k using Eqn. (1). Columns (2)–(6) control for student characteristics (gender, whether lives in the capital, monthly family income, mother's schooling, father's schooling and family size) and birth-year dummies. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Balance on covariates at the discontinuity

	Own effect		Spillover effect	
	β_1	Std. err.	γ_1	Std. err.
(a) Younger siblings				
1=July born			−0.013	(0.010)
Age	0.483***	(0.007)	0.003	(0.009)
1=Female	−0.003	(0.009)	0.002	(0.010)
1=Lives in the capital	0.011	(0.008)	−0.003	(0.010)
Monthly family income	9.157	(8.887)	4.398	(9.477)
Mother's schooling	0.056	(0.056)	−0.059	(0.057)
Father's schooling	0.048	(0.070)	−0.017	(0.076)
Family size	−0.010	(0.023)	0.005	(0.023)
1=Father employed	0.004	(0.010)	−0.001	(0.009)
1=Mother employed	−0.008	(0.014)	0.010	(0.011)
1=Public health coverage	−0.017	(0.013)	0.012	(0.012)
(b) Older siblings				
1=July born			−0.008	(0.007)
Age	0.420***	(0.006)	−0.005	(0.007)
1=Female	−0.004	(0.011)	−0.001	(0.007)
1=Lives in the capital	−0.003	(0.010)	0.011	(0.008)
Monthly family income	4.398	(9.477)	9.157	(8.887)
Mother's schooling	−0.059	(0.057)	0.056	(0.056)
Father's schooling	−0.017	(0.076)	0.048	(0.070)
Family size	0.005	(0.023)	−0.010	(0.023)
1=Father employed	−0.012	(0.009)	0.004	(0.006)
1=Mother employed	0.011	(0.010)	0.002	(0.011)
1=Public health coverage	0.005	(0.009)	−0.014	(0.012)

Notes: Table reports estimates of Eqns. (2) and (3) where Y_i and Y_j are a series of child and household characteristics. Column labelled β_1 reports the estimated discontinuity in each covariate at the threshold in own date of birth using a subsample of children born in June/July. Column labelled γ_1 reports the estimated discontinuity at the threshold in sibling date of birth using a subsample of children whose siblings are born in June/July. There are $N = 40{,}470$ younger siblings and $N = 43{,}437$ older siblings born in June/July. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Own effect of July birth on test scores at different ages

	Younger siblings β_1^Y	Older siblings β_1^O
SIMCE test score 4th grade	0.185***	0.136***
	(0.021)	(0.040)
SIMCE test score 8th grade	0.135***	0.111***
	(0.020)	(0.022)
SIMCE test score 10th grade	0.111***	0.075***
	(0.026)	(0.020)
High school GPA 9th-12th grade	0.090***	0.036**
	(0.020)	(0.017)
PSU test score 12th grade	0.091***	0.039***
	(0.016)	(0.011)

Notes: Estimates of β_1 from a linear version of Eqn. (2) using data only from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Each row is a separate regression with the test score outcome shown in the table. The effects are in terms of standard deviations. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Spillover effect of July birth on sibling PSU test scores.

	(1)	(2)	(3)	(4)	(5)
Older-to-younger:	-0.004	-0.002	0.014	-0.014	-0.017
$\mathbf{1}(B_i \geq 0)$	(0.017)	(0.016)	(0.013)	(0.016)	(0.017)
Effective obs.: Left				28,785	44
Effective obs.: Right				32,920	45
Optimal Bandwidth				43.43	44.49
Mean dep. var below cutoff	0.106	0.106	0.108	0.108	0.108
Observations	43,437	43,437	260,393	260,393	365
Controls		Yes	Yes	Yes	Yes
Younger-to-older:	0.067^{***}	0.056^{***}	0.048^{***}	0.063^{***}	0.060^{***}
$\mathbf{1}(B_i \geq 0)$	(0.022)	(0.018)	(0.012)	(0.024)	(0.019)
Effective obs.: Left				18,833	38
Effective obs.: Right				17,887	39
Optimal Bandwidth				27.58	38.04
Mean dep. var below cutoff	0.133	0.133	0.130	0.130	0.129
Observations	40,470	40,470	260,393	260,393	365
Controls		Yes	Yes	Yes	Yes

Notes: Table shows estimates of γ_1 from Eqn. (3). Column (1) reports the linear model shown in Figure 4 using data only from focal children born in June and July. Column (2) additionally controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Column (3) estimates a linear piecewise model (see Eqn. (1)) which uses all the cut-offs shown in Figure 2. Column (4) implements a non-parametric piecewise model using the methodology proposed by Calonico et al. (2014), with a polynomial of order one and weighted by a triangular kernel. Column (5) follows Cattaneo et al. (2024) and treats each day of birth as a separate observation. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Spillover effect of July birth on test scores from grade 4 up to grade 12

	Younger-to-older γ_1^O	Older-to-younger γ_1^Y
Test score 4th grade	0.040	-0.005
	(0.033)	(0.016)
Test score 8th grade	0.053^{*}	-0.049^{**}
	(0.028)	(0.022)
Test score 10th grade	0.082^{***}	-0.008
	(0.030)	(0.023)
High school GPA 9th-12th grade	0.012	-0.021
	(0.024)	(0.015)
PSU test score 12th grade	0.056^{***}	-0.002
	(0.018)	(0.016)

Notes: Estimates of γ_1 from a linear version of Eqn. (3) using data only from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Each row is a separate regression with the test score outcome shown in the table. The effects are in terms of standard deviations. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 8: Spillover effects on higher education outcomes

	Younger-to-older γ_1^O	Older-to-younger γ_1^Y
College enrolment	0.017	0.001
	(0.012)	(0.009)
College enrolment (5yr)	0.016^{**}	-0.005
	(0.008)	(0.007)
Selective Enrolment	-0.003	0.004
	(0.008)	(0.009)
Selective Enrolment (5yr)	-0.004	-0.003
	(0.009)	(0.008)
6-year graduation	0.006	0.009
	(0.008)	(0.010)
10-year graduation	0.021^{**}	-0.005
	(0.009)	(0.010)

Notes: Estimates of γ_1 from a linear version of Eqn. (3) using data only from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Each row is a separate regression with the binary outcome variable shown in the table. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 9: Heterogeneous spillover effects on PSU test scores

	Younger-to-older γ_1^O	Older-to-younger γ_1^Y
(1) All siblings (baseline)	0.056***	−0.002
	(0.018)	(0.016)
(2) Same sex	0.052**	−0.043**
	(0.021)	(0.017)
(3) Different sex	0.059**	0.044**
	(0.023)	(0.022)
(4) Same sex (male)	0.023	−0.056**
	(0.032)	(0.031)
(5) Same sex (female)	0.076**	−0.031
	(0.030)	(0.029)
(6) Different sex ($i =$ female $j =$ male)	0.035	−0.005
	(0.033)	(0.032)
(7) Different sex ($i =$ male $j =$ female)	0.084***	0.092***
	(0.033)	(0.032)
(8) Small age gap	0.074***	−0.004
	(0.021)	(0.020)
(9) Large age gap	0.037	−0.002
	(0.024)	(0.022)
(10) Low income	0.044*	0.004
	(0.022)	(0.024
(11) High income	0.080**	−0.009
	(0.032)	(0.023)
Observations	40,470	43,437

Notes: Estimates of γ_1 from Eqn. (3) for different subgroups, using the specification shown in Column (2) of Table 6. "Small age gap" is defined as those whose age difference is below the median in the estimation sample (3.12 years). "Low income" is defined as students from families whose income falls below the median of the distribution (384 USD). The effects are in terms of standard deviations. Standard errors are clustered at the running variable at daily level. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10: Spillover effect, controlling for time since sibling school start.

	Older-to-younger	Younger-to-older	Fully-interacted model	Controlling for time since sibling school start
	(1)	(2)	(3)	(4)
$\mathbf{1}(B_i > 0)$	-0.012	0.062^{***}	-0.012	-0.015
	(0.016)	(0.021)	(0.016)	(0.015)
$\mathbf{1}(B_i > 0) \times$ Older			0.074^{***}	0.068^{***}
			(0.024)	(0.024)
Years-from-shock dummies				Yes
Observations	118,590	87,989	206,579	206,579

Notes: Table shows estimates of γ_1 from a linear version of Eqn. (3) where the dependent variable is SIMCE test scores in grades 4, 8 and 10 and PSU test scores from grade 12. Estimates use data only from focal children born in June and July. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 11: Effects of July birth on parental investments

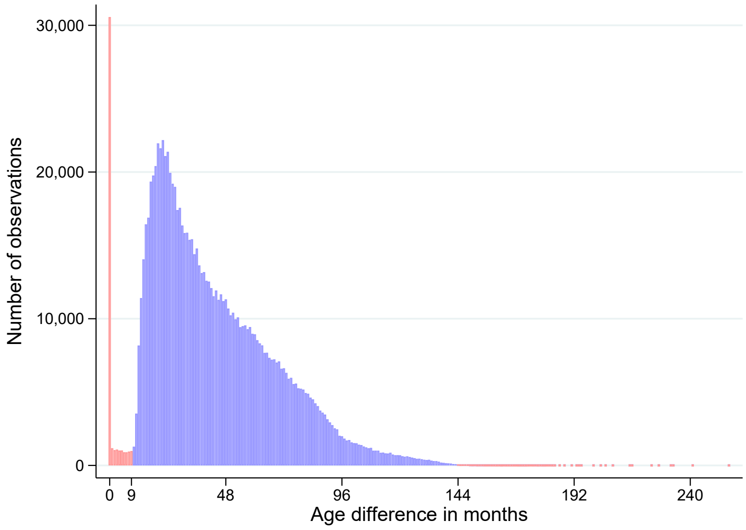
	Younger sibling	Older sibling	Low-income families	High-income families
	(1)	(2)	(3)	(4)
(a) Own effects of July birth on parental investment				
2nd grade (parental response)	-0.051	-0.063^{**}	-0.028	-0.087^{***}
	(0.034)	(0.028)	(0.037)	(0.030)
8th grade (student response)	-0.043	-0.053	-0.047	-0.049^{**}
	(0.029)	(0.034)	(0.032)	(0.024)

	Older-to-Younger	Younger-to-Older	Low-income families	High-income families
	(1)	(2)	(3)	(4)
(b) Spillover effects of July birth on parental investment				
2nd grade (parental response)	-0.016	0.162	-0.018	0.014
	(0.029)	(0.113)	(0.040)	(0.042)
8th grade (student response)	0.030	0.019	0.025	0.020
	(0.028)	(0.037)	(0.032)	(0.036)

Notes: Panel (a) shows estimates of β_1 from Eqn. (2) and panel (b) shows estimates of γ_1 from Eqn.(3) using data only from focal children born in June and July. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figures

Figure 1: Distribution of age differences (in months).



Notes: The observations in red correspond to twins who have the same date of birth (assumed to be twins), siblings pairs born less than nine months apart, and siblings whose age difference is greater than 12 years. These observations are removed from the sample.

Figure 2: School starting age and date of birth

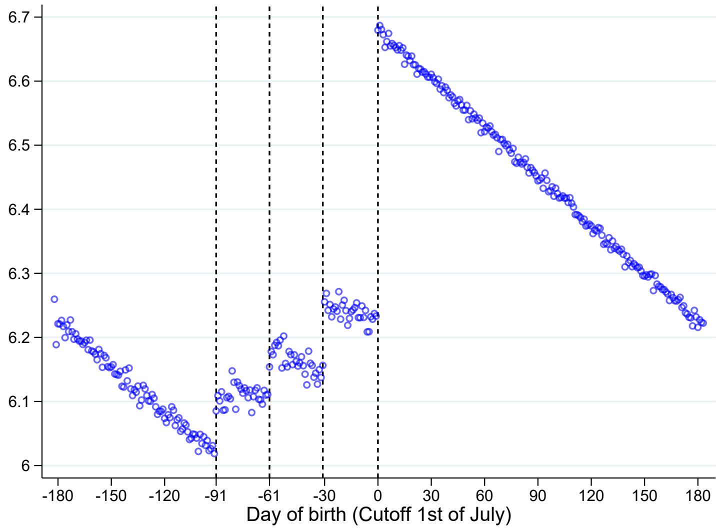
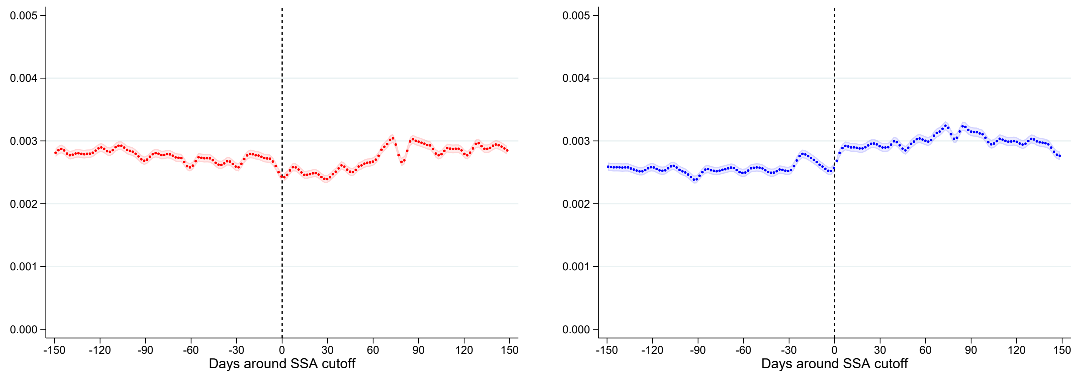


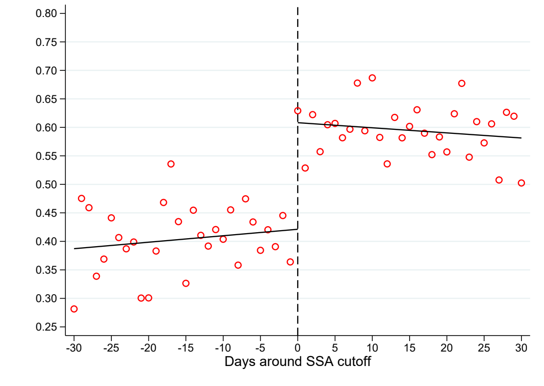
Figure 3: Estimated density of date of birth relative to school-entry cut-off date.



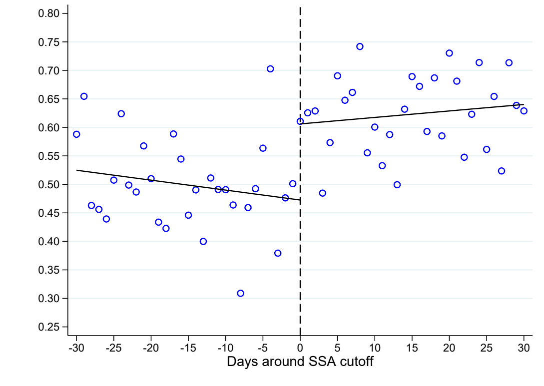
(a) Younger Sibling (b) Older Sibling

Notes: Panels (a) an (b) show an estimated density of date of birth within -/+150 days with shaded 95% confidence intervals using 200 points and a bandwidth of 5 days, for younger siblings and the older siblings, respectively. The dashed line corresponds to 1st of July.

Figure 4: Own effect of July birth on 4th grade SIMCE test scores



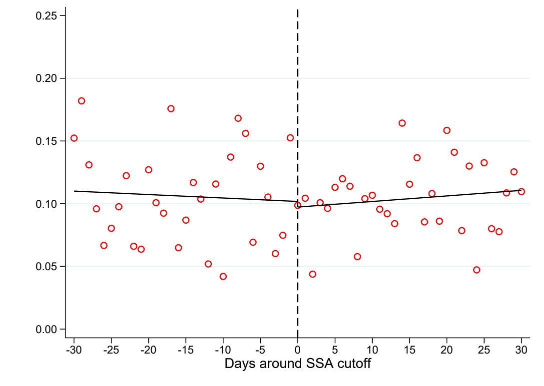
(a) Younger sibling



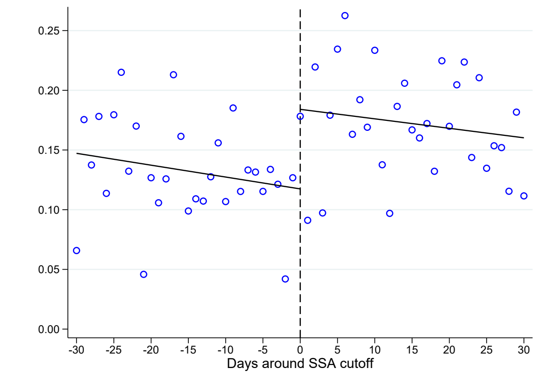
(b) Older sibling

Notes: Estimated using a linear version of Eqn. (2) using data only from focal children born in June and July.

Figure 5: Spillover effect of July birth on sibling PSU test scores.



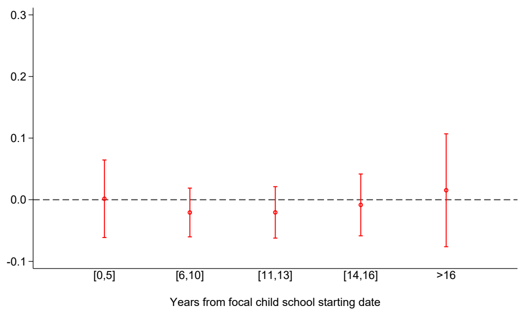
(a) Older-to-younger sibling



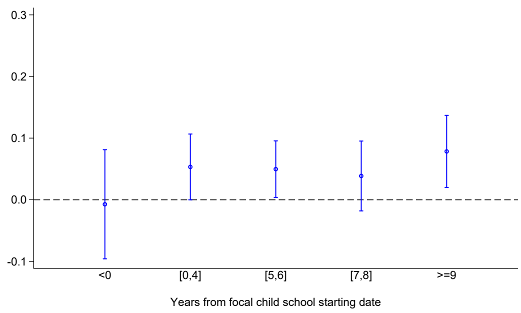
(b) Younger-to-older sibling

Notes: Estimated using a linear version of Eqn. (3) using data only from focal children born in June and July.

Figure 6: Variation in spillover effect by time between the sibling's test score outcome and the focal child's school start date.



(a) Older-to-younger spillovers



(b) Younger-to-older spillovers

Notes: Uses all SIMCE test scores in grades 4, 8 and 10 and PSU test scores in grade 12. Estimated using a linear version of Eqn. (3) using data only from focal children born in June and July. Standard errors are clustered at the running variable at daily level.

A Additional tables

Table A1: Data available for test scores and parental investment.

Cohort	SIMCE 4th grade	SIMCE 8th grade	SIMCE 10th grade	PSU 12th grade	Parental survey 2nd grade	Student survey 8th grade
2004	-	Yes	-	Yes	-	-
2005	-	-	Yes	Yes	-	-
2006	-	-	-	Yes	-	-
2007	-	-	-	Yes	-	-
2008	-	Yes	Yes	Yes	-	-
2009	-	-	-	Yes	-	-
2010	Yes	-	Yes	Yes	-	-
2011	-	Yes	-	Yes	-	-
2012	-	-	Yes	Yes	-	-
2013	Yes	Yes	-	Yes	-	Yes
2014	Yes	-	Yes	Yes	-	-
2015	Yes	Yes	Yes	Yes	-	Yes
2016	Yes	-	Yes	Yes	-	-
2017	Yes	Yes	Yes	Yes	-	Yes
2018	Yes	Yes	Yes	Yes	-	Yes
2019	Yes	Yes	Yes	Yes	-	-
2022	-	-	-	-	Yes	-
2023	-	-	-	-	Yes	-
2024	-	-	-	-	Yes	-
2025	-	-	-	-	Yes	-

Notes: Cohort refers to the year in which students graduate from high school (end of grade 12).

Table A2: Mean comparison test between matched and non-matched siblings

Variable	Sibling match (1)	No sibling match (2)	Diff.	Std. Error	Observations
1= Female	0.493	0.486	0.007***	0.001	875,570
Age difference	3.292	4.764	−1.473***	0.005	875,570
1 = Same-sex sibling	0.515	0.487	0.028***	0.001	875,570
1 = Public school	0.376	0.348	0.028***	0.001	875,570
1 = Voucher school	0.514	0.534	−0.020***	0.001	875,570
1 = Private school	0.110	0.118	−0.008***	0.001	875,570
1 = Lives in the capital	0.374	0.410	−0.035***	0.001	875,570
Mother's schooling	12.470	12.803	−0.333***	0.009	582,125
Father's schooling	12.368	12.643	−0.275***	0.010	582,125
School attendance	91.237	91.148	0.090***	0.024	826,666
Score 4th grade	0.322	0.302	0.020***	0.003	553,048
Score 8th grade	0.142	0.114	0.028***	0.005	171,740
Score 10th grade	0.200	0.168	0.031***	0.005	210,790

Notes: Table uses a sample of children who have a common pair of parental ID numbers and compares siblings who have a common surname-school-municipality (Column 1) and those who do not have a common surname-school-municipality (Column 2). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3: "Redshirting" and family characteristics.

	June born		July born	
	On-time	Delayed	On-time	Early
1=Female	0.55	0.50	0.52	0.55
1=Lives in the capital	0.31	0.51	0.41	0.43
Monthly family income	439.44	671.69	564.41	667.03
Mother's schooling	11.88	12.92	12.43	13.58
Father's schooling	11.89	13.06	12.50	13.95
Family size	4.51	4.56	4.52	4.54
1=Public school	0.43	0.26	0.34	0.18
1=Voucher school	0.52	0.49	0.51	0.48
1=Private school	0.05	0.24	0.15	0.34
Observations	20,098	21,070	41,488	416

Notes: Table excludes a very small number of children born in June who start school in the previous school year and July born children who start school in the following school year.

Table A4: RD estimates of the spillover effects of July birth on probability of completing high school and taking PSU test

	Younger-to-older γ_1^O	Older-to-younger γ_1^Y
(a) Probability of completing high school		
$\mathbf{1}(B_i \geq 0)$	0.0001	0.008*
	(0.005)	(0.004)
Observations	86,197	91,534
(b) Probability of taking PSU test		
$1(B_i \geq 0)$	0.003	0.010**
	(0.006)	(0.004)
Observations	45,779	49,216

Notes: Estimates of γ_1 from a linear version of Eqn. (3) using data only from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Panel (a) uses a sample of all sibling pairs reaching age 20 between 2004 and 2024 and the dependent variable is an indicator for high school completion by age 20. Panel (b) uses a sample of all high school graduates between 2004 and 2018, and the dependent variable is an indicator for taking the PSU test. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A5: Comparison of standard errors

(a) Own effects	Younger sibling	Older sibling
Coeff.	0.091	0.039
SE (Clustered by day of birth)	(0.016)	(0.011)
SE (Clustered by school)	(0.016)	(0.015)
SE (Heteroskedasticity Robust)	(0.016)	(0.016)
(b) Spillover effects	Older-to-Younger	Younger-to-Older
Coeff.	−0.002	0.056
SE (Clustered by day of birth)	(0.016)	(0.018)
SE (Clustered by school)	(0.016)	(0.016)
SE (Heteroskedasticity Robust)	(0.015)	(0.016)

Notes: Panel (a) shows estimates of β_1 from Eqn. (2) and panel (b) shows estimates of γ_1 from Eqn.(3) using data only from focal children born in June and July. Dependent variable is the PSU test score.

Table A6: Placebo RD estimates

	July vs. August		August vs. September		September vs. October	
	i Younger	i Older	i Younger	i Older	i Younger	i Older
(a) Own effect β_1						
$\mathbf{1}(B_i \geq 0)$	-0.002	0.031	-0.010	-0.052	0.060	-0.005
	(0.029)	(0.033)	(0.061)	(0.049)	(0.082)	(0.057)
(b) Spillover effect γ_1						
$\mathbf{1}(B_i \geq 0)$	-0.013	0.049^{*}	0.016	-0.011	-0.041	-0.009
	(0.025)	(0.029)	(0.049)	(0.056)	(0.061)	(0.074)
Observations	40,119	46,513	42,892	47,760	45,382	48,527

Notes: Estimates of γ_1 from a linear version of Eqn. (3) using data from focal children born in consecutive months not associated with school entry cut-offs, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Dependent variable is the PSU test score. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A7: RD estimates of β_1 and γ_1 for families with two or three siblings

Birth order	Birth order (sibling)		
(focal child)	1	2	3
1	0.040^{***}	0.009	0.009
	(0.011)	(0.014)	(0.031)
	[51,842]	[51,842]	[8,178]
2	0.054^{***}	0.075^{***}	-0.014
	(0.014)	(0.013)	(0.034)
	[48,451]	[48,451]	[7,867]
3	0.101^{***}	-0.007	0.051
	(0.030)	(0.039)	(0.036)
	[7,673]	[7,673]	[7,673]

Notes: The leading diagonal shows estimates of β_1 from Eqn. (2) and off-diagonal elements show estimates of γ_1 from Eqn. (3) using data from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Dependent variable is the PSU test score. Sample sizes for each estimate are reported in brackets. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8: Spillover effect on the probability of school transition for older siblings

	School transition (1)	Public (2)	Voucher (3)	Private (4)
$\mathbf{1}(B_i \geq 0)$	0.018^{**}	-0.001	0.004	0.001
	(0.007)	(0.004)	(0.004)	(0.002)
Observations	20,520	20,520	20,520	20,520

Notes: Estimates of γ_1 from Eqn. (3) using data from focal children born in June and July, with controls for student characteristics (gender, whether lives in the capital, monthly family income, father's schooling, mother's schooling and household size). Each column is a separate regression. Column (1) shows the probability of changing schools when the younger sibling enters school. Columns (2) through (4) present the probabilities of transitioning to a public, voucher, or private school, respectively. Standard errors are clustered at the running variable at daily level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A9: Descriptive statistics on parental investments

	(1) Difference Older − Younger	(2) Difference High income − Low income
(a) Parental investments: 2nd grade (Parent responses)		
Read to child	0.097^{***}	0.188^{***}
	(0.005)	(0.005)
Parent-child joint reading	0.041^{***}	-0.056^{***}
	(0.005)	(0.005)
Talk about their readings	0.038^{***}	-0.028^{***}
	(0.005)	(0.005)
Give books	-0.007	0.215^{***}
	(0.005)	(0.005)
Index of 2nd grade investments	0.037^{***}	0.109^{***}
	(0.005)	(0.005)
(b) Parental investments: 8th grade (Student responses)		
Always congrats for grades	0.025^{***}	0.109^{***}
	(0.004)	(0.004)
Always knows my grades	0.035^{***}	0.013^{***}
	(0.004)	(0.004)
Always willing to help with study	0.038^{***}	0.228^{***}
	(0.004)	(0.004)
Index of 8th grade investments	0.042^{***}	0.155^{***}
	(0.004)	(0.004)

Notes: Information from panel (a) comes from a parental survey of parents of 2nd grade children taken in 2012–2015 and comprises 175,803 observations. Information from panel (b) comes from a survey of 8th grade students in 2009, 2011, 2013 and 2014 and comprises 243,181 observations. The aggregate parental investment indices at the bottom of each panel are the standardised sum of each element. All measures are in standard deviations. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B Additional figures

Figure B1: Probability of starting school in the year in which the student turns 7

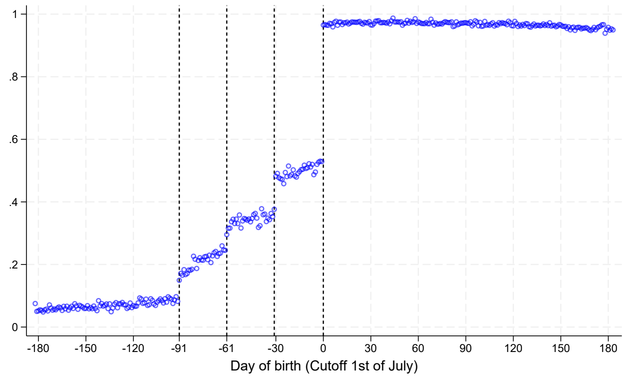
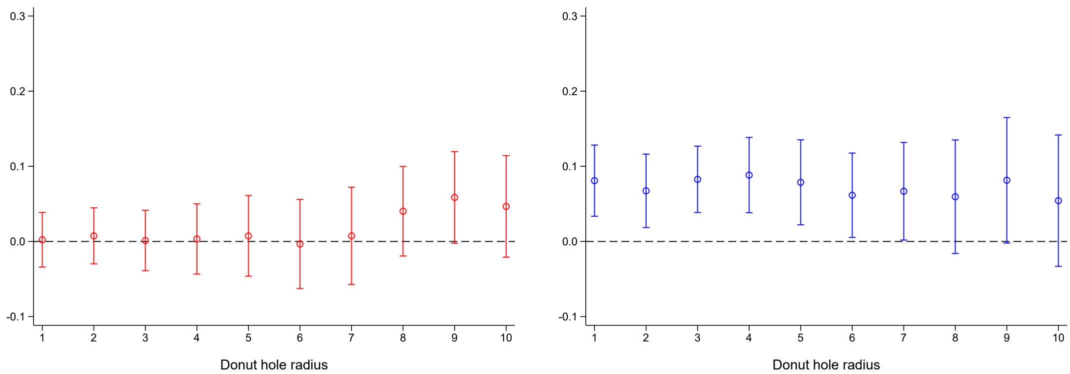


Figure B2: Spillover effect of July birth on PSU test scores, “donut hole” estimates.



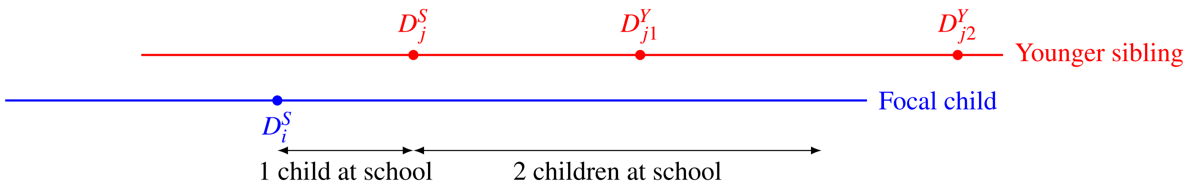
(a) Older-to-younger

(b) Younger-to-older

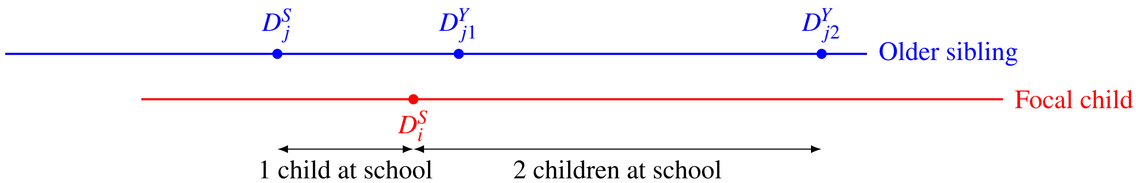
Notes: Panels (a) and (b) show the effects of sibling July birth on academic outcomes for younger siblings (older-to-younger) and older siblings (younger-to-older), removing students who are close to the cut-off within a radius of $+/-$ 10 days. Results are based on linear regressions within +/- 30 days. Standard errors are clustered at the running variable at daily level.

Figure B3: Sibling school start times and spillovers

(a) Older-to-younger spillovers

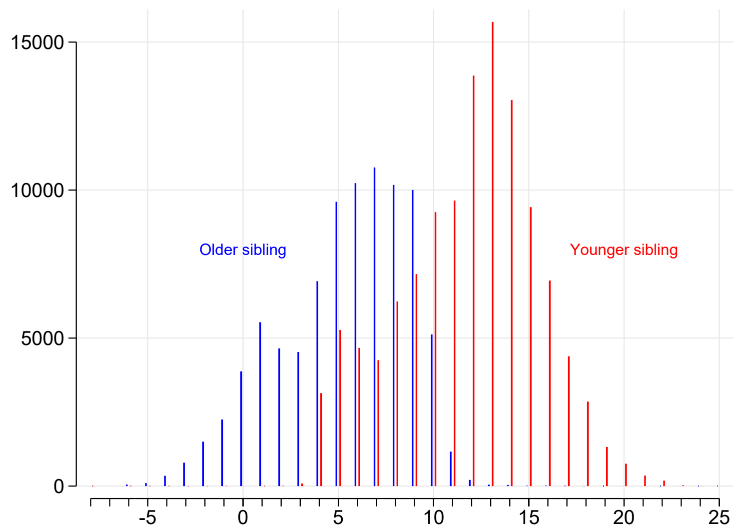


(b) Younger-to-older spillovers



Notes: Blue line indicates older sibling; red line indicates younger sibling. D_i^S indicates school start date of the focal child i. D_{j1}^Y and D_{j2}^Y indicate the academic outcome early and late in the sibling's school career. In the bottom panel, D_{jt}^Y is closer to D_i^S and parents have both children in school immediately after D_i^S.

Figure B4: Frequency distribution of time from sibling school start date to test date



Notes: Uses test scores taken in grades 4, 8, 10 and 12. Older siblings' tests are closer to the "shock" of their younger sibling starting school.