When the Household Becomes the School: Siblings, Parental Attention, and School Closures

Francisco Pardo

University of Texas at Austin

Click here for the latest version

This paper examines how family structure shapes educational outcomes when school closures shift instructional responsibilities to households. First, I document a new global stylized fact: children with siblings experienced substantially larger learning losses than only children during COVID-19 school closures. Using administrative data from Peru, I show that students with siblings lost 0.05 standard deviations more in GPA and standardized test scores, than only children, with effects increasing in family size and persisting years after reopening. Second, I show that these differences reflect a causal effect of siblings rather than correlated factors. To do this, I leverage a difference-in-differences approach that compares children with and without siblings before, during, and after school closures, and find that the estimates remain stable when controlling for observable heterogeneity. Instrumental variable estimates on exogenous changes in family size corroborate these results. Third, I show that parental time constraints is the key mechanism, with effects strongest in families who typically invest more time in their children. A regression discontinuity design exploiting school-entry age cutoffs shows that when younger siblings start school and spend less time at home, older siblings' performance and parental support improve. These spillovers disappeared during school closures when all children remained home. Overall, these results reveal how parental time dilution during school closures disproportionately disadvantaged larger families and highlight the essential role schools play in mitigating household constraints in education production.

JEL Codes: I21, I24, D13

Human capital accumulation at early ages is a key determinant of long-term outcomes such as educational attainment, employment, and wages. Education, in turn, is central to this process, but its production depends on multiple institutions, with households and schools playing crucial yet shifting roles. Parents provide the foundation for children's early learning, shaping cognitive and non-cognitive skills through time investments and the home environment. As children reach school age, formal instruction increasingly shifts to schools, allowing households to reallocate more of their time toward other activities such as leisure, employment, or childcare and support for other children in the household. In this context, the contribution of each input in the education production function varies with children's developmental stages and the availability of formal schooling.

Extended periods of school closures disrupt this normal process by abruptly shifting instructional responsibilities from schools to parents. This sudden reorganization of learning processes provides a unique opportunity to observe how families assume roles typically played by schools. However, the burden of this substitution varies systematically across family structures. In such contexts, having siblings may be either helpful or detrimental to students. On the one hand, they can act as mentors or support, particularly in a context where they have lost contact with their peers. On the other hand, household resources, whether they involve access to technology or parental attention, are also limited and may be diluted when shared among multiple children. If larger families struggle to replicate the role of schools, it reveals the limits of parental substitution and the distinct comparative advantages of each institution in education production.

The COVID-19 pandemic created an unprecedented disruption to schooling, affecting more than 1.6 billion students worldwide. In some countries, schools remained closed for up to two years, with learning losses roughly equivalent to a year's worth of schooling (Jakubowski, Gajderowicz and Patrinos, 2023). Recovery has been slow and uneven, with the largest setbacks among disadvantaged students, including those from low-income households or without reliable internet access (Haelermans et al., 2022). While research has documented these differential impacts, one critical but unexplored dimension of this heterogeneity is family structure.

This paper documents a new global stylized fact: children with siblings experienced substantially larger learning losses compared to only children during school closures. I provide a detailed analysis of this in Peru but show evidence of a similar pattern in both developing and developed countries using international test scores, particularly where closures were prolonged. These differences in learning losses are consistent across geographic settings, parental characteristics, and school types, suggesting a robust and general pattern. It also provides the first causal evidence that family structure mediates learning losses during large-scale educational disruptions. Siblings magnified learning losses, primarily through parental time constraints. These results link the literature on school-household substitution, family size, and pandemic learning losses, highlighting parents' limited attention as a key but often overlooked source of educational inequality. By revealing how the presence of siblings constrains substitution, these patterns illuminate the essential role schools play in the education production function under normal circumstances by offsetting the limitations that families face.

I use administrative data covering the universe of school enrollments in Peru's education system from 2014 to 2024, which includes information on all children from pre-kindergarten (K3) through grade 11, including their progression, GPA, and demographic and school characteristics. Linking children through mothers' IDs allows me to identify siblings among all enrolled students. I complement these data with standardized national examinations administered in grades 2 and 4 and with detailed parent surveys that provide measures of household resources, parental time investments, socioeconomic status, and educational expectations.

My analysis consists of three parts. First, I document that children with siblings experienced larger learning losses than only children during school closures. Second, I show that these differences reflect the causal effect of siblings rather than preexisting household differences. Finally, I examine several mechanisms underlying this effect, identifying parental time constraints as the most plausible channel.

I begin by documenting that children with siblings experienced larger declines in academic performance during the pandemic than only children. Similarly, children with more siblings performed worse than those with fewer siblings. Before the pandemic, both groups followed parallel trends in academic performance, which diverged sharply during and after school closures. Leveraging this similarity before the school closures, I use a difference-in-differences framework that compares changes in performance between families of different sizes who faced different childcare and time shocks. On average, children with siblings experienced learning losses 0.04 standard deviations greater than only children, and the effects reached up to 0.06 standard deviations among those with two or three siblings. These results hold for both GPA measures during and after closures and for standardized exams taken after schools reopened.

I then examine whether this divergence in performance is due to the number of children or to other factors correlated with family size. To verify this, I first show that the effects remain stable when controlling for observable parental and household characteristics, such as the mother's age, education, baseline socioeconomic status, and baseline student achievement. I then complement this analysis with a standard instrumental variables approach, using the sex composition of the first two children as an exogenous predictor of family size.[1] Within this framework, parents with two same-gender children are more likely to have an additional child, and the results confirm that an additional sibling reduces performance during school closures. Moreover, these effects are not limited to GPA, which may reflect reduced engagement of students with siblings during school closures. They also have lasting impacts, reducing standardized test scores three years after schools reopened by 0.05 standard deviations in both reading and math. The results further extend beyond immediate test scores to encompass broader educational trajectories. Using survey data on educational expectations, I find that parents with multiple children became systematically more pessimistic about their children's long-term educational prospects, lowering their expectations that their children would attain a graduate degree by 2 percentage points when schools reopened, from a baseline of around 40%. These shifts in expectations suggest that learning losses may permanently alter families' educational investment trajectories.

Finally, I investigate why additional siblings made families less adaptable to school closures, examining whether the effects vary by birth order, technological access, sibling age gaps, grade level, and parental socioeconomic status. I find no systematic differences by birth order or technological access, suggesting that neither lower resilience nor adaptability among later-born children, nor competition for technological resources, explain the effects. Instead, the results point to parental time constraints as the key mechanism: the effects are largest among elementary school children, where parental

[1] See Black, Devereux and Salvanes (2005), Angrist, Lavy and Schlosser (2010), and Black, Devereux and Salvanes (2010).

involvement is most intensive, and among families in the top three socioeconomic quartiles, where parents typically devote more time to children's learning. This pattern is consistent with lower socioeconomic status families investing little parental time regardless of family size. Using school starting-age cutoffs, I show that when younger siblings start school earlier—reducing the time they spend at home—older siblings' performance improves. I find a similar, though less robust, pattern for parental time investments. These positive spillover effects disappeared during school closures, when all children were home simultaneously, consistent with the dilution of parental time and attention.

This paper contributes to three strands of research. First, it provides the first causal evidence on how family structure mediates learning losses during large-scale educational disruptions in a setting where schools' instructional role dramatically shifts to households. It shows that while parents can partially substitute for schools, this substitution is constrained by family structure, with each additional child reducing educational quality. This finding connects to the literature on parental responses to school inputs. Research shows that parental inputs are less critical for cognitive development after age 5 (Attanasio et al., 2020; Attanasio, Cattan and Meghir, 2022). Also, parents reduce effort when their children attend a better school (Pop-Eleches and Urquiola, 2013) or face increased school resources (Houtenville and Conway, 2008; Fredriksson, Öckert and Oosterbeek, 2016). Other work shows that parents reduce private educational spending in response to anticipated school grants but not in response to unanticipated ones (Das et al., 2013). By examining a setting where schools' support is severely reduced, my results highlight the constraints families face in substituting for schools and the essential role schools usually play in overcoming these.

The paper also contributes to the literature on family size and the quality of education, providing new evidence on when a trade-off arises. Whether there is a quantity-quality trade-off in the number of children and the educational outcomes they achieve has been largely studied in economics. Research shows this trade-off often does not exist (Becker and Tomes, 1976; Black, Devereux and Salvanes, 2005; Angrist, Lavy and Schlosser, 2010). However, unexpected shocks can cause such trade-offs to emerge (Black, Devereux and Salvanes, 2010; Olof Åslund and Hans Grönqvist, 2010). For instance, parents can usually plan and adapt to having a new child in a way that does not affect the quality of education received by other children but having twins may alter that balance. This paper contributes to this literature by showing that school closures created an analogous, unanticipated increase in parental time requirements, activating otherwise latent trade-offs between family size and educational quality.

Finally, this paper contributes to the growing literature on learning losses from school closures by uncovering a key mechanism behind these effects. Prior work documents large and persistent losses, particularly among vulnerable populations (Haelermans et al., 2022; Singh, Romero and Muralidharan, 2022; Jakubowski, Gajderowicz and Patrinos, 2023; Goldhaber et al., 2023; Jack et al., 2023; Lichand and Doria, 2024). The results show a different aspect that sheds light on how these learning losses occurred: through the increased difficulties of translating education into the households when there are multiple children. While other studies find positive non-cognitive effects of siblings during the pandemic (Hughes et al., 2023; Lampis et al., 2023), as they provided emotional and linguistic support that buffered the loss of peer interactions, my results indicate that the effect on academic outcomes was negative.

My findings have important implications for education policy during crises. Differences in parental time constraints are a key source of variation in learning losses across family structures. Because even high-resource families cannot overcome this limitation, policies focused solely on providing technological or financial support may overlook an important source of inequality. Crisis response in education should therefore include additional support for larger families, such as targeted tutoring and remote learning structures that reduce parental supervision requirements. Descriptive evidence from international PISA test scores and from India, suggests that similar patterns emerge elsewhere, with larger effects in countries that experienced longer school closures. The widespread nature of these effects across contexts underscores the need for such policy adjustments to mitigate long-term human capital losses.

The rest of the paper proceeds as follows. Section Section I describes the Peruvian education system and how school closures where implemented, and Section II describes the data used. Section III shows a new stylized fact about siblings and only-children divergence in performance. Section IV establishes the causal relation described in the previous section by discussing the empirical strategy and main results and Section V the mechanisms. Section VI shows evidence of a similar relation occurring in the rest of the world. Section VII concludes.

I. Education System in Peru and School Closures

In this section, I describe the school education system in Peru, how it is structured, some overall statistics and then provide a description of how schools operated during closures and how remote learning was implemented in them.

Peru's basic education system consists of two mandatory levels: six years of primary education (grades 1-6) and five years of secondary education (grades 7-11). Around 5.5 million students are enrolled in these grades each year. According to household surveys, in 2019, 99% of children aged 7-11 were enrolled in primary education and 85% of those aged 12-16 were enrolled in secondary education. Most students are enrolled in public school, although private education is high compared to other developing countries, with almost 15% of students both in primary and secondary education enrolled in private institutions.

When the COVID-19 pandemic forced school closures in March 2020 at the beginning of the academic year, Peru's Ministry of Education rapidly launched "Aprendo en Casa" (Learn at Home), a national strategy to sustain educational services remotely. The strategy relied on three primary channels to reach students across diverse geographic and socioeconomic contexts: a web platform, television broadcasts, and radio programming. This multi-channel approach was designed to address Peru's significant digital divide, recognizing that while 86.6% of households had television access and 48.8% had radio, only 37.6% had internet connectivity, with only 6% in rural areas. The Ministry partnered with national broadcasters and produced educational content aligned with the National Basic Education Curriculum (CNEB), covering all levels of basic education.

Public education during school closures was designed to be delivered asynchronously. Teachers were expected to adapt centrally-produced experiences to local contexts and individual student needs, maintaining regular communication with families through available means, primarily WhatsApp, text messages, and phone calls. The strategy emphasized student autonomy in learning, family accompaniment (rather than substitution of teacher roles), and flexible differentiation based on each student's access conditions and circumstances.

Implementation revealed significant structural challenges that shaped how education was actually delivered. By July 2020, approximately 71% of students accessed content via television, with teachers predominantly using cellphones to distribute materials, provide guidance, and maintain student connection. Often, these cellphones belonged to parents, who also used them for work, further limiting children's access to learning materials and communication with teachers. The shift placed extraordinary demands on families, particularly mothers, 63.7% of whom reported accompanying their children while they watched the content. Regional governments and local municipalities supplemented national efforts by installing antennas, distributing printed materials, providing connectivity support, and developing complementary content. While many of these limitations also applied to those in private schools, the measures taken by them would be done independently, with some of them likely implementing remote learning with synchronous virtual classes. All of these environments potentially relied on higher parental involvement for its effectiveness.

Additionally, during school closures, the Ministry of Education decided to promote every student regardless of performance. A different grading scale was also adopted. In elementary schools, where students are graded A through D, during closures, students were only given grades A through C. In secondary school, they were graded from 11 to 20 instead of 0 to 20. To make grades comparable I assign a C or 11 to those who got a D or a grade between 0-10 during school normal operation.[2]

II. Data

I estimate the effects of family structure on educational outcomes before, during, and after school closures using the national population of enrolled students tracked through administrative records, combined with standardized exams and parental survey data covering the period 2014–2024. This comprehensive dataset allows me to examine effects even after schools reopened. Siblings are identified across sources using the mother's national identification number.

A. Administrative data on school progression and GPA (SIAGIE)

The SIAGIE (Sistema de Información de Apoyo a la Gestión Educativa) is a comprehensive administrative database maintained by Peru's Ministry of Education that tracks enrollment and academic records for all students in the country's education system. From 2014 to 2024, the system captures data across pre-kindergarten through 11th grade in both public and private institutions. The database contains detailed information on each student's school enrollment, grade level, grades by subject, passing status[3], sex, parents' education levels and parents' date of birth. Unique student identifiers allows individual students to be tracked across years. Additionally, I also had access to the dates of birth of all students enrolled in 2024.

B. Sibling Identification

The focus of my research is on family structure, specifically on identifying students with siblings. To do so, I use parental identification numbers to link students who share the same mother, which serves as a proxy for living in the same household. This information is available for 98% of students enrolled from pre-K through 11th grade.

[2] This change in grading policy is less relevant for elementary school given that only about 1% of students get a D. In secondary schools about 10% of students get a grade of 10 or lower so the adjustment helps to make grades comparable. However, unadjusted grades lead to similar results.

[3] I do not explore school progression as an outcome because the Ministry of Education adopted universal progression during the COVID-19 pandemic.

In Table 1, I show that 38% of students are only children, 32% have one sibling, 20% have two siblings and 10% have three siblings. Given that the identification of siblings is through enrollment data, this measure may underestimate the number of siblings each student has, which would potentially attenuate my results. Characteristics from panel A are similar in the first two columns, both with 79% of students in urban areas, two thirds of students in public schools, an average class size of 24 students and 39% of mothers with complete secondary education. Panel B shows some academic characteristics based on administrative SIAGIE data such as grade promotion and standardized measures of GPA at the class-year-school level.

C. Standardized National Examinations (ECE)

In addition to GPA, to measure academic performance, I use national standardized examinations conducted by the Peruvian Ministry of Education, also known as ECE (Evaluacion Censal de Estudiantes), which evaluate students in mathematics and reading skills. These evaluations were implemented across different academic years and grade levels. Second grade students were tested from 2007 to 2016 nationally and then in smaller representative samples in in 2019 and 2022. Fourth grades students were tested nationally in 2016, 2018 and 2024 and in smaller representative samples in 2019, 2022 and 2023.[4] The test is low-stakes for the students and measures their basic competencies in math and language at the end of the school year. This allows for a measure of learning losses that is not dependent on within school-grade variations. These tests are standardized with mean 0 and standard deviation of 1 in the base year of 2007, in order to have comparable measures across time. In Table 1, I show that only children score 0.1 standard deviations lower than children with one sibling in second grade mathematics and 0.02 standard deviations higher than children with two siblings. In second grade reading, only children score 0.04 standard deviations lower than children with one sibling and 0.12 standard deviations higher than children with two siblings.

[4]The national examinations were applied to schools with at least five students in the respective grade

D. Surveys

Along with standardized examinations, starting in 2015, the Ministry of Education surveyed parents, teachers and principals. These include information about socioeconomic status, parent's mother tongue, expectations for educational attainment, parental investment in education, access to internet and a computer, etc. Socioeconomic status is reported in a standardized index with mean 0 and standard deviation of 1. This is based on materials in the household, access to services, assets owned, and parent's education level. In Table 1, families with only one children have higher socioeconomic status on average, although only 0.03 standard deviations more than those with two children. Access to internet and a computer is also similar between these two groups with around a third of households having them. I show that most parents have high expectations for the maximum level of education that their children will achieve. 80% of parents of only children expect that to be college education or higher, similar to 81% of parents of children with one sibling.

Only children and children with one sibling are similar in most observable characteristics and those with more siblings tend to have lower performance and socioeconomic status. However, these are average differences across schools. Ultimately, I will be comparing students within a school where characteristics will be much more similar across families with different number of children.

III. Stylized Fact: Larger Learning Losses for Children with Siblings

A. Empirical Strategy

In this section, I document a new fact about family structure and student performance during school closures: children with siblings experienced larger learning losses than those without. I begin by examining how performance trends evolved during and after an extended period of school closures, using both GPA and standardized exams. I then analyze these patterns within a difference-in-differences framework, treating children without siblings as the control group, to estimate the effect of having a sibling on learning losses during school closures. This identification strategy is expanded in Section IV.

To make GPA measures more comparable across schools and over time, I make two adjustments. First, I impose the grading scale used during school closures on all other years, censoring lower grades as discussed in the previous section. Second, I standardize GPAs at the school-grade-year level to account for differences in teachers' grading standards and leniency.[5] In the case of standardized exams, scores are already comparable across schools and over time, so changes from year to year directly reflect gains or losses in performance. The plots reveal how both groups followed similar trajectories before 2020, but diverged sharply once schools closed, illustrating the onset of differential learning losses for children with siblings.

Underneath the comparison of these trends, and how the gap between both groups has changed after closures is a difference-in-difference framework. In order to have a more accurate estimate of the size of this change, I estimate the coefficient of those heterogeneous learning losses using the following equation for the difference-in-difference:

$$Y_{isgt} = \alpha + \delta Sib_i + \boldsymbol{\beta} Post_{it} Sib_i + \lambda_s + \mu_g + \tau_t + \varepsilon_{isgt} \tag{1}$$

Similarly, the corresponding estimate for the event study is as follows:

$$\begin{aligned} Y_{isgt} = \alpha + \delta_1 + \delta_2 Sib_i + \sum_{k=-5}^{-2} \delta_k (\mathbb{I}[t = 2020 + k] Sib_i) \\ + \sum_{k=0}^{4} \beta_k (\mathbb{I}[t = 2020 + k] Sib_i) + \lambda_s + \mu_g + \tau_t + \varepsilon_{isgt} \end{aligned} \tag{2}$$

where Y denotes the standardized GPA, Sib is an indicator variable taking the value one if the individual has siblings (or alternatively if the individual has one, two or three

[5]These adjustments help with interpretation of results, although the effects found are consistent with using unadjusted measures.

siblings), and *Post* is an indicator variable taking the value one if the year is 2020 and over, to account for the beginning of school closures. I also include a set of school (λ), grade (μ) and time (τ) fixed effects. The coefficient of interest is captured by β, which represents difference in achievement gaps between children with and without siblings. Standard errors are clustered at the school level.

B. *Performance before, during, and after school closures*

In Figure 1a, I present trends in fourth-grade standardized mathematics exams, which provide an absolute measure of learning that can be compared consistently across years. These exams were administered each November from 2016 to 2024 (except in 2017 and the years of school closures). Before closures, both groups performed similarly and experienced learning gains at comparable rates. After schools reopened, students were tested again in 2022 and 2023 in a representative sample, and all fourth graders were assessed in 2024. The emergence of a gap between children with siblings and only children is evident. This is particularly striking because not all of these students were enrolled in school when closures began. The 2024 cohort started first grade in 2021, and yet even they show larger losses among those with siblings. While standardized scores capture overall learning losses, they are measured as early as a year after schools reopened, when some recovery may have already occurred. To examine changes in real time, I next turn to GPA measures, which are available for every school year, including those affected by closures.

To ensure comparability of GPA measures across schools and years, I use standardized GPAs within each school-grade-year. In Figure 1b, I show how this measure evolved for children with and without siblings. The gap widens markedly during school closures and remains partially persistent after schools reopened. Note that because the variable is standardized within each year, the overall mean is always zero. Therefore, the apparent increase in scores among only children should be interpreted as a relative improvement, not an absolute gain in performance. In addition to the larger losses during school closures, after they reopened the gap remains larger than it was before.

I show results from the event study specification in Figure 2. Panel A shows that during school closures, children with siblings experienced larger learning losses of about 0.04sd and then stayed 0.01sd lower once schools reopened. In Panel B, I show that these effects are larger for those who had more siblings, although the main change occurs when going from zero to one sibling or from one to two; having more than two is associated with a similar learning loss.

IV. The Causal Effect of Siblings

A. *Empirical Strategy*

In Section III I showed a new stylized fact, that children who had siblings exhibited larger learning losses when compared to those who did not. I also provided an estimate for this difference in learning losses with a simple difference-in-difference estimate that compared both groups within each school, grade and year. However, families with different number of children may also differ in socioeconomic status, preferences for quality of education, mother's age, etc. Thus far, the difference-in-difference approach has dealt with this in two ways: (i) By including school fixed effects, comparisons are made between students within the same school, making both observed and unobserved parental characteristics more homogeneous while maintaining enough variation in the condition of having a sibling, and (ii) the time comparison accounts for differences whose effects are constant over time.

Difference-in-Difference with observed heterogeneity.— One potential concern could be that some of those differences that remain at the within school level interact with school closures and have heterogeneous effect, e.g. if richer parents or parents with more concern for the quality of education also have less children and deal better with remote learning. To address this, I include an extensive set of controls and their interactions with $Post$ for potential observed heterogeneities. Assuming these characteristics are correlated with unobserved ones (e.g. socioeconomic status correlated with preferences for educational quality) I can have a sense of the degree of potential endogeneity by unobserved characteristics based on the change of our estimated coefficient as proposed by Oster (2019). Data on mother's level of education and date of birth is available for the full sample, but in order to include a more comprehensive set of controls, I use information from baseline surveys on socioeconomic status and standardized exam achievements as well.

I complement equations (1) and (2) by including the term X_{ist} and its interaction with $Post_{it}$ to account for heterogeneous effects on observable characteristics. Our extended difference-in-difference estimate is the following:

$$Y_{isgt} = \alpha + \delta Sib_i + \boldsymbol{\beta} Post_{it} Sib_i + \gamma_1 X_{ist} + \gamma_2 Post_{it} X_{ist} + \lambda_s + \mu_g + \tau_t + \varepsilon_{isgt} \tag{3}$$

Similarly, the corresponding estimate for the event study is as follows:

$$\begin{aligned} Y_{isgt} &= \alpha + \delta_1 + \delta_2 Sib_i + \sum_{k=-5}^{-2} \delta_k (\mathbb{I}[t = 2020 + k] Sib_i) \\ &+ \sum_{k=0}^{4} \beta_k (\mathbb{I}[t = 2020 + k] Sib_i) + \gamma_1 X_{ist} + \gamma_2 Post_{it} X_{ist} + \lambda_s + \mu_g + \tau_t + \varepsilon_{isgt} \end{aligned} \tag{4}$$

If the endogeneity that jointly affects fertility decisions and performance during school closures is partly correlated with observable characteristics, controlling for these variables should mitigate some of the resulting bias.

Sex Composition as an Instrument for Family Size.— Previous literature has documented that parents have preferences for variety and so are more likely to have a third child if the previous first two children are both of the same sex (Angrist and Evans (1998)). As a result, if the sex composition is random, families that randomly had two same-sex children are more likely to have a third, all other things equal, making sex composition a potential instrument. The 2SLS equations that uses having the first two children of the same sex as an instrument for family size to estimate its effect on performance are as follows:

$$FamSize_{it} = \delta_0 + \delta_1 SameSex_{it} + \delta_2 X_{it} + \nu_{it} \tag{5}$$

$$Y_{it} = \beta_0 + \beta_1 \widehat{FamSize}_{it} + \beta_2 X_{it} + \epsilon_{it} \tag{6}$$

Where Y_{it} is an education outcome, which in this case is the standardized GPA, $FamSize$ is the amount of children in the family, $SameSex$ is a dummy variable that takes the value of one if the first two children are of the same sex, and X_{it} is a set of demographic characteristics such as mother's education level, mother's age and age at first birth and a birth order dummy for the second born child.

B. Results

I present evidence that the presence of a sibling is the causal driver of these effects, rather than unobserved heterogeneity correlated with family characteristics. I further demonstrate persistence beyond GPA, using standardized exams and parental expectations after schools reopened.

Difference-in-Difference with observed heterogeneity.— In Table 2 I show a DID estimate for sixth grade students between 2019 and 2020 using their second grade baseline characteristics given that in 2015 and 2016 students in that grade were tested and their parents surveyed.[6] Panel A shows the effects on standardized GPA for mathematics and Panel B for reading. In the first column, I show the estimate controlling for fixed effects but no other control for heterogeneous characteristics just as was done in equation (1). The size of the estimate for this sample is -0.051sd for mathematics and -0.034sd for reading. The rest of the columns progressively add other controls: dummy for the mother having higher education, both mother's age and age at first birth being

[6] I also identify two additional groups for which similar baseline information is available. I estimate a difference-in-differences model using 2018 and 2020 sixth-grade students with second- and fourth-grade baseline characteristics, and 2019 and 2020 fifth-grade students using fourth-grade baseline characteristics. All of these results are presented in Table A.1.

above 30, and dummies for socioeconomic status and performance in standardized test scores being above the mean. The main coefficient exhibits a significant stability across all these estimates, with a reduction of less than 10% of its initial value. Given the consistency of my estimates when including observed heterogeneity, it is unlikely that all the effect is driven by heterogeneous effects from unobserved characteristics.[7]

Exogenous change in family size.—The 2SLS estimates are presented in Table 3. I also include, for comparison, two sets of OLS coefficients estimated on exactly the same samples as the 2SLS estimates; the first set includes controls only for birth order, the second set includes all control variables. First, I estimate this for the years 2018-2019, when schools were operating normally. The first-stage estimate on $SameSex$, reported in Table 3, is 0.05. Similarly, Black and Devereux (2010) finds a coefficient of 0.082 using Norweign data, Angrist, Lavy and Schlosser (2010) find a coefficient of 0.073 using Israeli data and Conley and Glauber (2006) find that having two same-sex children increases the probability of having a third child by about 0.07 in the United States. This shows that like in other settings, Peru also exhibits a preference for variety in the sex composition of the children. The fourth column shows the 2SLS estimate of the an exogenous increase in the number of children in the family on the GPA performance of the first two children. This shows a slightly significant positive family size effect on GPA of 0.063. This estimate goes down to -.031 during school closures. However, this paper, rather than exploring the family size effects in a moment in time, focuses on how the penalty of larger families has *increased* during school closures. That is ultimately what the difference-in-difference estimate does. This is captured by the change in family size effect between both periods, equal to $-0.094 = (-0.031) - (0.063)$, and significant at the 1% level. It is worth noting that the point estimate between the DID and the IV strategies need not be equal. First, the DID captures an average treatment effect, while the IV identifies a local average treatment effect (LATE) for families whose fertility decisions are affected by having same-sex children. These are generally

[7]The second sample, which uses sixth-grade students from 2018 and 2020, shows a slightly larger reduction in the coefficient, up to 20%. The third sample, which uses fifth-grade students from 2019 and 2020, goes in the opposite direction, with slight increases in the coefficient.

parents of higher socioeconomic status, for whom the decision to have another child is less costly. In addition, the IV estimate is based on families with at least two children and measures the effect of an additional child, whereas the DID compares families with one versus multiple children. Despite these conceptual differences, both results point in the same direction: having siblings negatively affected learning outcomes during school closures.

C. *Persistence of Effects and Impacts Beyond GPA*

The estimates discussed above are based on GPA measures and established that the relationship identified in Section III is causal. Now, I provide a more detailed analysis of the effects of having a sibling on GPA losses and include results after schools reopened for standardized exams and parental expectations regarding educational attainment.

GPA.— I have already shown in Figure 2 that students with siblings in grades 1 through 11 experienced larger learning losses than only children during school closures. I now examine how these effects vary across grades and persist after schools reopened. As shown in Figure 3, results for elementary school students reveal larger and more persistent effects, with those who have two or three siblings experiencing GPA losses of up to 0.04 SD even three years after reopening. Estimating the difference-in-differences specification separately by grade (Figure 3b) confirms that younger students were more affected than older ones.

How do these effects vary across different types of schools? I examine elementary school students (grades 1 to 6) separately by urban and rural location, and by public and private school status. Figure A.1 shows that the effects are similar across all groups. This consistency is informative, as these schools adopted different approaches to remote learning during closures. Students in rural areas had less access to computers and the internet, and their lessons relied more on radio than those of students in urban areas. However, in both settings, most communication between teachers and families occurred through cellphones and WhatsApp. The similarity of effects between public and private schools is also notable. While public schools primarily relied on asynchronous instruction, private schools might have been quicker to adopt synchronous virtual classes.

It is possible that the drop in performance is a reflection of students in larger families having a harder time doing homework assignment or sending their materials on time without that translating into actual learning losses. For that reason, I also look into actual measures of performance from standardized test scores that were taken in the school after they had reopened.

Standardized Exams.— To have a better assesment about the actual consequences in learning losses that having a sibling had, I look at effects on standardized exams, which are comparable across students of different schools and time. Unlike GPA, standardized exams are not available for every year and every grade. They were not taken during school closures, and in elementary school, only second and fourth grade students take it. In some cases this is done for the full population of students in the grade and in others only for a representative sub-sample.[8] For second grade and fourth grade exams, I can perform a DID estimation between 2019 and 2022, using tested students from a representative sample of the population. In fourth grade, I can also compare the full population of students using 2018 and 2024 national examinations. Table 4 shows the estimate for the effect of having siblings during school closures on standardized test scores. In 2nd grade, the effect is -0.034sd for mathematics and -0.023sd for reading. Losses in 4th grade are larger than those in second grade with up to -0.052sd for mathematics and -0.036sd for reading when looking at the same sample of years. Results using the national examinations of 2018 and 2024 point in the same direction. For fourth grade, results are even higher than -0.1sd for those who had 3 siblings. The larger magnitude in panel C is particularly interesting when noting panels A and C are testing the same cohort of students in the post period: those who were in first grade in 2021. That is, the effect in learning losses in standardized test scores is not only persistent but the gap increases with time after schools have reopened.

[8]Prior to COVID, the test was administered to the full population of second-grade students from 2007 to 2016, and then to representative subsamples in 2018, 2019, 2022, and 2023. The full population of fourth-grade students was tested in 2016, 2018, and 2024, while only representative subsamples were tested in 2019, 2022, and 2023.

These results highlight some important facts. First, although part of the losses in GPA seem to recover after schools reopened, actual performance in standardized exams seems to have an effect beyond school closures that similarly does not attenuates over time. Second, even students who had not started school yet in 2020 have experience learning losses. This is important as it speaks of the potential relevance of my research question to the loss of institutional support during Pre-K and K education. Those students can also be impacted by the presence of a sibling that may constraint household resources.

Expectations.— Along with the standardized examinations, parents were surveyed and asked about the maximum level of education they expected the student would obtain. As is common, these expectations tend to exceed actual attainment levels: Almost 80% of parents expect their children to obtain a higher education degree and 40% expect them to obtain a master's or PhD degree. In Table 4, I show significant reductions in these expectations among parents of students with siblings. Results are noisier in panels A and B for effects on expectations over higher education degrees, with generally negative point estimates but not always significant. However, Panel C, which uses a larger sample, shows more consistent results. Both expectations for higher education and for a graduate degree decline as a consequence of having a sibling, by an average of 2.5 and 1.7 percentage points, and by as much as 5.7 and 3.5 percentage points for those with three siblings—almost a 10% decline from baseline levels.

V. Mechanisms

I have shown that having a sibling had a negative effect on learning, and that this holds across different segments of the population in Peru. Moreover, the magnitude of the learning loss increases with the number of siblings, is larger for younger students, and extends to GPA, standardized test scores, and parental expectations. There are several plausible explanations for why siblings could negatively affect learning during school closures. In this section, I examine three potential mechanisms that may account for these patterns. First, I test if birth-order differences drive the results, since first-born children typically receive more parental attention. Second, I explore the possibility that siblings competed for limited material or technological resources, such as computers, or that they were disruptive to each other. Third, I assess whether the dilution of parental time and attention explains the observed effects, using heterogeneity by socioeconomic status and age, as well as a regression discontinuity design based on school starting-age cutoffs.

A. *Birth Order*

Research has shown positive effects on being the first-born on educational outcomes (Behrman and Taubman (1986), Price (2008)). An important channel for that is the difference in parental investment that each children receives at each stage of their lives, as first-born benefit from a period of undivided attention. To isolate the role of birth order, I restrict the sample to first-born children. This way, I compare the same birth-order children for both those with and without siblings.[9] In Figure 4a I show results of the effect of having a sibling only considering first-born children. The size of the effects are similar to those obtained when using all children in the sibling sample, which suggest that even though birth order may play a role at explaining different levels of education overall, it does not have a differential effect in learning losses due to school closures. Results in the following sections only include the oldest sibling in the sibling sample.

B. *Sibling disruptions*

During school closures, students not only spent more time at home but also relied on it as the primary environment for studying, reviewing lessons, and taking tests. Because learning took place within the household rather than at school, siblings rather than peers could have influenced this process by disrupting one another or by competing for access to technology and learning materials.

To test whether technological constraints explain the effects, I split the sample by

[9] When using the younger sibling instead, and comparing that to only children, results are consistent.

household access to computers and the internet. If competition for these resources were driving the results, I could expect households who lack those resources to not exhibit the negative effect of having a sibling. However, in panel C and D of Table 5, I show that the negative effects are present even in households with neither a computer or internet to access remote education easily. This suggests that the results are not driven by siblings competing for material resources.

Looking into siblings being distracting to each other, it is plausible for students who are closer in age to interact more and in that sense, to be more disruptive. I look at results by age gap in Figure 4b. Having a sibling 1-2 years apart has a similar effect as having a sibling 3-5 years apart which suggest that age gap is not a meaningful factor in explaining the effect of a sibling. However, when looking at siblings of the same age, this negative effect disappears and even becomes slightly positive. These confounds both sibling and parental mechanisms. Students of the same age are students attending the same grade which likely share classes and assignments.[10] On one hand, it is more likely that in this context having a sibling is beneficial as cooperation is more plausible, especially when most of these are twins. On the other hand, it is easier for parents to keep up with two children doing the same work than with two children doing different schoolwork. This leads to the final mechanism explored: parental time and attention.

C. Parental time and attention

I next examine whether parental time constraints explain the negative effects of having siblings during school closures. When schools closed, the role of teachers and schools in the education production function diminished, and the learning environment shifted almost entirely to the household. Parents therefore had to take a more active role and allocate additional time to support their children's learning. However, parents have limited time, and those with multiple children may have been more constrained and less able to dedicate sufficient attention to each child. I find this to be the most plausible mechanism driving the observed results.

[10]Because I do not have dates of birth for the full sample, age gap has been proxied by the difference in grades between siblings

If parental time is binding and having siblings dilutes this resource, I expect to see larger effects where parental time investment is both more necessary and more common. If parents do not devote much time to their children, those with and without siblings would be equally affected by the lack of support. In contrast, parents who usually invest more time in their children face a tradeoff when that time must be divided among several children. Consistent with this hypothesis, as shown in Figure 3b, the effects are larger for younger students, where parental involvement tends to be greater. Similarly, in Table 5, I find stronger sibling effects among families in the upper socioeconomic quartiles and weaker or null effects among those in the lowest quartile, where parental time investment is already limited. This pattern aligns with descriptive evidence from 2015 in Figure A.2a, which shows that parental time investment is substantially lower for households in the bottom part of the socioeconomic index distribution.[11]

In Table A.2, I also find larger effects among families whose parents held higher educational expectations and for students in the top quartile of baseline achievement. This pattern may reflect compensatory behavior, where parents allocate additional time to children who are performing worse. This interpretation contrasts with evidence from Giannola (2024), who finds that parents invest more in higher-achieving children and more so when constrained.

To further this analysis, I exploit the school starting age (SSA) cutoff in Peru to identify the spillover effects of younger siblings starting school on older siblings' academic performance. Under this rule, children who turn six years old by March 31 are required to enroll in first grade in that same academic year, while those who turn six after the cutoff must wait until the following year. This cutoff creates a sharp discontinuity in school entry timing that is otherwise unrelated to previous family characteristics. When a younger sibling starts school, they presumably require fewer hours of parental care, which frees up time for parents and may improve the academic outcomes of older siblings through increased parental attention. Panel A of Figure 5 shows that the

[11] Parents are asked whether they help their children with homework, discuss what they did in school, or clarify their questions. A standardized index of parental investment is constructed using these responses.

distribution of ages around the cutoff is smooth, and Panel B confirms that the first stage is sharp, indicating that the rule is strictly enforced. I use this discontinuity to estimate the effect of a younger sibling starting school, and therefore spending less time at home, on the GPA of older siblings who remain in school.

$$Y_{if} = \gamma_0 + \beta_1 ABOVE_{jf} + f(AgeCutoff_{jf}, ABOVE_{jf}) + \tau_t + \epsilon_{if} \tag{7}$$

Where Y_{if} is the standardized GPA of child i in family f, based on the running variable, $AgeCutoff_{jf}$, which is the age of the youngest sibling j by the cutoff date. The variable $ABOVE_{jf}$ captures the discontinuity around the cutoff and $f(\cdot)$ includes local linear controls for the running variable. τ_t captures year fixed effects.

When schools operate normally, being above the cutoff implies that students begin first grade and therefore spend less time at home. This reduction in childcare needs for parents may generate positive spillovers by increasing the time they have available for their other children already in school. I show these positive spillover effects in Figure 5 and in panel A of Table 6 for the optimal bandwidth[12] in column 1 and for different bandwidth sizes in columns 2 through 6. However, during school closures, all children remained at home before and after they reached school-starting age. In Panel B, I show that the positive spillovers from a younger sibling entering school disappear during closures. More precisely, I estimate a *Diff-in-RD*, comparing the regression discontinuity estimates during closures with those obtained when schools were open, using the following equation:

$$\begin{aligned} Y_{ift} = {} & \delta_0 + \beta_2 ABOVE_{jf} \cdot Closures_t \\ & + \alpha_0 ABOVE_{jf} + \alpha_1 Closures_t + f(AgeCutoff_{jf}, ABOVE_{jf}) + \tau_t + \epsilon_{ift} \end{aligned} \tag{8}$$

[12]Based on Calonico et al. (2017).

The effect of interest is captured by β_2. Effects for mathematics and reading are shown in panel C and F of Table 6. The negative effects of -0.012sd and -0.013sd in column 1 show the estimates for the optimal bandwidth for mathematics and reading respectively[13]. When schools are closed, and children have to stay at home rather than start school when they turn 6, parents cannot longer benefit from the increased time availability that this creates if schools were operating normally. This is reflected by the negative effects in panels C and F which reflect the loss of positive spillovers that older siblings would have otherwise benefited from.

To explore the mechanisms behind the positive spillovers from having a sibling at school rather than at home, I use reported measures of parental time investment from the surveys. Although these measures are not available during or before school closures, I can examine their spillover effects in 2022 and 2023, when schools had reopened. For these years, I construct an index of parental investment based on several questions about the frequency of activities such as explaining topics, helping with schedules, asking questions about school, etc. In Panel D of Figure 5, I find a small positive effect on the parental investment index for fourth-grade students when a younger sibling starts school rather than remains more time at home. Results in Table 7 also show increases in specific activities, with parents spending more time explaining topics and helping with schedules. However, although point estimates remain positive, most effects are no longer statistically significant when using the optimal bandwidth, except for "help with schedule." Because the sample for this measure is relatively small, part of the loss of significance may reflect limited statistical power.

Overall, the evidence points to parental time constraints as the main channel. Birth order and material competition do not explain the observed patterns, and although there is some indication of positive spillovers in specific cases such as twins, there is no consistent evidence that siblings were disruptive otherwise. The heterogeneity analysis and school starting-age discontinuity results are consistent with what would

[13]Because Panels A and B and panels D and E use different samples, the optimal bandwidth was estimated for each of them and the minimum was chosen for each pair.

be expected if parental time constraints were the mechanism behind sibling-related learning losses during school closures, when parents faced greater demands for childcare and educational support.

VI. External Validity

Do children with siblings also perform worse during closures in other countries? I provide evidence about this from two different sources. First, I use PISA scores from 2012 and 2022 to compare how the gap in learning losses widens between children who had siblings and those who did not in most developed and developing countries, particularly in those who experienced larger losses. Second, I provide evidence of a similar heterogeneity in India with data from Singh, Romero and Muralidharan (2022).

A. PISA

The Programme for International Student Assessment (PISA) is a worldwide study that measures the educational performance of 15-year-old students, It is conducted mainly in OECD countries but also includes some non-OECD participants. It is conducted every three years and the assesment is on skills in mathematics, reading, and science literacy. The last 5 tests were performed in 2009, 2012, 2015, 2018, and 2022, however, neither the 2015 or the 2018 tests ask about sibling status. Because of this, I use the most recent pre-covid test to compare how the gap has changed by sibling status between 2012 and 2022 in each country. In Figure 6a I show how the score distributions shifted leftward for both groups, but more markedly for those with siblings. To look closer at this I first estimate the learning losses among children with siblings in each country and children without siblings (only children). Then I estimate how different these learning losses are for both groups. My estimate is analogous to a simple 2x2 DID estimate without controls: $DID = \Delta_{2022-2012}Sibs - \Delta_{2022-2012}OC$. Like in our previous analysis, if children with siblings experienced larger losses, this difference will take a negative value.

Additionally, using data from UNESCO on length of school closures for each country, I estimate the amount of weeks each country had of full or partial closures. In Figure 6b I notice two patterns (i) most developed and developing countries are below the 0 line. This means that between 2012 and 2022, children with siblings also experienced larger losses, and (ii) the size of these larger losses is highly correlated with the duration of school closures. This provides some suggestive evidence that across different lengths of closures and policy measures to deal with school closures, children with siblings had a harder time adjusting to remote schooling.

B. India

I provide additional evidence from another country that, like Peru, experienced prolonged school closures: India. In Peru, standardized exams are not administered until one year after schools reopened, allowing time for learning recovery that may attenuate some of the effects. In contrast, in the setting studied by Singh, Romero and Muralidharan (2022), students were tested at the end of school closures. Children scored about 0.7sd lower in mathematics and 0.34sd lower in Tamil (the local language) in December 2021 compared to children of the same age and gender in the same villages in August 2019. The authors find greater learning losses among children whose mothers had not completed high school (12th grade).

To complement their analysis, in Table 8 I extend their results by adding a dimension of heterogeneity that is the focus of my paper: whether the children had siblings. Interestingly, I find substantially larger learning losses among children with siblings. The size of these are -0.11 standard deviations in mathematics and -0.065 in Tamil. As in the Peruvian case, the estimates are robust when controlling for heterogeneity by the mother's education and are larger relative to those differences.

VII. Conclusions

This paper has found evidence of a so far overlooked issue regarding family structure and school closures: that larger families struggled more to fill the role left by schools, that the losses caused by this are persistent, and that they are likely caused by parents being unable to substitute the role of teachers given time constraints that become more prevalent when having to attend multiple children. Even though children lose their peers, having a sibling is not beneficial in terms of their academic performance. It might however be beneficial in other dimensions such as emotional and mental health as suggests evidence from Hughes et al. (2023) and Lampis et al. (2023).

Descriptive evidence from international PISA test scores and from detailed data from India suggest that this pattern may be occurring in other countries as well, with larger effects in those that experienced longer school closures. Given the persistence of some of the effects found in Peru and the persistence of learning losses still to be recovered in other parts of the world, these results highlight the importance of considering family structure both for remediation programs but also for a better design of remote environments to avoid imposing larger costs to larger families. More broadly, my results speak about the important role that schools and educational services play in mitigating the costs of larger families due to their inability of adequately substitute them in the production of education.

REFERENCES

Angrist, Joshua D., and William N. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *The American Economic Review*, 88(3): 450–477. Publisher: American Economic Association.

Angrist, Joshua, Victor Lavy, and Analia Schlosser. 2010. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." *Journal of Labor Economics*, 28(4): 773–824. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Attanasio, Orazio, Raquel Bernal, Michele Giannola, and Milagros Nores. 2020. "Child Development in the Early Years: Parental Investment and the Changing Dynamics of Different Dimensions."

Attanasio, Orazio, Sarah Cattan, and Costas Meghir. 2022. "Early Childhood Development, Human Capital, and Poverty." *Annual Review of Economics*, 14(Volume 14, 2022): 853–892. Publisher: Annual Reviews.

Becker, Gary S., and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, 84(4, Part 2): S143–S162. Publisher: The University of Chicago Press.

Behrman, Jere R., and Paul Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics*, 4(3): S121–S145. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Black, Sandra E., and Paul J. Devereux. 2010. "Recent Developments in Intergenerational Mobility."

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 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, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2010. "Small Family, Smart Family?: Family Size and the IQ Scores of Young Men." *Journal of Human Resources*, 45(1): 33–58. Publisher: University of Wisconsin Press Section: Articles.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2017. "Rdrobust: Software for Regression-discontinuity Designs." *The Stata Journal*, 17(2): 372–404. Publisher: SAGE Publications.

Conley, Dalton, and Rebecca Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources*, 41(4): 722–737. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics*, 5(2): 29–57.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2016. "Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule." *Journal of Human Resources*, 51(4): 832–868. Publisher: University of Wisconsin Press Section: Articles.

Giannola, Michele. 2024. "Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Survey Experiment." *The Economic Journal*, 134(658): 671–727.

Goldhaber, Dan, Thomas J. Kane, Andrew McEachin, Emily Morton, Tyler Patterson, and Douglas O. Staiger. 2023. "The Educational Consequences of Remote and Hybrid Instruction during the Pandemic." *American Economic Review: Insights*, 5(3): 377–392.

Haelermans, Carla, Madelon Jacobs, Rolf van der Velden, Lynn van Vugt, and Sanne van Wetten. 2022. "Inequality in the Effects of Primary School Closures Due to the COVID-19 Pandemic: Evidence from the Netherlands." *AEA Papers and Proceedings*, 112: 303–307.

Houtenville, Andrew J., and Karen Smith Conway. 2008. "Parental Effort, School Resources, and Student Achievement." *Journal of Human Resources*, 43(2): 437–453. Publisher: University of Wisconsin Press Section: Articles.

Hughes, Claire, Luca Ronchi, Sarah Foley, Caoimhe Dempsey, Serena Lecce, and I.-Fam Covid Consortium. 2023. "Siblings in lockdown: International evidence for birth order effects on child adjustment in the Covid19 pandemic." *Social Development*, 32(3): 849–867. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/sode.12668.

Jack, Rebecca, Clare Halloran, James Okun, and Emily Oster. 2023. "Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts." *American Economic Review: Insights*, 5(2): 173–190.

Jakubowski, Maciej, Tomasz Gajderowicz, and Harry Anthony Patrinos. 2023. "Global learning loss in student achievement: First estimates using comparable reading scores." *Economics Letters*, 232: 111313.

Lampis, Valentina, Sara Mascheretti, Chiara Cantiani, Valentina Riva, Maria Luisa Lorusso, Serena Lecce, Massimo Molteni, Alessandro Antonietti, and Marisa Giorgetti. 2023. "Long-Lasting Effects of Changes in Daily Routine during the Pandemic-Related Lockdown on Preschoolers' Language and Emotional–Behavioral Development: A Moderation Analysis." *Children*, 10(4): 656. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

Lichand, Guilherme, and Carlos Alberto Doria. 2024. "The lasting impacts of remote learning in the absence of remedial policies: Evidence from Brazil." *Proceedings* *of the National Academy of Sciences*, 121(22): e2316300121. Publisher: Proceedings of the National Academy of Sciences.

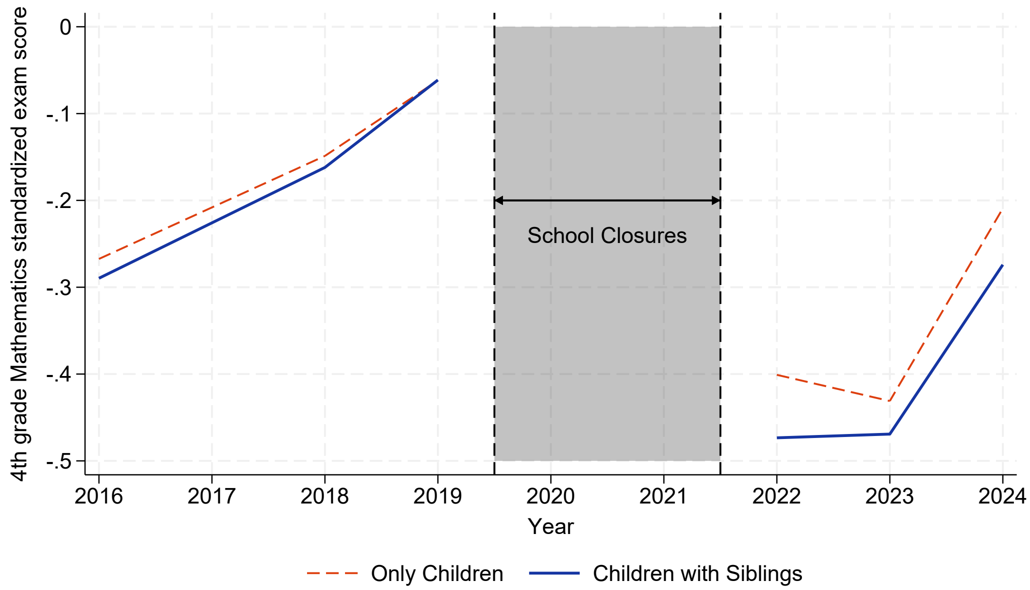
Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204. Publisher: ASA Website _eprint: https://doi.org/10.1080/07350015.2016.1227711.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review*, 103(4): 1289–1324.

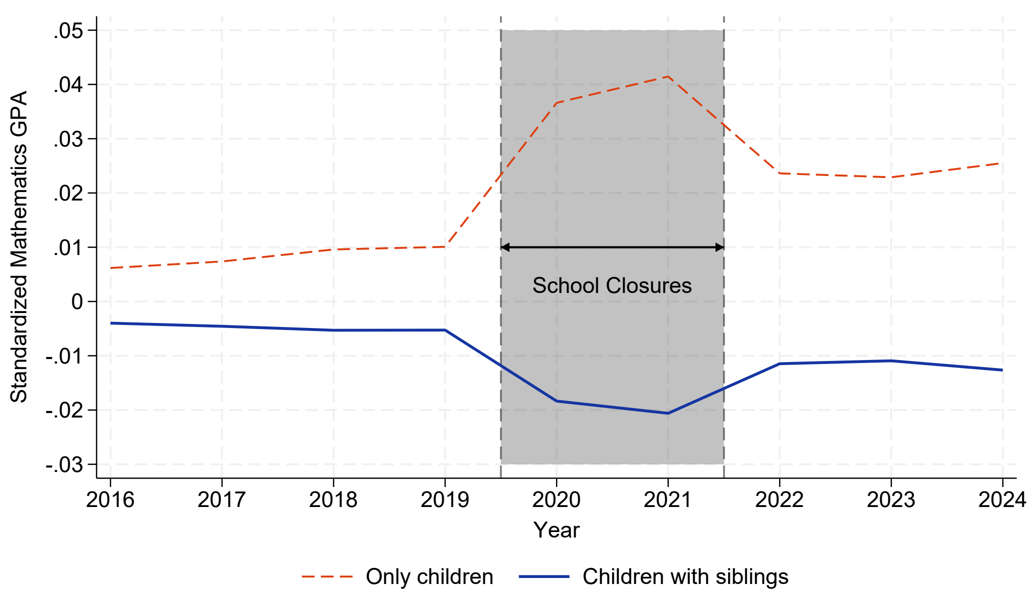
Price, Joseph. 2008. "Parent-Child Quality Time: Does Birth Order Matter?" *Journal of Human Resources*, 43(1): 240–265. Publisher: University of Wisconsin Press Section: Articles.

Singh, Abhijeet, Mauricio Romero, and Karthik Muralidharan. 2022. "Covid-19 Learning Loss and Recovery: Panel Data Evidence from India."
ΩÅslund and Grönqvist

Åslund, Olof, and Hans Grönqvist. 2010. "Family size and child outcomes: Is there really no trade-off?" *Labour Economics*, 17(1): 130–139.



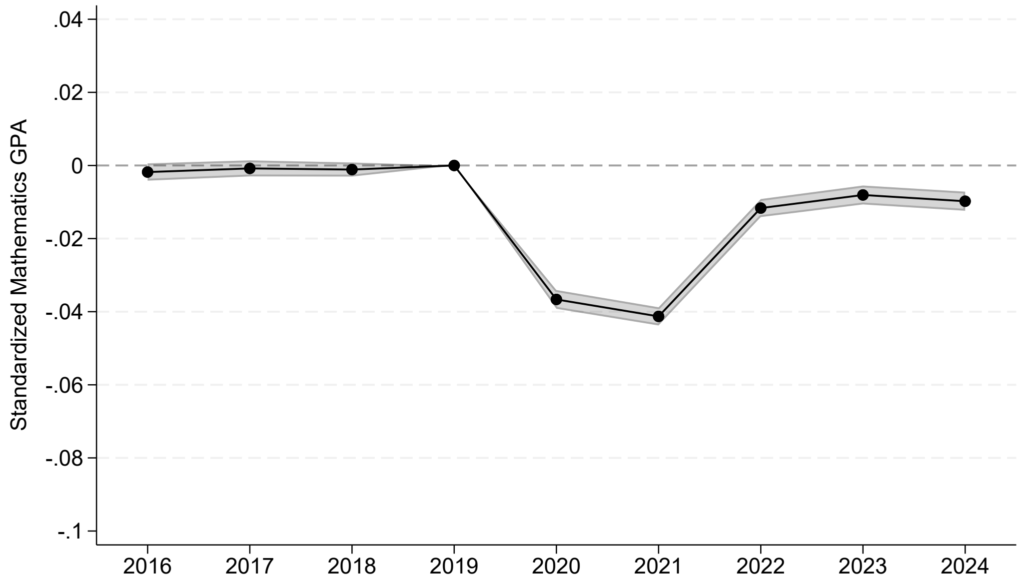
(a) Average Standardized Mathematics Exam in 4th grade



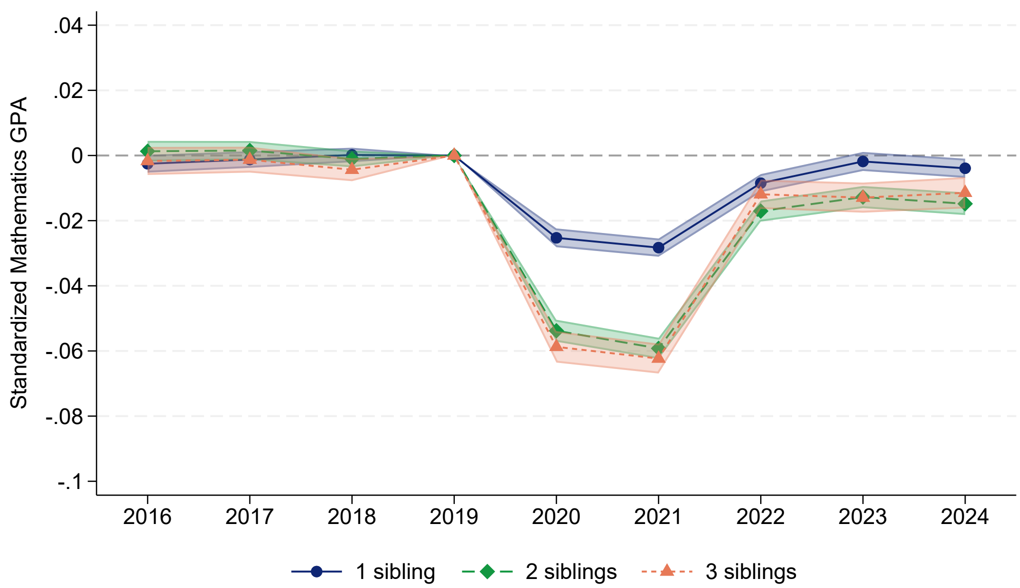
(b) Average GPA standardized within school-grade-year from 1st-6th grade

Figure 1. : Trends in education outcomes for only children and children with siblings

Notes: In *both panels*, the X-axis shows the year in which performance was measured, indicating the years of school closures (2020-2021). In the *top panel*, the Y-axis represents the average standardized exam score for fourth grade mathematics. These scores are comparable across time, with the year 2007 set as reference (mean 0 and standard deviation of 1). These exams are usually taken in November, close to the end of the academic year. The exam was not taken in 2017 or in 2020-2021. In the *bottom panel*, the Y-axis represents average GPA in mathematics from grades 1 through 11, standardized at the school-grade-year level. In both cases, results are show for the sample of children with and without siblings.



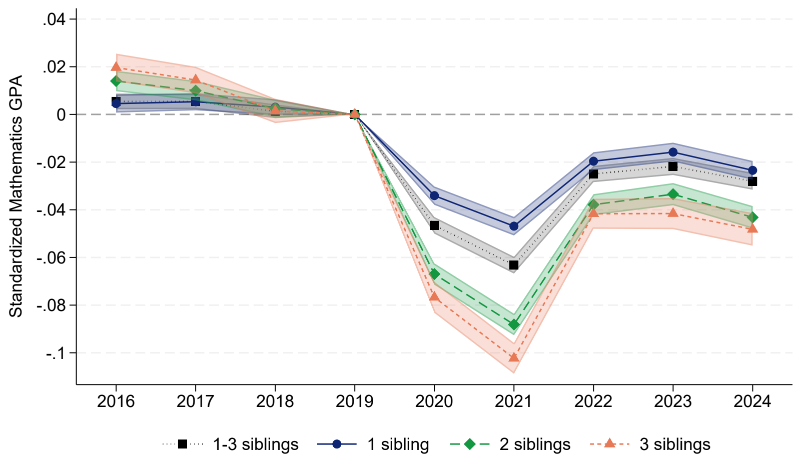
(a) Effect of having a sibling on mathematics GPA



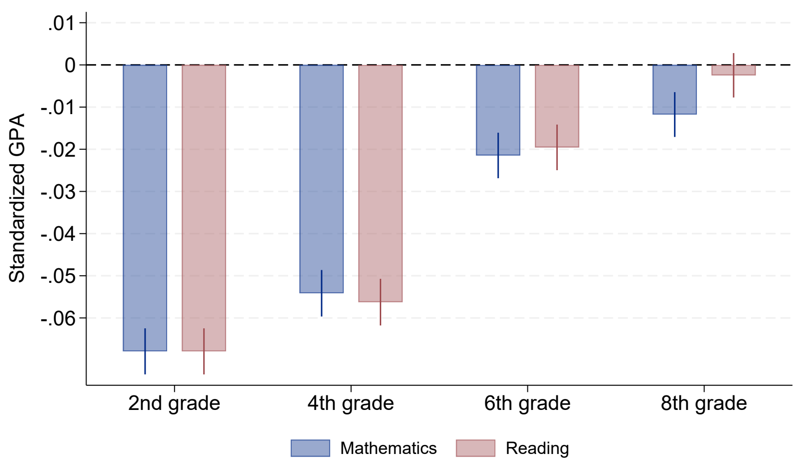
(b) Effect of having a sibling on mathematics GPA by number of siblings

Figure 2. : Event Study: Effect of having a sibling

Notes: The *top panel* shows the results from the event study using equation (2) for the GPA in mathematics from grades 1 through 11, standardized at the school-grade-year level. The results are measures for the sample of students with siblings relative to those without siblings. The Sib_i dummy takes the value of one if the student has siblings. In the *bottom panel* I do the same but only considering those with one, two or three siblings respectively compared to those without siblings. That is, estimates for students with one sibling compared to those without siblings are shown in blue circle markers, with two sibling compared to those without siblings are shown in green diamond markers, and with three sibling compared to those without siblings are shown in red with triangle markers. Estimated standard errors, reported in parentheses, are clustered at the primary school level. Shaded areas show the 95% confidence intervals.



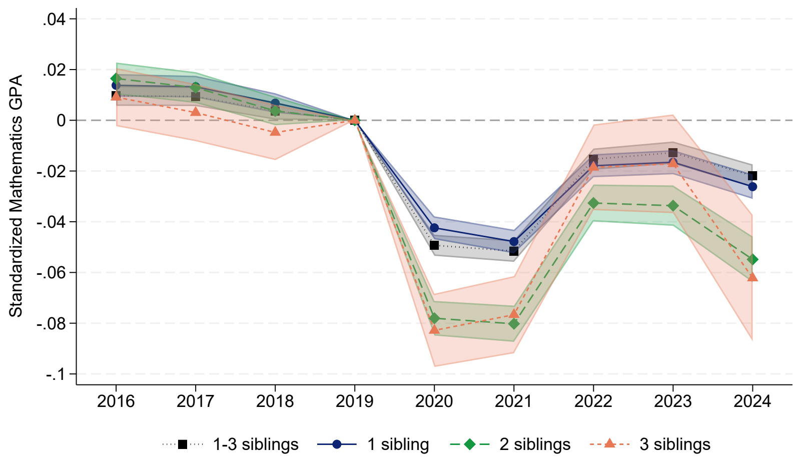
(a) Effect of having a sibling on mathematics GPA for elementary school children



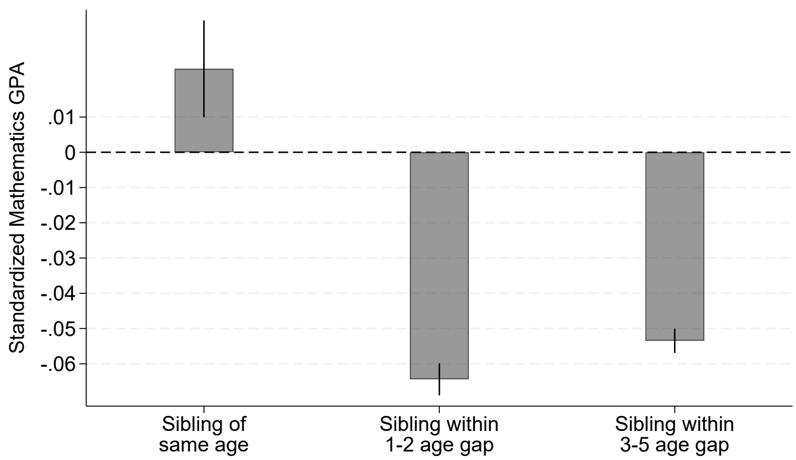
(b) Effect of having a sibling on mathematics GPA by grade

Figure 3. : Effect of having a sibling during elementary school

Notes: The *top panel* shows results from the event study using equation (2) for the GPA in mathematics from grades 1 through 6, standardized at the school-grade-year level. The Sib_i dummy takes the value of one if the student has siblings. Estimates for students with one, two or three siblings compared to those without siblings are shown in black square markers, with one sibling compared to those without siblings are shown in blue circle markers, with two sibling compared to those without siblings are shown in green diamond markers, and with three sibling compared to those without siblings are shown in red with triangle markers. Standard errors are clustered at the primary school level. Shaded areas show the 95% confidence intervals. The *bottom panel* shows results from the difference-in-difference estimates using equation (3) for the GPA in mathematics and reading standardized at the school-grade-year level for grades 2, 4, 6, and 8 for the period 2016-2021. The control variables (X_{ist}) used to account for other potential heterogeneity are student's sex, mother's age, mother's age at first birth, and mother's level of education. Standard errors are clustered at the primary school level. Vertical lines in each bar show the 95% confidence intervals.



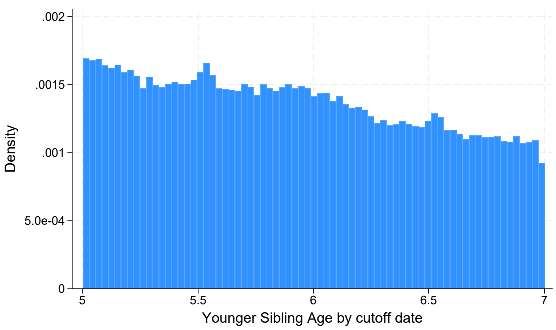
(a) Effect of having a sibling on mathematics GPA for first-born children



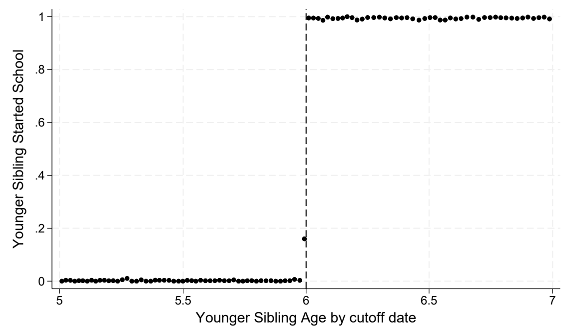
(b) Effect of having a sibling on mathematics GPA by age gap

Figure 4. : Mechanisms: Birth order and age gap

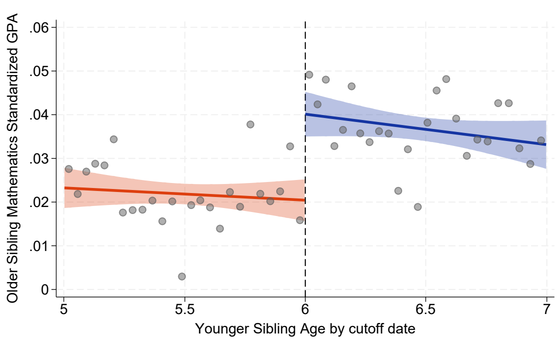
Notes: The *top panel* shows results from the event study using equation (2) for the GPA in mathematics from grades 1 through 6, standardized at the school-grade-year level. Only the sample of first-born children is consider for both those with or without siblings. The Sib_i dummy takes the value of one if the student has siblings. Estimates for students with one, two or three siblings compared to those without siblings are shown in black square markers, with one sibling compared to those without siblings are shown in blue circle markers, with two sibling compared to those without siblings are shown in green diamond markers, and with three sibling compared to those without siblings are shown in red with triangle markers. Standard errors are clustered at the primary school level. Shaded areas show the 95% confidence intervals. The *bottom panel* shows results from the difference-in-difference estimates using equation (3) for the GPA in mathematics standardized at the school-grade-year level for the period 2016-2021. Three samples were used when comparing to those without siblings: those with a sibling of the same age, with a sibling within 1-2 years of age, and with a sibling within 3-5 years of age. The control variables (X_{ist}) used to account for other potential heterogeneity are student's sex, mother's age, mother's age at first birth, and mother's level of education. Standard errors are clustered at the primary school level. Vertical lines in each bar show the 95% confidence intervals.



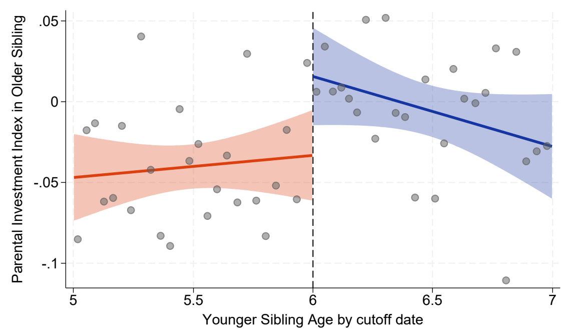
(a) Distribution of younger sibling's ages around cutoff date



(b) First Stage



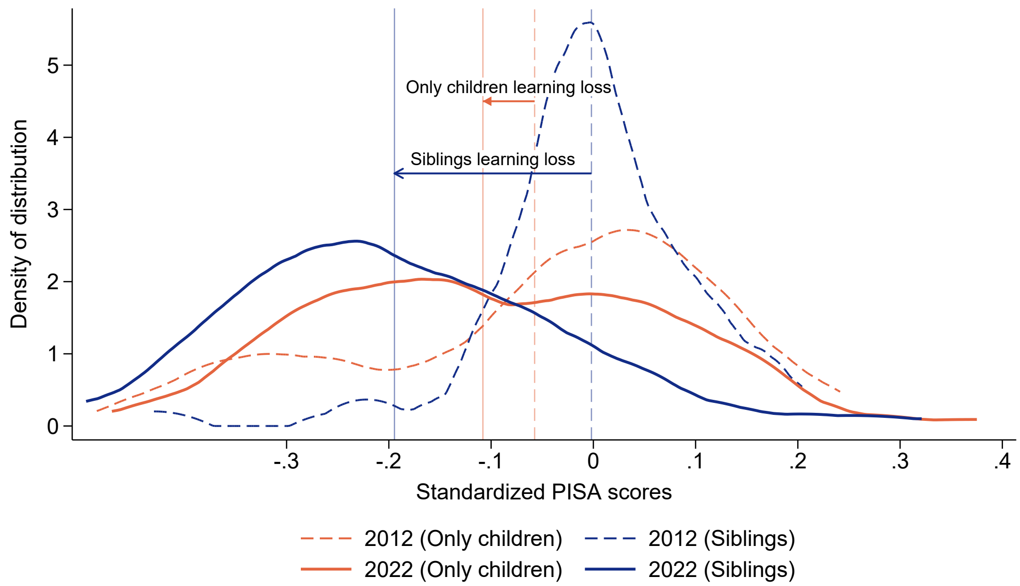
(c) Spillover effect on mathematics GPA



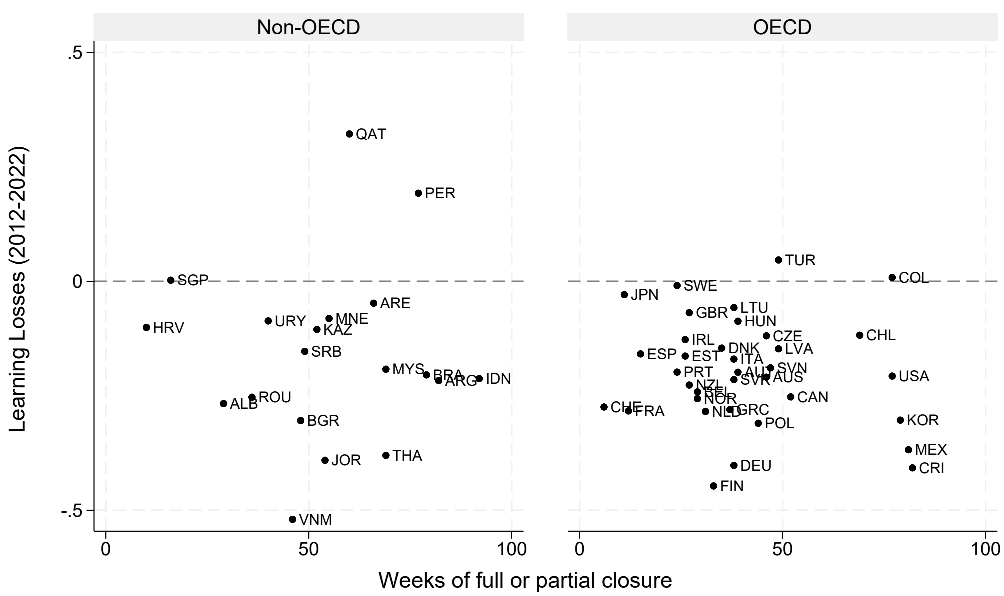
(d) Spillover effect on parental investment

Figure 5. : Regression Discontinuity during normal school operation: Spillover effect of having a younger sibling start school

Notes: This figure shows results from the regression discontinuity described in equation . The *top left panel (a)* shows the distribution of the running variable: age of the younger sibling around the starting age cutoff by March 31st. The distribution is smooth around the cutoff which provides evidence of no manipulation. The *top right panel (b)* shows the first stage. The Y-axis is the percentage of students who enroll in first grade during that year. The sharp discontinuity shows the almost perfect compliance of the school-starting age rule. The *bottom left panel (c)* shows the estimate of the regression discontinuity equation for the standardized mathematics GPA of the older sibling for years when schools were open (2016-2019, 2022-2023). The *bottom right panel (d)* shows the estimate of the regression discontinuity equation for the investment index reported by parent's of fourth grade older siblings for the years 2022-2023. The variable was not available in years 2016-2019. For both panels c and d, the bandwidth used was 365 days. In table Table 6, I show results for smaller bandwidths, for reading and for the *difference-in-RD* estimate comparing years of closures with years when schools were open. In Table 7, I show results for smaller bandwidths for the investment index and three of its components. Shaded areas show the 95% confidence intervals.



(a) Learning gaps in Mathematics between 2012 and 2022 for only children and children with siblings



(b) Change in learning gaps by duration of school closure for OECD and Non-OECD countries.

Figure 6. : International evidence of larger learning losses for children with siblings

Notes: The *top panel* shows the distribution of average country PISA scores in mathematics for the 'only children' (red) and 'sibling' (blue) sample for 2012 and 2022 in dashed and solid lines respectively. Individual scores were first standardized using the mean a standard deviation from 2012 and then averaged at the country-year level for both children with and without siblings. The vertical lines show the mean country averages for each of the distributions and the horizontal arrows show the learning loss from 2012 to 2022. The *bottom panel* the Y-axis shows the estimated learning loss for each country as $DID = \Delta_{2022-2012}Sibs - \Delta_{2022-2012}OC$. and the X-axis shows the number of weeks the country had full or partial school closures.

Table 1—: Descriptive Statistics

	Only children	1 sibling	2 siblings	3 siblings
	(1)	(2)	(3)	(4)
Sample share	0.38	0.32	0.20	0.10
Panel A: School and household characteristics				
Urban	0.79	0.79	0.71	0.58
Public School	0.67	0.70	0.80	0.89
Class Size	24.44	24.61	23.46	21.34
Mother with complete secondary	0.39	0.39	0.36	0.28
Panel B: Academic characteristics				
Grade promotion	0.96	0.97	0.96	0.94
Standardized GPA - Mathematics	0.02	0.05	-0.02	-0.10
Standardized Exam - Reading	0.02	0.05	-0.02	-0.10
Observations	1,238,826	1,079,501	683,613	316,529
Panel C: Academic Performance (2nd grade)				
Standardized Exam - Mathematics	0.65	0.75	0.63	0.43
Standardized Exam - Reading	0.77	0.81	0.65	0.44
Observations	819,485	702,368	432,584	193,909
Panel D: Household Characteristics (2nd grade)				
Socio-Economic Index	0.11	0.08	-0.15	-0.44
Internet	0.34	0.32	0.25	0.16
PC	0.35	0.35	0.28	0.19
Education expectation: 5-year college	0.80	0.81	0.75	0.66
Observations	90,998	93,439	58,021	24,544

Notes: This table shows descriptive statistics for four samples of the student population: students without siblings (only children) and students with one, two or three siblings. The first row shows the sample share. The statistics correspond to students from grades 1 through 6 for the year 2016. The first row shows the distribution of students between these four groups. Then, *panel A* shows statistics using on school and mother's education using administrative enrollment data. *Panel B* uses the same data to show academic characteristics of the students. *Panel C* reports baseline academic performance from standardized exams students took in 2nd grade from 2007-2016. *Panel D* reports information from baseline survey responses of second grade students in 2015-2016

Table 2—: DID estimates of having a sibling controlling for confounding heterogeneity

	GPA				
	(1)	(2)	(3)	(4)	(5)
Panel A: 2019–2020 6th grade DID with 2nd grade baseline (Mathematics GPA)					
Sibling x Post	-0.051***	-0.048***	-0.048***	-0.047***	-0.047***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Observations	193,960	193,960	193,960	193,960	193,960
Panel B: 2019–2020 6th grade DID with 2nd grade baseline (Reading GPA)					
Sibling x Post	-0.034***	-0.031***	-0.031***	-0.031***	-0.031***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
Observations	192,367	192,367	192,367	192,367	192,367
Mother's Ed		X	X	X	X
Mother's age			X	X	X
SES				X	X
Score					X

Notes: This table shows results from the *difference-in-difference* estimates using equation (3) for sixth grade students in 2019 and 2020 using second grade baseline characteristics. *Panel A* shows results for GPA in mathematics and *panel B* for reading, both standardized at the school-grade-year level. Column 1 shows the estimates without any controls for heterogeneous effects, analogous to what is done in equation (1). Column 2 adds a dummy for the mother having higher education. Column 3 adds dummies for both mother's age and age at first birth being above 30. Column 4 adds a dummy for baseline socioeconomic status being above the mean. Column 5 adds a dummy for baseline standardized test scores being above the mean. Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.

Table 3—: Effect of Family Size on mathematics GPA

	OLS (no controls)	OLS (controls)	First Stage	Second Stage	Observations
Panel A: Pre-Covid (2018-2019)					
Instrument: first two children same sex			0.050*		3,300,349
(Sample: first and second children in families with two or more births)			(0.001)		
Number of children in family	−0.081*	−0.070*		0.063*	
	(0.001)	(0.001)		(0.022)	
Panel B: Covid (2020-2021)					
Instrument: first two children same sex			0.047*		2,809,126
(Sample: first and second children in families with two or more births)			(0.001)		
Number of children in family	−0.119*	−0.101*		−0.031	
	(0.001)	(0.001)		(0.025)	
Panel C: Difference between COVID and Pre-COVID estimate					
				-.094***	
				(0.033)	

Notes: This table shows results from the *same-sex IV* estimates using equations (5) and (6) using the sample of first and second born children in families with two or more births. *Panel A* shows estimates during years when schools were open and *panel B* shows estimates during years of school closures. Column 1 and 2 show the simple OLS estimates with no controls and controlling for demographics. Column 2 adds a dummy for the mother having higher education. Column 3 adds dummies for both mother's age and age at first birth being above 30. Column 4 adds a dummy for baseline socioeconomic status being above the mean. Column 5 adds a dummy for baseline standardized test scores being above the mean. Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.

Table 4—: DID estimates of having a sibling on Standardized Exams and Expectations

	Has a sibling	1 sibling	2 siblings	3 siblings
	(1)	(2)	(3)	(4)
Panel A: 2nd grade students (2019, 2022)				
Mathematics	-0.034***	-0.024**	-0.072***	-0.079***
	(0.009)	(0.009)	(0.012)	(0.019)
Reading	-0.023***	-0.018**	-0.061***	-0.064***
	(0.008)	(0.009)	(0.011)	(0.018)
Max Expectation: 5-year college	-0.007**	-0.015***	-0.007	0.012
	(0.003)	(0.004)	(0.005)	(0.009)
Max Expectation: Graduate level	-0.015***	-0.015***	-0.029***	-0.011
	(0.005)	(0.005)	(0.007)	(0.011)
Observations	226,592	153,784	110,671	77,517
Panel B: 4th grade students (2019, 2022)				
Mathematics	-0.052***	-0.028***	-0.081***	-0.149***
	(0.009)	(0.010)	(0.012)	(0.018)
Reading	-0.036***	-0.010	-0.066***	-0.131***
	(0.009)	(0.010)	(0.012)	(0.017)
Max Expectation: 5-year college	-0.005	-0.009**	0.004	-0.024***
	(0.004)	(0.004)	(0.006)	(0.009)
Max Expectation: Graduate level	-0.012**	-0.010*	-0.009	-0.027***
	(0.005)	(0.006)	(0.007)	(0.010)
Observations	183,473	121,339	90,833	63,819
Panel C: 4th grade students (2018,2024)				
Mathematics	-0.040***	-0.029***	-0.065***	-0.102***
	(0.004)	(0.005)	(0.006)	(0.009)
Reading	-0.062***	-0.050***	-0.096***	-0.132***
	(0.004)	(0.005)	(0.006)	(0.009)
Max Expectation: 5-year college	-0.025***	-0.019***	-0.041***	-0.057***
	(0.002)	(0.002)	(0.003)	(0.004)
Max Expectation: Graduate level	-0.017***	-0.015***	-0.032***	-0.035***
	(0.002)	(0.003)	(0.003)	(0.005)
Observations	758,901	533,684	399,912	299,759

Notes: This table shows results from the *difference-in-difference* estimates using equation (3) for second and fourth grade students controlling for the mother having higher education for both mother's age and age at first birth being above 30. The first column shows the estimate of having a sibling when compared to those without a sibling. The second to fourth column show the estimate of having one, two a three siblings respectively when compared to those without a sibling. Each panel shows results for standardized exam scores in mathematics and reading as well as for reported parental expectations on the student obtaining a 5-year college or a graduate degree. *Panel A* shows results for representative sample of second grade students who take the exam in 2019 (pre) and 2022 (post). *panel B* shows results for representative sample of fourth grade students who take the exam in 2019 (pre) and 2022 (post). *panel C* shows results for the universe of fourth grade students who take the exam in 2018 (pre) and 2024 (post). Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.

Table 5—: DID estimates of having a sibling on mathematics GPA by baseline resources

	1 sibling	2 siblings	3 siblings
	(1)	(2)	(3)
Panel A: Low SES Households (Q1)			
	-0.001	-0.024***	-0.058***
	(0.008)	(0.009)	(0.014)
Observations	282,258	238,260	201,756
Panel B: High SES Households (Q4)			
	-0.037***	-0.065***	-0.134***
	(0.008)	(0.014)	(0.034)
Observations	249,817	193,842	173,621
Panel C: Households with neither PC nor Internet			
	-0.019***	-0.049***	-0.083***
	(0.006)	(0.008)	(0.012)
Observations	468,143	381,437	322,583
Panel D: Households with both PC and Internet			
	-0.028***	-0.056***	-0.076***
	(0.007)	(0.011)	(0.023)
Observations	371,603	294,298	261,885

Notes: This table shows results from the *difference-in-difference* estimates using equation (3) for the mathematics GPA standardized at the school-grade-year level for students in sixth, seventh and ninth grade for which baseline characteristics from standardized test scores and socioeconomic characteristics was available. Column 1 through 3 show the estimate of having one, two a three siblings respectively when compared to those without a sibling. I identified four subsamples of students for grades that were tested before and during closures that also had baseline characteristics: sixth grade students in 2019 and 2020 with baseline characteristics from second grade standardized exams in 2015 and 2016, sixth grade students in 2018 and 2020 with baseline characteristics from fourth grade standardized exams in 2016 and 2018, seventh grade students in 2019 and 2021 with baseline characteristics from fourth grade standardized exams in 2016 and 2018, and ninth grade students in 2019 and 2020 with baseline characteristics from second grade standardized exams in 2018 and 2019. *Panel A* shows results for students in the first (lowest) quartile of socioeconomic status. *Panel B* shows results for students in the fourth (highest) quartile of socioeconomic status. *Panel C* shows results for students with no computer or internet at home. *Panel D* shows results for students with both computer and internet at home. Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.

Table 6—: Effects of younger sibling starting school on older sibling GPA

	Standardized mathematics GPA					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Spillover effects on mathematics when schools are open						
Younger sibling goes to school	0.025***	0.025**	0.021**	0.024***	0.020***	0.019***
	(0.007)	(0.012)	(0.009)	(0.007)	(0.005)	(0.004)
Observations	316,576	103,214	206,349	307,449	601,655	1,203,142
Bandwidth (days)	93	30	60	90	180	365
Panel B: Spillover effects on mathematics when schools are closed						
Younger sibling goes to school	0.006	0.001	0.011	0.003	-0.001	0.002
	(0.012)	(0.022)	(0.015)	(0.013)	(0.009)	(0.006)
Observations	101,002	32,857	65,643	97,896	191,203	378,757
Bandwidth (days)	93	30	60	90	180	365
Panel C: Diff-in-RD estimate of spillover on mathematics (closed - open)						
Younger sibling goes to school	-0.012*	-0.004	-0.018**	-0.012	-0.012**	-0.014***
	(0.007)	(0.013)	(0.009)	(0.007)	(0.005)	(0.004)
Observations	417,578	136,071	271,992	405,345	792,858	1,581,899
Bandwidth (days)	93	30	60	90	180	365
Panel D: Spillover effects on reading when schools are open						
Younger sibling goes to school	0.016**	0.018	0.016*	0.016**	0.020***	0.019***
	(0.007)	(0.013)	(0.009)	(0.007)	(0.005)	(0.004)
Observations	334,542	101,752	203,526	303,286	593,439	1,186,652
Bandwidth (days)	100	30	60	90	180	365
Panel E: Spillover effects on reading when schools are closed						
Younger sibling goes to school	0.005	0.019	0.012	0.002	0.004	-0.002
	(0.012)	(0.022)	(0.016)	(0.013)	(0.009)	(0.006)
Observations	107,418	32,492	65,143	97,185	189,831	376,132
Bandwidth (days)	100	30	60	90	180	365
Panel F: Diff-in-RD estimate of spillover on reading (closed - open)						
Younger sibling goes to school	-0.013*	-0.004	-0.015*	-0.011	-0.016***	-0.012***
	(0.007)	(0.013)	(0.009)	(0.007)	(0.005)	(0.004)
Observations	441,960	134,244	268,669	400,471	783,270	1,562,784
Bandwidth (days)	100	30	60	90	180	365

Notes: This table shows the estimates from the regression discontinuities using equations (5) and (8). The running variable is the younger sibling age at the school-starting age date cutoff (March 31st). Results show the effect of the younger sibling starting school on the older sibling's mathematics and reading GPA standardized at the school-grade-year level. *Panel A and D* show results for mathematics and reading respectively for the years in which schools were open (2016-2019 and 2022-2023). *Panel B and E* show results for mathematics and reading respectively for the years in which schools were closed (2020-2021). *Panel C and F* show the *Diff-in-RD* estimate using equation (8). Column 1 shows results for the optimal bandwidth and columns 2 through 6 for different bandwidths between 30 and 365 days. The optimal bandwidth used in *panels A-C* or *panels D-F* the minimum optimal bandwidth of between *panels A-B* or *panels D-E*. ***p<0.01, **p<0.05, *p<0.1.

Table 7—: Effects of younger sibling on parental time investment in older sibling

	Standardized mathematics GPA					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Spillover effects on parental time investment index						
Younger sibling	0.019	-0.039	-0.019	0.001	0.024	0.032**
goes to school	(0.022)	(0.043)	(0.031)	(0.025)	(0.018)	(0.013)
Observations	34,524	8,675	17,315	25,774	50,183	102,383
Bandwidth (days)	123	30	60	90	180	365
Panel B: Spillover effects on parental time investment (explains topics)						
Younger sibling	0.035	-0.071	-0.029	0.008	0.038**	0.035***
goes to school	(0.022)	(0.044)	(0.032)	(0.026)	(0.018)	(0.013)
Observations	31,752	7,996	15,984	23,743	46,152	94,166
Bandwidth (days)	123	30	60	90	180	365
Panel C: Spillover effects on parental time investment (helps with schedule)						
Younger sibling	0.040*	-0.040	-0.010	0.017	0.055***	0.048***
goes to school	(0.022)	(0.044)	(0.031)	(0.026)	(0.018)	(0.013)
Observations	31,458	7,924	15,819	23,523	45,730	93,197
Bandwidth (days)	123	30	60	90	180	365
Panel D: Spillover effects on parental time investment (search for information)						
Younger sibling	0.002	-0.004	-0.021	-0.007	0.002	0.015
goes to school	(0.023)	(0.045)	(0.032)	(0.026)	(0.019)	(0.013)
Observations	31,605	7,980	15,905	23,626	45,955	93,620
Bandwidth (days)	123	30	60	90	180	365

Notes: This table shows the estimates from the regression discontinuities using equation (5). The running variable is the younger sibling age at the school-starting age date cutoff (March 31st). Results show the effect of the younger sibling starting school on the parental time investment index for the older sibling in fourth grade in the year 2022 and 2023, which is when they were surveyed. *Panel A* shows results for the overall parental time investment index standardized with mean 0 and standard deviation of 1. *Panel B* shows results for wether parents explain topics the student does not understand. *Panel C* shows results for wether parents help students make a schedule or set a time for studying. *Panel D* shows results for wether parents help students search for material or additional information for their homework. Column 1 shows results for the optimal bandwidth and columns 2 through 6 for different bandwidths between 30 and 365 days. ***p<0.01, **p<0.05, *p<0.1.

Table 8—: Learning loss between August 2019 and December 2021 in India - Based on Singh, Romero and Muralidharan (2022)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Math score (in SD)						Tamil score (in SD)					
Wave 1 (Dec 2021)	-.73***	-.74***	-.76***	-.75***	-.68***	-.71***	-.35***	-.35***	-.37***	-.38***	-.32***	-.34***
	(.031)	(.038)	(.042)	(.049)	(.037)	(.046)	(.02)	(.023)	(.027)	(.029)	(.024)	(.031)
Male × Dec 21		.023						-.0074				
		(.041)						(.022)				
Mother Edu: Gr. 9-11 × Dec 21			.019			.021			.0015			.004
			(.053)			(.053)			(.03)			(.03)
Mother Edu: Gr. 12+ × Dec 21			.09*			.084*			.06**			.057**
			(.049)			(.049)			(.025)			(.025)
SES Decile × Dec 21				.0046						.0061		
				(.0075)						(.0039)		
Has Siblings (2-10 yrs old) × Dec 21					-.11**	-.10**					-.065**	-.062**
					(.041)	(.041)					(.025)	(.026)
N. of obs.	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083
R-squared	.33	.33	.33	.33	.33	.33	.31	.31	.31	.31	.31	.31

Notes: This table is based on Table 1 of Singh, Romero and Muralidharan (2022), but adding two columns per panel to include the heterogeneity by sibling. Observations across panels have been adjusted to information in this new variable. The original table estimates learning losses between August 2019 and December 2021 from COVID-19 for mathematics and Tamil ((the local language) using equation $Y_{it} = \alpha_v + \beta_1 Dec2021_t + \beta_2 X_{it} + \epsilon_{it}$ where X_{it} includes demographic characteristics such as sex, mother's education, and socioeconomic status interacted with $Dec2021$. Their results are shown in columns 1 through 4 and 7 through 10. In columns 5, 6, 11, and 12 I include the estimate on learning losses of having a sibling (2-10 years old) and similarly find heterogeneity that is 15%-20% of the total losses. ***p<0.01, **p<0.05, *p<0.1..

Appendix: NOT FOR PUBLICATION

Appendix A: Additional Tables and Figures

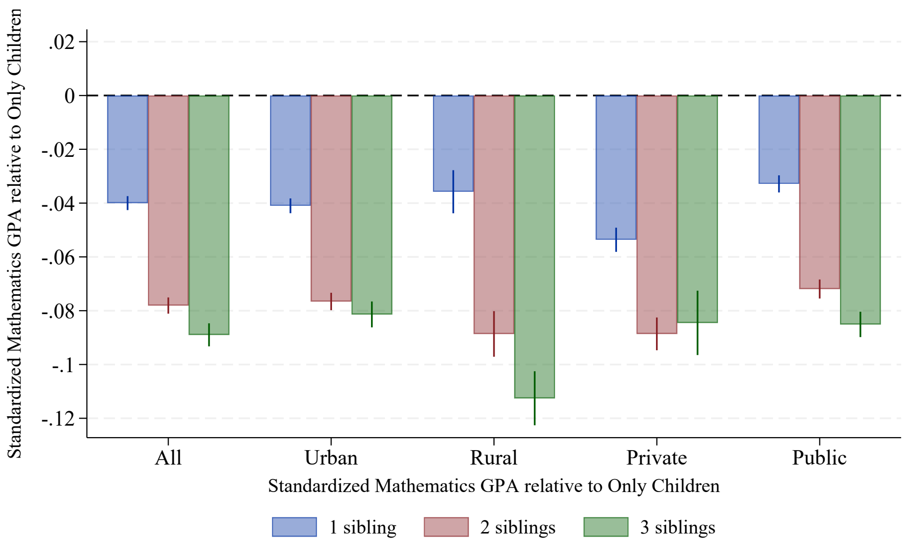
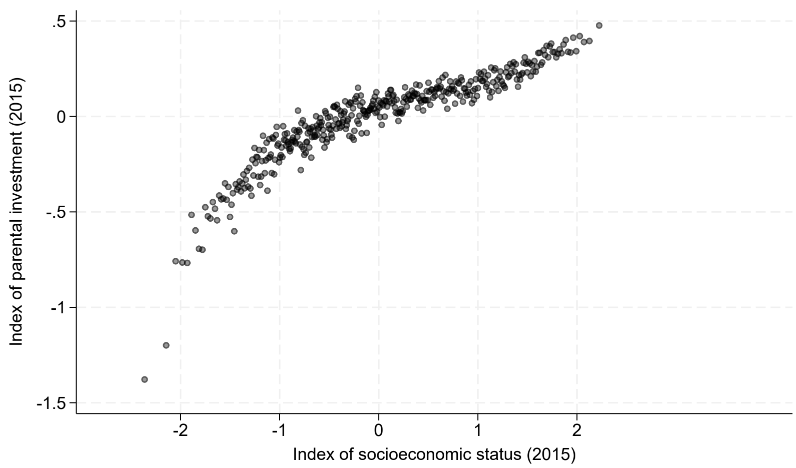
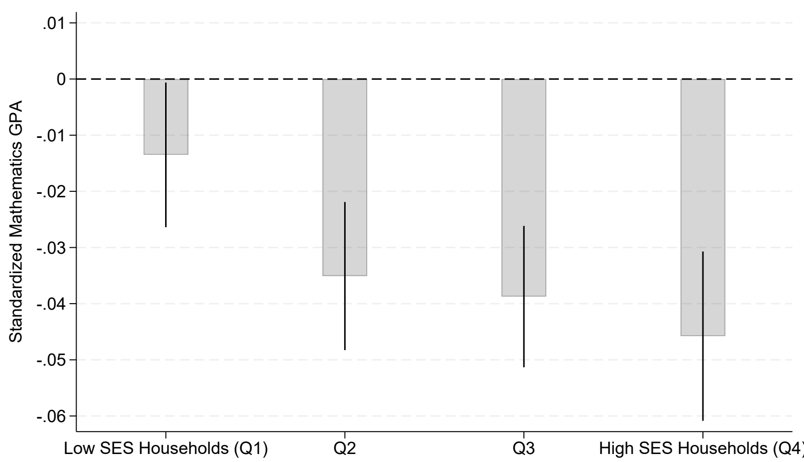


Figure A.1. : Effect of having a sibling on mathematics GPA by school characteristics

Notes: This figure shows results from the *difference-in-difference* estimates using equation (3) for the mathematics GPA standardized at the school-grade-year level for different samples of the population: all students and student in urban schools, rural schools, private schools and public schools. Each column shows results for the effect of having one, two or three siblings respectively compared to those without siblings. Standard errors are clustered at the primary school level. Vertical lines in each bar show the 95% confidence intervals.



(a) Relationship between parental SES and time investment



(b) Heterogeneity of sibling effect by socio-economic status

Figure A.2. : Parental investment and socio-economic status

Notes: The *top panel* shows the relationship between parental time investment index (Y-axis) and socioeconomic status index (X-axis) for the year 2015 in which the parental investment index was observed. The *bottom panel* shows results from the *difference-in-difference* estimates using equation (3) for the mathematics GPA standardized at the school-grade-year level for students in sixth, seventh and ninth grade for which baseline characteristics from standardized test scores and socioeconomic characteristics was available as described in the footnote of Table 5.

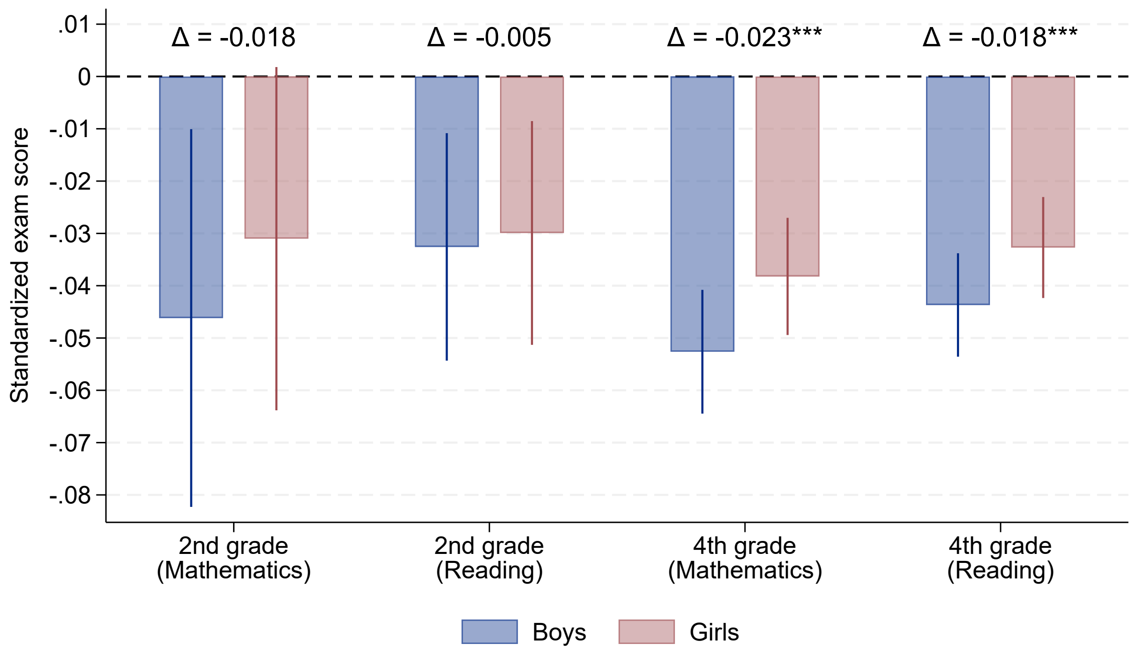


Figure A.3. : Heterogeneous effects on standardized test scores by gender

Notes: This figure shows results from the *difference-in-difference* estimates of having a sibling using equation (3) for standardized exam scores in mathematics and reading for second and fourth grade students. In each case I estimate the effect for the sample of boys and girls and show the difference on top of the bars. Standard errors are clustered at the primary school level. Vertical lines in each bar show the 95% confidence intervals. ***p¡0.01, **p¡0.05, *p¡0.1

Table A.1—: DID estimates of having a sibling controlling for confounding heterogeneity

	Mathematics GPA					Reading GPA				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 2019-2020 6th grade DID with 2nd grade baseline										
Sibling x Post	-0.051***	-0.048***	-0.048***	-0.047***	-0.047***	-0.034***	-0.031***	-0.031***	-0.031***	-0.031***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
Mother Ed. x Post		-0.045***	-0.044***	-0.024*	-0.026**		-0.049***	-0.048***	-0.036***	-0.038***
		(0.012)	(0.012)	(0.013)	(0.013)		(0.012)	(0.012)	(0.013)	(0.013)
Age x Post			-0.005	-0.005	-0.003			-0.006	-0.005	-0.003
			(0.017)	(0.017)	(0.017)			(0.017)	(0.017)	(0.017)
Age 1st x Post			0.002	-0.002	-0.002			0.003	0.001	0.000
			(0.010)	(0.010)	(0.010)			(0.011)	(0.011)	(0.011)
2nd grade SES x Post				-0.047***	-0.052***				-0.022**	-0.029***
				(0.011)	(0.011)				(0.011)	(0.011)
2nd grade Score x Post					0.146***					0.151***
					(0.014)					(0.014)
Observations	193,960	193,960	193,960	193,960	193,960	192,367	192,367	192,367	192,367	192,367
Panel B: 2018-2020 6th grade DID with 2nd & 4th grade baseline										
Sibling x Post	-0.042***	-0.040***	-0.036***	-0.035***	-0.033***	-0.044***	-0.042***	-0.038***	-0.038***	-0.035***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mother Ed. x Post		-0.015***	-0.011*	-0.004	-0.058***		-0.009	-0.004	0.001	-0.050***
		(0.006)	(0.006)	(0.007)	(0.007)		(0.006)	(0.006)	(0.007)	(0.007)
Age x Post			-0.004	-0.005	0.006			0.009	0.009	0.017*
			(0.010)	(0.010)	(0.010)			(0.010)	(0.010)	(0.010)
Age 1st x Post			0.021***	0.019***	0.016***			0.019***	0.017***	0.017***
			(0.006)	(0.006)	(0.006)			(0.006)	(0.006)	(0.006)
4th grade SES x Post				-0.016***	-0.060***				-0.010*	-0.051***
				(0.006)	(0.006)				(0.006)	(0.006)
4th grade Score x Post					0.231***					0.212***
					(0.006)					(0.006)
2nd grade Score x Post					0.104***					0.106***
					(0.008)					(0.008)
Observations	572,522	572,522	572,522	572,522	529,994	565,811	565,811	565,811	565,811	523,853
Panel C: 2019-2020 5th grade DID with 4th grade baseline										
Sibling x Post	-0.038***	-0.037***	-0.041***	-0.040***	-0.039***	-0.032***	-0.032***	-0.037***	-0.037***	-0.037***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Mother Ed. x Post		-0.053***	-0.056***	-0.043***	-0.074***		-0.051***	-0.055***	-0.047***	-0.074***
		(0.010)	(0.010)	(0.011)	(0.011)		(0.011)	(0.011)	(0.011)	(0.011)
Age x Post			-0.029**	-0.031**	-0.019			-0.017	-0.018	-0.008
			(0.014)	(0.014)	(0.013)			(0.014)	(0.014)	(0.013)
Age 1st x Post			-0.012	-0.013	-0.007			-0.021**	-0.022**	-0.016
			(0.010)	(0.010)	(0.010)			(0.010)	(0.010)	(0.010)
4th grade SES x Post				-0.030***	-0.062***				-0.015	-0.044***
				(0.010)	(0.010)				(0.010)	(0.010)
4th grade Score x Post					0.296***					0.269***
					(0.010)					(0.010)
Observations	384,813	384,813	384,813	384,813	384,813	380,309	380,309	380,309	380,309	380,309

Notes: This table shows results from the *difference-in-difference* estimates using equation (3) and extends those in Table 2. *Panel A* shows results for GPA in mathematics and reading standardized at the school-grade-year level for sixth grade students in 2019 and 2020 controlling for second grade baseline characteristics. *Panel B* shows results for sixth grade students in 2018 and 2020 controlling for second and fourth grade baseline characteristics. *Panel C* shows results for fifth grade students in 2019 and 2020 controlling for fourth grade baseline characteristics. Column 1 and 6 show the estimates without any controls for heterogeneous effects, analogous to what is done in equation (1). Column 2 and 7 add a dummy for the mother having higher education. Column 3 and 8 add dummies for both mother's age and age at first birth being above 30. Column 4 and 9 add a dummy for baseline socioeconomic status being above the mean. Column 5 and 10 add a dummy for baseline standardized test scores being above the mean. Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.

Table A.2—: DID estimates of having a sibling on mathematics GPA by baseline achievement and expectations

	TWFE Effect on Mathematics GPA		
	1 sibling	2 siblings	3 siblings
	(1)	(2)	(3)
Panel A: Low achievement students (Q1)			
	-0.002	-0.041***	-0.103***
	(0.008)	(0.009)	(0.015)
Observations	238,723	201,254	175,478
Panel B: High achievement students (Q4)			
	-0.047***	-0.099***	-0.149***
	(0.008)	(0.012)	(0.023)
Observations	352,745	272,727	236,028
Panel C: Low Education Expectations: Below 5-year college			
	0.008	-0.022	-0.064**
	(0.014)	(0.017)	(0.026)
Observations	81,715	69,691	60,595
Panel D: High Education Expectations: 5-year college and beyond			
	-0.031***	-0.065***	-0.097***
	(0.004)	(0.006)	(0.011)
Observations	980,339	780,019	675,846

VIII. Appendix B: Other Covid Shocks

If the additional sibling is causing learning losses, is this due to school closures or other related shocks? I provide some evidence that suggests that other shocks, although important for overall learning losses, may be unrelated to how having a sibling affected them.

A. School Closures

A small set of schools reopened operations either fully or partially in 2021. I estimate the event study for that group of schools and show the results in panel A of Figure B.1 which shows attenuated effects on 2021, consistent with classes returning to in person.

B. Income Shocks

First, in panel B of Figure B.1, I show, using data from household surveys, that average household income per capita declined sharply during the first half of 2020 but had mostly recovered by 2021. The much larger shock in the first year of closures is inconsistent with the roughly equal-sized effects observed in Figure 2.

Second, because I have access to national data, I estimate the effect of having a sibling separately for each region and compare it with regional variation in GDP growth. In panel C of Figure B.1, I show that the two are essentially uncorrelated: regions that experienced greater GDP losses do not exhibit larger effects of having a sibling.

Finally, while there is no information on income, there is a socioeconomic index based on household characteristics. In Table B.1, I show results for 2022 and 2023, which are generally not significant, although in a few cases they point in the direction opposite to negative income shocks as a mechanism. That is, the socioeconomic status of larger families is either the same of has slightly improved relative to that of only children. One caveat is that this index is more rigid than income, and families could experience income shocks without an immediate effect on the socioeconomic index, which is based on materials in the household, access to services, assets owned, and parent's education level.

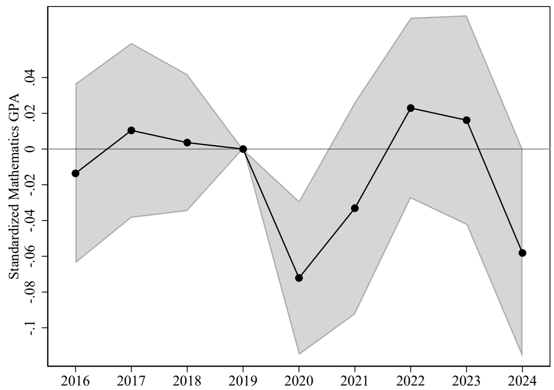
C. Health Shocks

Similar to the analysis of income shocks, I estimate the effect of having a sibling separately for each region and compare it with regional variation in excess deaths from COVID-19. In panel D of Figure B.1, I show that the severity of the pandemic, measured by death rates, was uncorrelated with the magnitude of the estimated sibling effect.

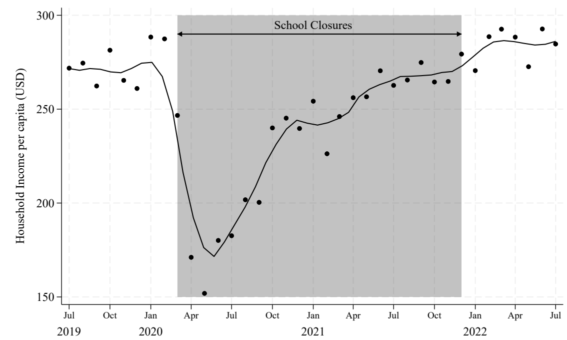
Table B.1—: DID estimates of having a sibling on socioeconomic index

	TWFE			
	Has a sibling	1 sibling	2 siblings	3 siblings
	(1)	(2)	(3)	(4)
Panel A: 2nd grade students (2019, 2022)				
Socio Economic Index	0.013	0.007	0.005	0.006
	(0.009)	(0.009)	(0.019)	(0.069)
Observations	96,352	86,837	57,119	49,746
Panel B: 4th grade students (2019, 2022)				
Socio Economic Index	-0.011	-0.011	-0.030*	-0.038
	(0.009)	(0.009)	(0.016)	(0.039)
Observations	86,858	74,479	50,776	42,382
Panel C: 4th grade students (2018,2024)				
Socio Economic Index	0.013***	0.008*	-0.011	-0.034
	(0.004)	(0.004)	(0.009)	(0.032)
Observations	397,895	352,349	260,225	230,647

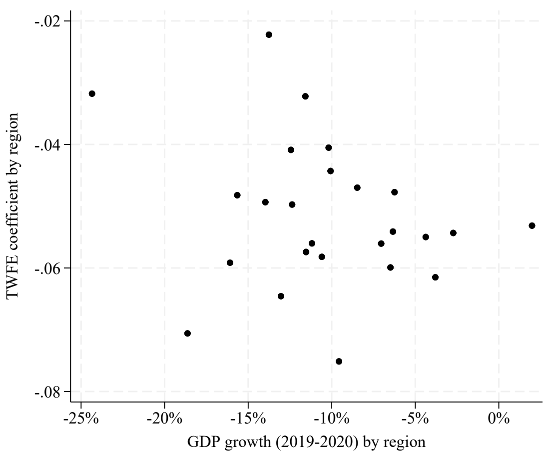
Notes: This table shows results from the *difference-in-difference* estimates of having a sibling on socioeconomic status using equation (3) for second and fourth grade students controlling for the mother having higher education for both mother's age and age at first birth being above 30. The first column shows the estimate of having a sibling when compared to those without a sibling. The second to fourth column show the estimate of having one, two a three siblings respectively when compared to those without a sibling. Each panel shows results for the effect on the socioeconomic status index based on materials in the household, access to services, assets owned, and parent's education level. . *Panel A* shows results for representative sample of second grade students who take the exam in 2019 (pre) and 2022 (post). *panel B* shows results for representative sample of fourth grade students who take the exam in 2019 (pre) and 2022 (post). *panel C* shows results for the universe of fourth grade students who take the exam in 2018 (pre) and 2024 (post). Estimated standard errors, reported in parentheses, are clustered at the primary school level. ***p<0.01, **p<0.05, *p<0.1.



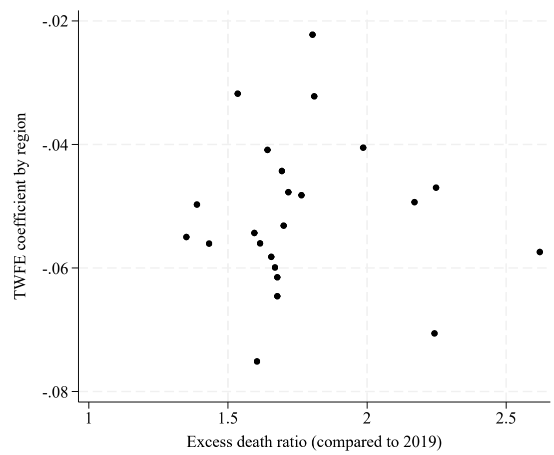
(a) Effects of having a sibling for schools that reopened in 2021



(b) Monthly average household income per capita



(c) Effect of having a sibling and GDP growth



(d) Effect of having a sibling and COVID excess death ratio

Figure B.1. : Other shocks

Notes: This figure shows evidence that school closures and not other covid related shocks are what is driving the effects. In the *top left panel (a)*, I estimate the event study from equation (2) for schools that reopened early in 2021. As expected, the effect on learning losses of having a sibling is attenuated for those schools in 2021. In the *top right panel (b)* I show the trend in household income per capita estimated using household surveys from 2020-2022 and show that income losses are concentrated in the first part of 2020, and yet, learning losses are equally present in 2020-2021. In the *bottom left panel (c)* I plot the estimate of the effect of having a sibling from equation (1) for each region (Y-axis) against the GDP growth from 2019 to 2020 (X-axis) and show that there is no relationship. In the *bottom right panel (d)* I plot the estimate of the effect of having a sibling from equation (1) for each region (Y-axis) against the excess death rate in 2020 compared to 2019 (X-axis) and show that there is no relationship.