# Sibling Spillover in Rural China: A Story of Sisters *

Cynthia Bansak† Xuan Jiang‡ Guanyi Yang§

April 17, 2021

**Abstract**

We study the sibling spillover effects on school performance of the elder sibling from the younger sibling using data on multi-children households in rural China. We use the variation in the younger sibling's schooling status to parse out the direction of the impact of spillover effects and the arbitrary school enrollment cutoff dates imposed by the Chinese Law of Compulsory Education as exogenous variation in the timing of school eligibility. We find a significant increase in school test scores of elder siblings when their younger siblings begin school. The strongest spillover effects occur when both children are daughters. Such increases in test scores come from a more intense academic atmosphere within a household when both children enroll in school and are not attributed to parental education investments or attitudes. Our findings suggest that policies promoting girls' education, pre-school education programs, and after school public resources can have multiplier effects through sibling spillovers.

*JEL classification:* E24, C68, J30.

*Keywords:* Human capital, peer effect, sibling spillover, rural China, sisters, girls, school enrollment, intra-household allocation, intra-houshold externality.

*We thank Shan Li, Zhiqiang Liu, Daniel Millimet, Juan Pedroza, Yang Song, and Xianfang Xiong for helpful comments on earlier drafts of this paper as well as participants at the 2020 CHLR International Symposium on Human Capital and Labor Markets, the 2020 SEA conference, and the 2019 APPAM meeting.

†Bansak: cbansak@stlawu.edu. Department of Economics, St. Lawrence University

‡Jiang: jiang.445@osu.edu. Department of Economics, The Ohio State University

§Yang: gyang@ColoradoCollege.edu. Department of Economics & Business, Colorado College

# 1 Introduction

Investment in children's human capital has lifetime impacts. Childhood interventions can reduce inequalities and promote intergenerational mobility (e.g. Keane and Wolpin, 1997; Heckman, 2008; Black, Devereux, et al., 2011; Huggett, Ventura, and Yaron, 2011; Yang, 2018). In spite of its importance, most of our understanding of children's human capital production is limited to observations at the household level and from parental resource allocations to their children (e.g. Yang and Bansak, 2020; Del Boca, Flinn, and Wiswall, 2013; Guryan, Hurst, and Kearney, 2008; Ramey and Ramey, 2010; Blau and Currie, 2006; Blau, 1999). This paper extends the analysis to *within household* dynamics and investigates an externality between siblings on school performance in multi-children households in rural China.

We find that an elder child's school performance improves significantly when their younger sibling begins elementary school. The increase in school performance does not rely on changes of family resources allocated to the elder sibling, but merely results from a positive externality from the younger sibling's studying to the elder sibling's academic achievements. We refer to this externality as a sibling spillover effect. This positive spillover effect differs significantly across sibling gender-pairs, and is particularly driven by the younger sisters' school enrollment. To the best of our knowledge, we are the first to examine a spillover effect in school performance from the younger sibling to an elder sibling.

The biggest challenge in identifying the sibling spillover is to disentangle it from factors outside of the sibship (or common family factors) and to identify the direction of impact between siblings.[1] We use the variation in the younger sibling's schooling status (started school or not) to parse out the impact of spillover effects *from the younger sibling to the elder one*. We exploit the arbitrary birth date cutoff imposed by the Chinese Law of Compulsory Education as an exogenous variation in schooling status. The law stipulates that children

[1] As Black, Breining, Figlio, Guryan, Karbownik, Nielsen, Roth, and Simonsen (2017) review, siblings growing up in the same household are the ultimate peers, sharing the same parents, similar genetics and experiences.

must be at least 6 years old or above on August 31 to be admitted into primary school in the new school year. This arbitrary day of birth generates exogenous variations in the younger sibling's eligibility of school enrollment, allowing us to attribute changes in the elder sibling's school performance to that of the enrollment of younger sibling.

We use data from the Rural Household Survey (RHS) of the Longitudinal Survey on Rural Urban Migration in China (RUMiC) to conduct our analysis. Having an in-school younger sibling leads to a 6.12 percentage point increase in math scores for a student relative to a same-grade peer who has a pre-school (not in school) sibling. Such an effect remains after controlling for differing gender ratios at the county-level, an approximation of the local son-preference culture, and for the annual education expenditures on the student. It is also robust to changes in the younger sibling's age range.

When further exploring the gender heterogeneity in sibling spillovers, we find that the second child being a sister creates an even stronger effect on the spillover to the elder child's test scores. Having a younger sister in school leads to a 13.8 percentage point increase on the elder sister's math score and a 12 percentage point increase on their Chinese score. The elder brother also sees a large increase in his estimated test scores, although this impact is not statistically significant possibly due to small sample size.

We conjecture that the positive spillover is a result of a positive at-home externality that occurs when both siblings are studying. When the younger child begins school, the elder child may experience an increase in sibling studying time relative to sibling play time, a heightened household attention to school performance, or more frequent household discussions on school matters. Elder siblings more exposed to this heightened academic environment at home may see this emphasis on schoolwork translate to a better performance in school. Though difficult to directly measure, we consider siblings with more time at home and a stronger home connection to experience a stronger spillover externality. We approximate the at-home time by investigating the different spillover effect between elder children attending boarding school and elder children attending day school. A child attending boarding school has the least interaction with the siblings at home. As a result, we find that the positive sibling spillover nearly disappears for an elder sibling in boarding school. Additionally, the positive spillover effect also diminishes for elder siblings attending a school that is further away from home compared to those who live closer. To further justify this potential mechanism of the positive spillover effects, we test for changes in parental resource allocation and supervision attitude on the elder child when the younger sibling starts school and find no evidence that the increase in the elder child's school performance is due to increased parental resource allocation or attention.

Overall, this paper joins others in providing a new perspective on sibling spillovers. Frameworks providing insight on human capital production among siblings within a family typically take one of three approaches: they assess the quantity-quality trade-off needed with multiple children (e.g. Becker, 1960; Becker and Lewis, 1973; Becker and Tomes, 1976; Blake, 1981; Hanushek, 1992; Liang and Gibson, 2018), they allow for birth order effects (e.g. Black, Grönqvist, and Öckert, 2018; Breining, Doyle, Figlio, Karbownik, and Roth, 2020; Lin, Pantano, and Sun, 2020), and most recently, they look for possible sibling interactions (e.g. Joensen and Nielsen, 2018; Black et al., 2017; Breining, 2014; Yi, Heckman, Zhang, and Conti, 2015). To date, the literature on sibling spillovers has largely focused on developed countries and the impact of older children on younger children. We contribute to the literature by studying the spillover effect of younger siblings on the school performance of elder siblings.

We also contribute to the sibling spillover literature regarding developing economies, where educational resources are scarce and gender inequality in educational resource allocation tends to be larger. Only one paper we know of has looked at sibling spillovers in China and this was from the older to the younger child (Biavaschi, Giulietti, and Zimmermann, 2015). The findings of our paper reveal a strong sibling spillover in school work, driven by sisters, consistent with the result from Jakiela, Ozier, Fernald, and Knauer (2020) who studied siblings in Kenya. However, Jakiela et al. (2020) were examining the impact from an older sibling to a younger one, not the other way around. Our results suggest that policy geared towards daughters and sisters in developing communities like rural China should produce the largest spillovers and ultimately largest increases in human capital development over time.

# 2 Data

We use data from the Rural Household Survey (RHS) of the Longitudinal Survey on Rural Urban Migration in China (RUMiC) to conduct our analysis.[2] The 2008 and 2009 waves of RHS in RUMiC are the only publicly available waves from RUMiC and we are using them to identify geographic information, individual demographic characteristics, family relationships, parental education and income, and schooling information for all children under 16 in surveyed households. The 2009 wave contains exam scores in Chinese language and math for all in-school children, providing us a valuable and reliable measures of school performance.[3]

Another reason that RUMiC suits our research goals is that it covers a relatively large number of multi-children households. Over a quarter (27 percent) of the children in RHS of RUMiC live in two-children household, second to the single-child families (69 percent), leaving only a small share of children in households with three or more children. We focus our analysis on two-children households and check robustness of our main estimates on three-children households.[4]

Our main sample used for the analysis in this paper includes 671 individuals (elder children) from two-children households. Table 1 provides a description of summary statistics

[2]RUMiC is a collaboration project between the Australian National University, Beijing Normal University and the Institute for the Study of Labor (IZA). RUMiC consists of three independent surveys: the Urban Household Survey (UHS), the Rural Household Survey (RHS) and the Migrant Household Survey (MHS). Each of the three surveys include comprehensive information on household and personal characteristics, detailed health status, employment, income, training and education of adults and children, social networks, family and social relationships, life events, and mental health measures of the individuals.

[3]Although self-reported, test scores are likely accurate due to the fact that most Chinese schools give each student a booklet that has their scored recorded for each semester during the school year (Biavaschi et al., 2015; Chen, Huang, Rozelle, Shi, and Zhang, 2014).

[4]Removing the households with more than two children also reduces the concern for the heterogeneity of family size to sibling peer effect (Chen, Chou, Wang, and Zhao, 2019). Our results remain when we relax the sample to include three-children households, accounting for 3 percent of all families in our sample.

Table 1: Summary Statistics

| | **Younger Sibling Not in School** | | **Younger Sibling in School** | | |
|---|---|---|---|---|---|
| | **Obs. 263** | | **Obs. 408** | | |
| Variable | Mean | SD | Mean | SD | Difference |
| Age | 11.278 | 3.022 | 15.007 | 2.262 | -3.729*** |
| Female | 0.658 | 0.475 | 0.669 | 0.471 | -0.011 |
| Birthweight | 3217.795 | 427.043 | 3230.686 | 447.014 | -12.892 |
| Chinese score | 81.078 | 12.378 | 78.361 | 12.436 | 2.717*** |
| Math score | 83.069 | 11.115 | 80.003 | 13.219 | 3.066*** |
| Distance to school (km) | 4.521 | 7.648 | 10.601 | 37.584 | -6.079** |
| Boarding school | 0.319 | 0.467 | 0.519 | 0.5 | -0.200*** |
| Parents' Worry | 0.681 | 0.467 | 0.686 | 0.465 | -0.006 |
| Annual edu expenses (yuan) | 1093.966 | 2197.67 | 2060.673 | 3002.964 | -966.707*** |
| Younger sib's age | 3.973 | 1.726 | 11.012 | 2.903 | -7.039*** |
| Younger sib is female | 0.388 | 0.488 | 0.341 | 0.475 | 0.047 |
| Younger sib' birthweight | 3307.11 | 467.977 | 3284.828 | 479.934 | 22.282 |
| Father's wage | 1534.422 | 3213.169 | 1306.9 | 1993.444 | 227.523 |
| Mother's wage | 549.156 | 1959.935 | 384.064 | 643.573 | 165.092 |
| Father Never Been to School | 0.011 | 0.007 | 0.007 | 0.004 | 0.004 |
| Father Literacy Class | 0.011 | 0.006 | 0.007 | 0.004 | 0.004 |
| Father Elementary School | 0.122 | 0.020 | 0.174 | 0.019 | -0.052* |
| Father Junior Middle School | 0.677 | 0.0289 | 0.662 | 0.023 | 0.015 |
| Father Senior Middle School | 0.122 | 0.020 | 0.120 | 0.016 | 0.002 |
| Father Specialized Secondary School | 0.023 | 0.009 | 0.017 | 0.006 | 0.006 |
| Father Polytechnic college | 0.034 | 0.011 | 0.012 | 0.005 | 0.021* |
| Father Undergraduate | 0 | 0 | 0 | 0 | 0 |
| Mother Never Been to School | 0.030 | 0.011 | 0.029 | 0.008 | 0.001 |
| Mother Literacy Class | 0.011 | 0.007 | 0.017 | 0.006 | -0.006 |
| Mother Elementary School | 0.221 | 0.026 | 0.301 | 0.023 | -0.081** |
| Mother Junior Middle School | 0.631 | 0.029 | 0.603 | 0.024 | 0.028 |
| Mother Senior Middle School | 0.065 | 0.015 | 0.047 | 0.010 | 0.018 |
| Mother Specialized Secondary School | 0.027 | 0.010 | 0.002 | 0.002 | 0.024*** |
| Mother Polytechnic college | 0.011 | 0.007 | 0 | 0 | 0.011** |
| Mother Undergraduate | 0.004 | 0.004 | 0 | 0 | 0.004 |

Notes: This table presents the summary statistics of the main variables that we used for analysis, separately by the schooling status of the younger siblings. Specifically, in the upper panel, we show the demographic characteristics, school characteristics, test scores, and parents' educational expense and worry of the elder sibling—the individuals of interest in our analysis. In the bottom panel, we show demographic information of the younger sibling, parent's wage and parents' educational attainments. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

split by whether a younger child is in school or not. The top panel of Table 1 describes the average characteristics of the elder child, the subject of interest in our paper. The bottom panel presents the younger child's characteristics and each parent's wage and educational attainment. The key dependent variables for the study are the elder sibling's Chinese score and Math score. We standardize the test scores to percentage terms to remove differences in scaling. Even so, test scores might differ by school grade, school or school district. The RUMiC dataset provides no geographic identification disaggregated below city-level, so we do not observe school or school district information. Instead, we control for children's grade fixed effects, geographic location fixed effects, and family characteristics in all of our analysis to address this concern. The average scores for both Chinese and math are around 80 percent. The average scores are lower for elder siblings from families with a younger child in school. One reason might be that test scores tend to go down as one moves to a higher grade. Families with both children in school are spending more money on educational expenditures, and are more likely to have the elder sibling in boarding school and in a school further away from home.

We find no statistical difference in mother's and father's wage income between families with a younger child in school and those with a younger child not enrolled in school. There is no gender and birth-weight differences among elder and younger siblings between the two groups. We also do not find statistical difference in parental attitude on the elder child, measured as parental worry, between the two groups. There is no significant difference in mother's and father's junior middle school completion rate—the most common educational attainment in our data—between the two groups. This suggests that the children and households with younger sibling in and not in schools are quite similar.

# 3 Empirical strategy

We examine the impact of a younger sibling's schooling status on the educational outcomes of their older sibling, and refer to this as the sibship spillover effect. We regress the elder sibling's schooling performance on an indicator of whether the younger sibling is in school and control for a host of demographic and economic characteristics of the parents and both siblings:

$$S_i^E = \alpha + \beta inSchool_i^Y + \Lambda X_i^E + \Gamma X_i^Y + \Pi X_i + \varepsilon_i \tag{1}$$

where $E$ stands for the elder child, and $Y$ stands for the younger child. Thus, $S_i^E$ is the elder sibling's schooling performance, Chinese score or Math score, for person i. The variable $inSchool_i^Y$ is an indicator whether the younger sibling is in school. The coefficient on $inSchool_i^Y$, $\beta$, is our main coefficient of interest.

We include a number of controls for both siblings and household characteristics to absorb idiosyncratic variation. $X_i^E$ is a vector of the elder sibling's characteristics, including the elder sibling's birth weight, gender, and (schooling) grade fixed effects. $X_i^Y$ is a vector of the younger sibling's characteristics, including birth weight, gender, and a linear trend of age. We also control for ethnicity and province fixed effects for the household, both parents' ages, wages, educational attainment, and educational expenses on the elder sibling in $X_i$.[5] We cluster all standard errors at the city level.

## 3.1 Identification

A main concern to the estimate, $\beta$, in Equation 3 is that the timing of the younger sibling's school enrollment may not be exogenous. For instance, parents may delay their younger child's schooling due to financial hardship thus this may correlate to the school performance of the elder child. Parents could also experiment on the younger child's school entrance

[5] We impute zeroes for missing wages of parents and include an indicator for a missing wage in our regressions.

based on their observation of the older child's performance, considering an age-school relationship. To account for such potential endogeneity on the spillover between the younger child's schooling and the older child's test scores, we take advantage of an arbitrary cutoff birth date for compulsory school entry age, stipulated by the Chinese Compulsory Education Law, and implement an instrumental variable indicating the younger child's eligibility to enroll in school based on the policy.

### 3.1.1 China's compulsory education and education in rural China

China passed its Compulsory Education Law in 1986 that aims to establish universal education for all school-aged children. The universal education includes nine years of schooling spanning from elementary school to junior high school.[6] Depriving their school-aged children from enrollment by parents is considered a crime. Across China, schools have September 1 as the official first day of a school year. The implementation of the Compulsory Education requires that children with their sixth birthday before August 31 must enroll in an elementary school on September 1 of the same year. Children who reach six years old after August 31 must wait until the following school year for enrollment.

Importantly, the birth date of August 31 becomes an arbitrary cutoff for schooling among six-year-olds. Having a birth date before or after August 31 creates an exogenous variation in school start dates. Fang, Eggleston, Rizzo, Rozelle, and Zeckhauser (2012) document that there were variations in the enforcement of the school starting cutoff date in the early implementation of the Law. However, Zhang and Xie (2018) find that the actual birth date eligible for school entry is now largely in compliance with the August 31 cutoff. Huang,

[6] The elementary school enrolls students from age six (grade one) and ends at age twelve (grade six), with a curriculum focusing on Chinese, math and physical education. Urban elementary schools also include music, drawing, science, and morality and ethics. Some advanced urban schools introduce English courses. Junior high schools cover a more comprehensive curriculum, including Chinese, Mathematics, English, Biology, Politics, Geography, History, Physical Education as the main subjects that students have to take mid-term and final exams in each semester. Although we do not have test scores for subjects other than Chinese and math for the junior high students, one could argue that Chinese and math are two of the most important subjects in the sense of how much weight is given to these tests, even in the senior high school exams and college entrance exams.

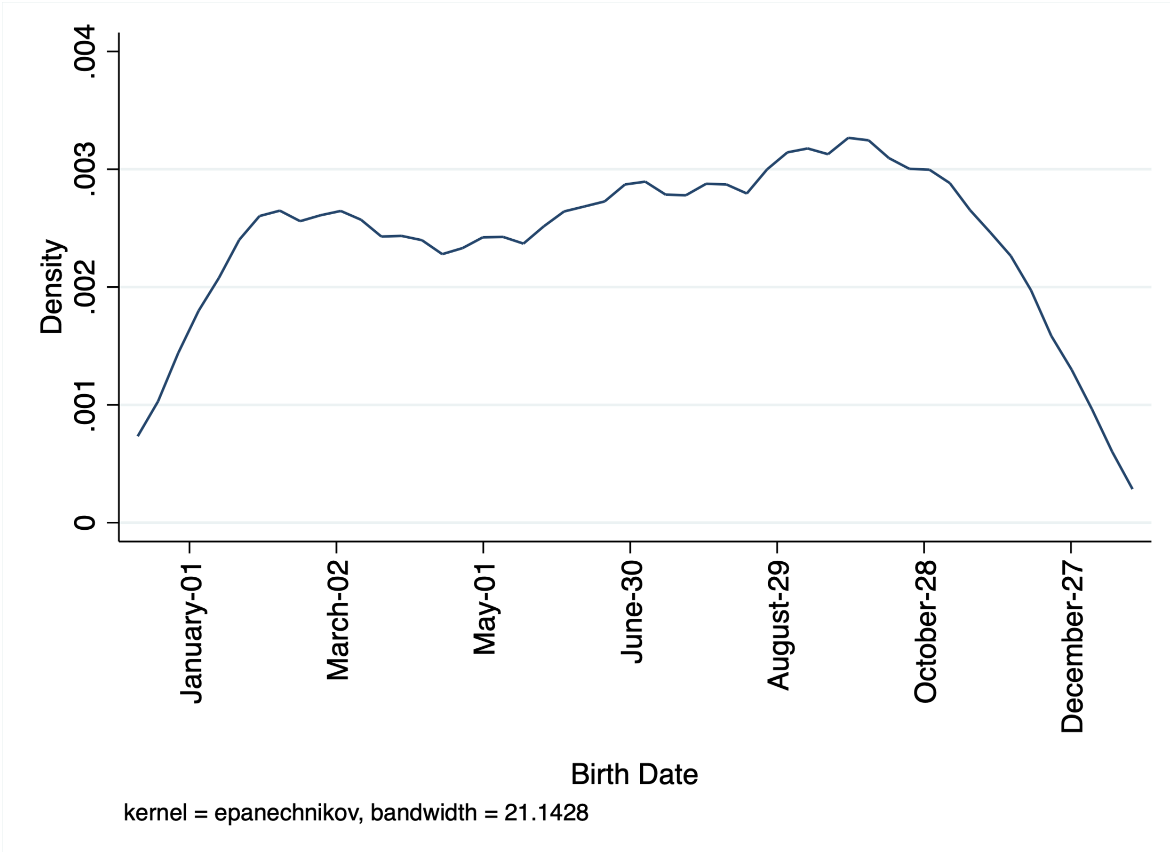


Figure 1: Density of Children's Birth Date

Notes: This figure plots the kernel density of birth dates (month and date) of all younger children in our sample. The purpose of this figure is to show there is no breaks in the density around August and September.

Zhang, and Zhao (2020) show evidence of birth planning around the school entry date in certain areas. In Figure 1, we find that there is no obvious evidence of birth planning around August 31 in our sample. Along these lines, it is also reasonable to assume that the exact birth date of one child is not correlated with the school performance of another child.

### 3.1.2 Instrumental variable

Incorporating the arbitrary birth date cutoff, we employ a two-stage estimation using the arbitrary birth date cutoff as an instrumental variable for the younger child's school eligibility. Our first stage estimation is specified as:

$$inSchool_i^Y = \alpha + \theta Ineligible_i^Y + \Lambda X_i^E + \Gamma X_i^Y + \Pi X_i + \varepsilon_i \tag{2}$$

where $Ineligible_i^Y$ is an indicator for the younger sibling's lack of eligibility of enrolling in school. $Ineligible_i^Y = 1$ identifies the younger children who had their sixth birthday after August 31 in the survey year; $Ineligible_i^Y = 0$ otherwise.[7] Thus, $\theta$ provides the propensity of a child to be in school. We would expect $\theta$ to be negative as a six year old should not be in school if one was born in the last four months of the year. Our estimate of interest then will be $\beta^{IV}$ from the second stage:

$$S_i^E = \alpha + \beta^{IV} \widehat{inSchool}_i^Y + \Lambda X_i^E + \Gamma X_i^Y + \Pi X_i + \varepsilon_i \tag{3}$$

# 4 Results

In this section, we present our main results, the effects of sibship spillover and the heterogeneity across sibling gender composition. We further explore a potential mechanism associated with the spillover.

## 4.1 Sibship spillover

We estimate the impact of younger sibling's in-school status on the elder sibling's schooling performance, measured by Chinese and math test scores. We present the results in Table 2. The top panel of Table 2 reports the IV estimation results and the bottom panel presents the OLS estimates for comparison purposes. The odd-numbered columns have the standardized Chinese test score as the outcome variable and the even-numbered columns have the standardized math test score as the outcome variable. We control for the elder sibling's grade fixed effects, both sibling's birth weight (a commonly used proxy for health), a linear trend of the younger sibling's age, family characteristics including ethnicity and province fixed effects, both parent's age, income, and educational attainment in all columns. Our instrument is strong, with a Kleibergen-Paap rk Wald F-statistic of 140 for our baseline

[7] We use birth date (year, month, date) to construct the variable $Ineligible_i^Y$.

Table 2: Sibling Spillover Effects

| | (1) Chinese | (2) Math | (3) Chinese | (4) Math | (5) Chinese | (6) Math |
|---|---|---|---|---|---|---|
| ***IV Estimates:*** | | | | | | |
| Younger Sib in School | 3.592 | 5.857** | 3.610 | 5.856** | 3.799 | 6.115** |
| | (2.333) | (2.757) | (2.390) | (2.728) | (2.786) | (3.064) |
| Elder Sib is Female | | | 3.663*** | 2.873*** | 2.698*** | 1.812* |
| | | | (0.842) | (0.921) | (0.972) | (0.978) |
| Younger Sib is Female | | | 0.457 | 0.214 | 0.653 | -0.227 |
| | | | (1.013) | (1.121) | (1.071) | (1.216) |
| County Children Female Ratio | | | -23.61 | -11.49 | -16.46 | -2.765 |
| | | | (17.39) | (19.61) | (15.50) | (19.20) |
| ln(Edu Expense) | | | | | -0.0965 | -0.0907 |
| | | | | | (0.466) | (0.546) |
| First Stage F-Stats | 139.295 | 139.295 | 140.701 | 140.701 | 95.0327 | 95.0327 |
| Observations | 671 | 671 | 671 | 671 | 574 | 574 |
| ***OLS Estimates:*** | | | | | | |
| Younger Sib in School | 0.876 | 1.327 | 0.726 | 1.215 | 1.114 | 0.960 |
| | (1.556) | (1.640) | (1.577) | (1.654) | (1.868) | (2.000) |

Notes: This table presents the IV and OLS estimates of the main analysis. The upper panel shows the IV estimates of three different specifications. The lower panel shows the corresponding OLS estimates. Each column is a separate regression. The dependent variable in odd-numbered columns is Chinese test score and the dependent variable in even-numbered columns is Math test score. The test scores are standardized to 100-point scale. The key variable of interest is a dummy variable of younger sibling enrolled in school. Columns (1) and (2) are the results of the baseline model, including both sibling's birth weight, a linear trend of younger sibling's age, elder sibling's school grade fixed effects, both parent's age, income, and education, the household's ethnicity, province fixed effects. Columns (3) and (4) are adding both sibling's gender and county level children girl ratio. Columns (5) and (6) further control for the log annual educational expense on the elder sibling. All standard errors are clustered at city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

model. We present our first stage estimates in Table A.2. The estimates in columns (1) and (2) of Table 2 show that having the younger sibling enrolled in school improves the elder child's test scores in math by 5.86 percentage points, and it is statistically significant at the 5% level. The impact on the Chinese score has a relatively smaller magnitude and is not statistically significant.

Columns (3) and (4) add both sibling's gender and a county-level girl ratio to capture potential gender-specific educational resource allocation based on a son-preference culture rooted in rural China[8] We approximate the son-preference culture by exploring the shares

[8] A substantial literature documents the son-preference culture in Asian communities and its impact on

of girls in each county. The county-level girl ratio is calculated by using the total number of girls under age 18 divided by the total number of children under age 18. The estimates of interest are highly consistent with the ones in columns (1) and (2). We further control for the log of annual educational expense on the elder sibling in Columns (5) and (6), capturing the possibility for parents to switch resources away from or towards the elder sibling when the younger one begins school. This is our preferred specification and we find that having an in-school younger sibling will have a non-significant positive effect on Chinese test scores and a 6.12 percentage point significant increase in math scores for an older sibling, relative to a same-grade peer who has a pre-school sibling.

The IV estimates show that the school enrollment of the second child has a statistically strong, and economically meaningful positive effect on the elder child's mathematics performance, but less so in the Chinese performance. This compares to the prevalent null effects in the OLS estimation, suggesting a downward bias in OLS estimates. The OLS estimates of the control variables are similar in both economic and statistical significance compared the IV estimates and are presented in Table A.1.

## 4.2 Robustness of results

Our instrument considers all children's sixth birth date after August 31, 2009 as ineligible for school enrollment and all with prior birth dates as eligible. One concern to our IV estimation is that the effect is driven by those with birth date further away from the cutoff date. In this sub-section, we narrow the age window of the younger siblings to check for the robustness of our results and present restricted sample results from our preferred specification in Table 3. We would ideally only include a narrow bandwidth of the younger sibling's birth date in our estimation for a local effect. Unfortunately, this is not possible as we suffer from a small sample size of our analytical dataset.

resource allocation to children of opposite genders (e.g. Blau, Kahn, Brummund, Cook, and Larson-Koester, 2020; Kaul, 2018; Hong Chew, Yi, Zhang, and Zhong, 2017; Das Gupta, Zhenghua, Bohua, Zhenming, Chung, and Hwa-Ok, 2003).

Table 3: Robustness: Sibling Spillover Effects

| | (1) Chinese Ages 4-9 | (2) Math Ages 4-9 | (3) Chinese Ages 4-8 | (4) Math Ages 4-8 | (5) Chinese Ages 5-9 | (6) Math Ages 5-9 | (7) Chinese Ages 5-8 | (8) Math Ages 5-8 |
|---|---|---|---|---|---|---|---|---|
| Younger Sib in School | 3.926** | 5.450*** | 2.760 | 5.602** | 3.704* | 5.914*** | 1.626 | 5.965** |
| | (1.833) | (2.071) | (2.121) | (2.403) | (2.000) | (2.046) | (2.579) | (2.728) |
| Elder Sib is Female | 3.084** | 1.753 | 2.041 | 1.250 | 3.546** | 1.282 | 2.278 | 0.360 |
| | (1.421) | (1.454) | (1.553) | (1.574) | (1.685) | (1.606) | (1.871) | (1.777) |
| Younger Sib is Female | 2.013 | 2.293* | 2.338* | 2.307* | 3.989*** | 4.593*** | 4.518*** | 4.895*** |
| | (1.314) | (1.339) | (1.378) | (1.361) | (1.291) | (1.532) | (1.314) | (1.627) |
| County Children Female Ratio | -12.39 | -2.920 | 6.947 | 16.13 | -28.60 | -29.47 | -10.11 | -10.49 |
| | (18.43) | (18.30) | (21.84) | (19.98) | (19.89) | (18.22) | (25.21) | (20.44) |
| ln(Edu Expense) | 0.0328 | 0.0952 | -0.541 | -0.473 | 0.265 | 0.706 | -0.396 | 0.146 |
| | (0.423) | (0.507) | (0.396) | (0.537) | (0.406) | (0.507) | (0.472) | (0.573) |
| Observations | 243 | 243 | 209 | 209 | 195 | 195 | 161 | 161 |

Notes: This table presents the IV estimates from running the same specification in columns (5) and (6) in Table 2 on subsamples with different age ranges of the younger siblings. Columns (1) and (2) show estimates from the sample with younger sibling's age ranging 4-9. Columns (3) and (4) show estimates from the sample with younger sibling's age ranging 4-8. Columns (5) and (6) are estimates from the sample with younger sibling's age ranging 5-9. Columns (7) and (8) are estimates from the sample with younger sibling's age ranging 5-8. All standard errors are clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (1) and (2) restrict the younger siblings to ones with ages between 4 and 9 in year 2009; Columns (3) and (4) restrict the the younger sibling's ages to be within 4 and 8 in year 2009; Columns (5) and (6) restrict the younger sibling's age to fall between 5 and 9 years old in year 2009, and Columns (7) and (8) limit the age to be between 5 and 8 in year 2009. As the age band shrinks, the number of observations falls sharply. Notwithstanding this limitation, we find that the younger sibling's schooling status has a robust positive effect on the elder sibling's math score regardless of sample restrictions. The estimates on Chinese test score are also robust in terms of the magnitudes when restricting the sample. If anything, we gain significance for most of the estimates of Chinese test scores with this conservative specification.

Another potential concern with our identification strategy is that the spillover effect from the younger sibling to elder sibling might come from the age gap between siblings rather than a younger sibling's in-school effect. For example, the younger sibling may be older when entering school age, displaying a closer age gap to the elder sibling than the younger ones who are pre-school. Liu and Li (2016) and Zhang and Xie (2018) suggest that siblings of different age gaps may have varying school performance. To address this concern, we run regressions on a sub-sample of older children with only pre-school younger siblings, and with the variable of interest as the younger sibling's age. We report the results in Table 4. The first two columns include linear controls for age; the second two have age dummies; and the last two columns have grade level dummies for the older child. These combinations of age controls were used to estimate an age gap effect in an exhaustive manner. In no case did we find any statistically significant relationship between the younger sibling's age and the older sibling's test scores and our falsification test rules out this competing explanation.

Table 4: Falsification: Is it a Sibling's Age Gap Story?

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Chinese (100) | Math (100) | Chinese (100) | Math (100) | Chinese (100) | Math (100) |
| Younger Sib's Age | 0.480 | -0.0478 | 0.260 | 0.143 | -0.138 | 0.138 |
| | (0.632) | (0.353) | (0.530) | (0.341) | (0.412) | (0.358) |
| Observations | 263 | 263 | 263 | 263 | 263 | 263 |

Notes: This table presents the results of regressions of elder sibling's test scores on the age of pre-school younger sibling. The variable of interest is the younger sibling's age. Columns (1) and (2) are controlling for a linear term of the elder sibling's age; columns (3) and (4) are controlling for elder sibling's age dummies. Columns (5) and (6) control for the elder sibling's grade dummies. All standard errors are clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## 4.3 Story of sisters

Literature documents that girls consistently outperform boys in school (e.g. Duckworth and Seligman, 2006; Goldin, Katz, and Kuziemko, 2006; Buchmann, DiPrete, and McDaniel, 2008; DiPrete and Jennings, 2012; Lundberg, 2017; Jiang, 2020). Compared to boys, girls begin school more prepared, spend more hours on homework, and have fewer disruptive behaviors (e.g. DiPrete and Jennings, 2012; Becker, Hubbard, and Murphy, 2010; Goldin et al., 2006). We find comparable evidence in which the sibling spillover is driven by girls. We present these findings by exploring the heterogeneity of spillover across different sibship gender pairs in Table 5.

By sub-setting our sample to specific gender-pairs, the sample size decreases considerably.

Table 5: Heterogeneity of spillover by siblings gender pairs

| | **Elder F + Younger M** | | **Elder F + Younger F** | | **Elder M + Younger M** | | **Elder M + Younger F** | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Chinese | Math | Chinese | Math | Chinese | Math | Chinese | Math |
| Younger Sib in School | 0.306 | 1.324 | $12.01^{**}$ | $13.75^{**}$ | 0.0137 | 7.747 | 16.61 | 16.56 |
| | (4.006) | (3.931) | (5.620) | (6.347) | (8.831) | (11.78) | (15.40) | (13.63) |
| County Children Female Ratio | -28.82 | -9.057 | 14.23 | 2.624 | $-61.19^{**}$ | -39.28 | -53.94 | 1.545 |
| | (18.96) | (21.74) | (26.27) | (31.24) | (24.40) | (30.14) | (41.68) | (42.48) |
| Observations | 263 | 263 | 121 | 121 | 104 | 104 | 86 | 86 |

Notes: This table presents the IV estimates of the main regression by different subset of sample. Each column is a separate regression. The dependent variable is either Chinese test score or Math test score, in 100-point scale. The key variable of interest is a dummy variable of younger sibling enrolled in school. Control variables include both sibling's birth weight, younger sibling's age, elder sibling's grade fixed effects, both parent's age, income, and education, the household's ethnicity, province fixed effects, county level children girl ratio, and log total educational expenditure on the elder sibling. Column (1)-(2), (3)-(4), (5)-(6), and (7)-(8) are subsamples of different gender pairs, elder sister-younger brother, elder sister-younger sister, elder brother-younger brother, and elder brother-younger sister, respectively. All standard errors are clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Nevertheless, we find striking positive spillover effects for sibling pairs who are both girls (see columns (3) and (4)). Having a younger sister in school leads to a 12 percentage point increase in the elder sister's Chinese score and a 13.8 percentage point improvement in her math score. We also observe an economically strong result of the impact of a younger sister on an elder brother as seen in columns (7) and (8), although the standard errors are large due to the small sample. The results presented in the columns where younger brothers are attending schools (columns (1), (2), (5), and (6)) show null effects on the elder sibling's Chinese performance and statistically insignificant effects on the elder sibling's math performance.

Our findings in Table 5 contribute to the sibling gender literature by showing that having a younger sister in school provides a strong and positive impact on the elder sibling's school performance, compared to having a younger brother being in school.

## 4.4 Home connections and potential mechanisms

As we find a strong positive impact from the second child's enrollment status on the older child's school performance, we hypothesize that this improvement is due to a positive at-home externality from both siblings studying together. When the younger child is enrolled in school, the older child would likely be exposed to a more intense academic atmosphere at home. This could be reflected as an increase in studying time among siblings at home relative to sibling play time, a heightened household attention to school performance, or to more frequent household discussions of school matters.

Conditional on the variables provided by our data, we infer such mechanisms through the variations on the elder sibling's home-school distance and whether the older sibling attends boarding school. Children spend more time at home if they do not attend boarding school. A shorter home-school distance may also suggest a tighter social connection between a child and home. Following our hypothesis, we test whether children who spend more time at home and have stronger social connections at home are more likely positively influenced by the intense home academic atmosphere when the younger sibling begins school. This would be evident by observing a large sibling spillover when the older child's school is closer to home. To show the heterogeneous spillover effects on this margin, we add an indicator of whether the elder sibling is in a boarding school and an interaction between the boarding school indicator and an indicator whether the younger sibling is in school to our main specification. The results are shown in Panel A, Table 6. In Panel B, we replace the boarding school indicator in Panel A with an indicator of whether the elder sibling's school distance is above the median. In both panels, we estimate the main regression on both the full sample and the subsample with only female younger siblings.

The results in Table 6 support our hypotheses. In Panel A, elder children who are in boarding school experience significantly less sibling spillover effect from the younger sibling. The differences in the positive spillover effects between elder children who are in boarding school and those who are not range from 4.1 (full sample) to 13.3 (sub-sample) percentage points for the Chinese scores. Similarly, the reduction of the positive spillover effect for math score is as large as 14.2 percentage points for the elder children with a female younger sibling. The negative estimate of the interaction term nearly offsets the positive spillover

Table 6: Sibling spillover by elder child's boarding school status and home-school distance

| | All Siblings | | Younger Sibling is Female | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Chinese | Math | Chinese | Math |
| **Panel A: Boarding School** | | | | |
| Younger Sib in School | 5.674** | 7.163** | 18.96*** | 19.22*** |
| | (2.839) | (3.544) | (5.939) | (6.392) |
| Younger Sib in School×Elder Sib in Boarding School | -4.105* | -3.442 | -13.30*** | -14.17*** |
| | (2.408) | (2.491) | (4.159) | (4.307) |
| Elder Sib in Boarding School | 1.830 | 1.212 | 4.409* | 4.709** |
| | (1.995) | (1.952) | (2.599) | (2.357) |
| **Panel B: Above Median Distance** | | | | |
| Younger Sib in School | 5.266* | 6.668** | 12.33** | 14.16** |
| | (2.806) | (3.388) | (5.962) | (6.277) |
| Younger Sib in School×Elder Sib in School > Median Distance | -3.930* | -1.403 | -3.775 | -5.443 |
| | (2.272) | (2.248) | (4.370) | (3.917) |
| Elder Sib in School > Median Distance | 1.475 | 0.737 | 0.434 | 0.929 |
| | (1.690) | (1.480) | (3.243) | (2.744) |
| Observations | 574 | 574 | 207 | 207 |

Notes: The key independent variable in panel A is a dummy of the elder sibling goes to a boarding school. The key independent variable in panel B is a dummy of the elder sibling's school distance is above median. Each column is a separate regression. The first two columns show results from the full sample and the last two columns show results from a subsample with only female younger sibling households.

effect from younger sibling's enrollment. This suggests that elder children in boarding school lose the opportunity of their positive peer effects through a reduction of home time. Panel B shows consistent, although weaker, findings compared to Panel A: elder children who attend schools that are farther from home experience a 3.9 percentage points reduction in the positive sibling's spillover in Chinese scores, relative to their counterparts who attend schools closer to home. The reduction effects shown in Panel A is much stronger than the ones in Panel B because the child would have much less home time when attending a boarding school.

Studies attribute a significant portion of human capital development to parental attention and monetary investment on children (e.g. Guryan et al., 2008; Blau and Currie, 2006; Yum, 2016; Del Boca et al., 2013; Yang and Bansak, 2020). One might be concerned that alternative channels such as a direct adjustment of educational resources or attention from parents to the elder child when the younger child starts school might dominate the increased in performance, rather than the at-home externality we show above. We conduct additional tests in Table A.3 and do not find significant changes in education expenditure and parental attention to the elder child when the younger child begins school. Thus, we infer that the improvement in elder sibling's performance when the younger sibling starts school is not derived from changes in parental education spending or parental concerns or attention.

## 5 Conclusion

We find that having a younger sibling enrolled in school will significantly increase the elder sibling's test scores in both Chinese and math. We refer to this result as the positive sibling spillover effect from the younger sibling to the elder. The main driver of the positive spillover comes from the younger sibling being a female. Our mechanism suggests that the more the in-school siblings interact with each other at home, the better the elder sibling performs. This set of findings contributes to the literature of gender inequality in education and adds empirical evidence to the positive externalities of women's educational improvement. Thanks to these spillovers, policies that promote young women's education may well have important effects on other children as well, suggesting a potential multiplier effect of a policy.

Our results also advise that it could be beneficial to increase the time period in which spillover occur by offering pre-school children earlier education programs. In this case, siblings would be in school together longer. This could be done by introducing pre-K programs such as Head Start in the US. In such programs, children younger than the compulsory enrollment birth date would have access to school readiness programs at an younger age. Sibling spillover effects could be larger if additional after school public services were made freely available, such as library access. They could enhance at-home academic culture and further facilitate sibling spillovers. These offerings could be targeted at low income families to reduce future income inequality.

# References

Becker, G. S. (1960). *An Economic Analysis of Fertility, Demographic and economic change in developed countries*. Princeton University Press.

Becker, G. S., W. H. Hubbard, and K. M. Murphy (2010). Explaining the worldwide boom in higher education of women. *Journal of Human Capital 4*(3), 203–241.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *Journal of political Economy 81*(2, Part 2), S279–S288.

Becker, G. S. and N. Tomes (1976). Child endowments and the quantity and quality of children. *Journal of political Economy 84*(4, Part 2), S143–S162.

Biavaschi, C., C. Giulietti, and K. F. Zimmermann (2015). Sibling influence on the human capital of the left-behind. *Journal of Human Capital 9*(4), 403–438.

Black, S. E., S. Breining, D. N. Figlio, J. Guryan, K. Karbownik, H. S. Nielsen, J. Roth, and M. Simonsen (2017). Sibling spillovers. Technical report, National Bureau of Economic Research.

Black, S. E., P. J. Devereux, et al. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics 4*, 1487–1541.

Black, S. E., E. Grönqvist, and B. Öckert (2018). Born to lead? the effect of birth order on noncognitive abilities. *Review of Economics and Statistics 100*(2), 274–286.

Blake, J. (1981). Family size and the quality of children. *Demography 18*(4), 421–442.

Blau, D. and J. Currie (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education 2*, 1163–1278.

Blau, D. M. (1999). The effect of income on child development. *Review of Economics and Statistics 81*(2), 261–276.

Blau, F. D., L. M. Kahn, P. Brummund, J. Cook, and M. Larson-Koester (2020). Is there still son preference in the united states? *Journal of Population Economics*, 1–42.

Breining, S., J. Doyle, D. N. Figlio, K. Karbownik, and J. Roth (2020). Birth order and delinquency: Evidence from denmark and florida. *Journal of Labor Economics 38*(1), 95–142.

Breining, S. N. (2014). The presence of adhd: Spillovers between siblings. *Economics Letters 124*(3), 469–473.

Buchmann, C., T. A. DiPrete, and A. McDaniel (2008). Gender inequalities in education. *Annu. Rev. Sociol 34*, 319–337.

Chen, C., S.-Y. Chou, C. Wang, and W. Zhao (2019). The effect of the second child on the anthropometric outcomes and nutrition intake of the first child: Evidence from the relaxation of the one-child policy in rural china. *The BE Journal of Economic Analysis & Policy 20*(1).

Chen, X., Q. Huang, S. Rozelle, Y. Shi, and L. Zhang (2014). Effect of migration on childrenâs educational performance in rural china. In *Chinaâs economic development*, pp. 206–224. Springer.

Das Gupta, M., J. Zhenghua, L. Bohua, X. Zhenming, W. Chung, and B. Hwa-Ok (2003). Why is son preference so persistent in east and south asia? a cross-country study of china, india and the republic of korea. *The Journal of Development Studies 40*(2), 153–187.

Del Boca, D., C. Flinn, and M. Wiswall (2013). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

DiPrete, T. A. and J. L. Jennings (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research 41*(1), 1–15.

Duckworth, A. L. and M. E. Seligman (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of educational psychology 98*(1), 198.

Fang, H., K. N. Eggleston, J. A. Rizzo, S. Rozelle, and R. J. Zeckhauser (2012). The returns to education in china: Evidence from the 1986 compulsory education law. Technical report, National Bureau of Economic Research.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic perspectives 20*(4), 133–156.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *Journal of Economic perspectives 22*(3), 23–46.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of political economy 100*(1), 84–117.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic inquiry 46*(3), 289–324.

Hong Chew, S., J. Yi, J. Zhang, and S. Zhong (2017). Risk aversion and son preference: Experimental evidence from chinese twin parents. *Management Science 64*(8), 3896–3910.

Huang, C., S. Zhang, and Q. Zhao (2020). The early bird catches the worm? school entry cutoff and the timing of births. *Journal of Development Economics 143*, 102386.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54.

Jakiela, P., O. Ozier, L. Fernald, and H. Knauer (2020). Big sisters.

Jiang, X. (2020). Family planning and women's educational attainment: Evidence from the one-child policy. *Contemporary Economic Policy 38*(3), 530–545.

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

Kaul, T. (2018). Intra-household allocation of educational expenses: Gender discrimination and investing in the future. *World Development 104*, 336–343.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of political Economy 105*(3), 473–522.

Liang, Y. and J. Gibson (2018). Do siblings take your food away? using china's one-child policy to test for child quantity-quality trade-offs. *China Economic Review 48*, 14–26.

Lin, W., J. Pantano, and S. Sun (2020). Birth order and unwanted fertility. *Journal of Population Economics 33*(2), 413–440.

Liu, D. and X. Li (2016). Children born in july and august: a study on the age regulation in primary school and studentâs education access and development. *The Journal of Chinese Sociology 3*(1), 22.

Lundberg, S. J. (2017). Father absence and the educational gender gap.

Ramey, G. and V. A. Ramey (2010). The rug rat race. *Brookings Papers on Economic Activity*, 129.

Yang, G. (2018). Endogenous skills and labor income inequality.

Yang, G. and C. Bansak (2020). Does wealth matter? an assessment of china's rural-urban migration on the education of left-behind children. *China Economic Review 59*, 101365.

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

Yum, M. (2016). Parental time investment and intergenerational mobility. Technical report, Working Paper Series.

Zhang, C. and Y. Xie (2018). Does the school cutoff date cause the disadvantage for children born in july and august? *The Journal of Chinese Sociology 5*(1), 4.

# Appendix A Additional Tables and Figures

We present the OLS estimates of the impact of the younger sibling's enrollment status on test scores of the older child in Table A.1. When the younger sibling enters school, there is no statistically significant effect in any of the OLS regressions at the standard levels. Interestingly, compared to the IV regression in Table 2, all other variables are similar in scale and level of significance.

Table A.1: Sibship spillover - OLS

| | (1) Chinese | (2) Math | (3) Chinese | (4) Math | (5) Chinese | (6) Math |
|---|---|---|---|---|---|---|
| Younger Sib in School | 0.876 | 1.327 | 0.726 | 1.215 | 1.114 | 0.960 |
| | (1.556) | (1.640) | (1.577) | (1.654) | (1.868) | (2.000) |
| Elder Sib is Female | | | 3.698*** | 2.928*** | 2.732*** | 1.878* |
| | | | (0.864) | (0.959) | (1.011) | (1.038) |
| Younger Sib is Female | | | 0.430 | 0.171 | 0.624 | -0.284 |
| | | | (1.055) | (1.171) | (1.119) | (1.269) |
| County Children Female Ratio | | | -23.99 | -12.11 | -16.71 | -3.233 |
| | | | (18.07) | (20.23) | (16.49) | (20.16) |
| ln(Edu Expense) | | | | | -0.106 | -0.108 |
| | | | | | (0.483) | (0.559) |
| Observations | 671 | 671 | 671 | 671 | 574 | 574 |

Notes: This table presents the OLS estimates of the main regression. Each column is a separate regression. The dependent variable is either Chinese test score or Math test score, in 100-point scale. The key variable of interest is a dummy variable of younger sibling enrolled in school. Column (1) and (2) are the results of the baseline model, including both sibling's birth weight and age, both parent's age, income, and education, the household's ethnicity, province fixed effects. Column (3) and (4) are adding both sibling's gender and county level children girl ratio. Column (5) and (6) further control for the yearly log educational expense on the elder sibling. Column (7) and (8) replace elder sibling's age with grade fixed effects. Column (9) and (10) narrow the age range of younger sibling to age 4-8. All standard errors are clustered at city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We show the first stage results in Table A.2, using the August 31 birth date cutoff through the compulsory education. The coefficients on the IV are all negative and significant. These results are robust across various samples at the 1 percent level of significance.

In Table A.3, we run the preferred specification of our main IV regression with the annual educational expense and an indicator of parent's worry as the dependent variable. Though we control for the annual education expenditure in our main regressions, columns (1) and (2)

Table A.2: First Stage

| | (1) Chinese | (2) Math | (3) Chinese | (4) Math | (5) Chinese | (6) Math |
|---|---|---|---|---|---|---|
| Younger Sib Ineligible | -0.623*** | -0.623*** | -0.624*** | -0.624*** | -0.613*** | -0.613*** |
| | (0.0528) | (0.0528) | (0.0526) | (0.0526) | (0.0629) | (0.0629) |
| First Stage F-Stats | 139.295 | 139.295 | 140.701 | 140.701 | 95.0327 | 95.0327 |
| Observations | 671 | 671 | 671 | 671 | 574 | 574 |

Notes: This table presents the first stage of the IV estimation and the Kleibergen-Paap rk Wald F-statistics, corresponding to Table 2.

provide additional evidence showing that educational expenditure on the elder sibling does not change when the second child begins school. We use an additional variable, an indicator of whether parents "worry" about the child, to measure the possible variation in parent's attention.[9] Although this measure is based on a subjective question, it is worth showing as supportive evidence. Columns (3) and (4) show that having a second child in school does not have changes on parental worry for the elder child.

Table A.3: Educational expense and parental attention to the elder when the younger starts school

| | (1) 2009 Edu Expense | (2) 2008&2009 Edu Expense | (3) 2009 Parents' Worry | (4) 2008&2009 Parents' Worry |
|---|---|---|---|---|
| Younger Sib in School | -0.462 | -0.225 | -0.0659 | 0.0126 |
| | (0.394) | (0.300) | (0.108) | (0.0699) |
| Observations | 610 | 1485 | 671 | 1628 |

Notes: This table shows estimates from regressions with dependent variable as either parents' educational expenditure on or "worry" about the elder child. All regressions include the same set of independent variables in columns (5) or (6) in Table 2. In columns (1) and (2), we examine what happens to total education expense on the elder child when the second child enters school. In columns (3) and (4), we use self-reported parental worry as an indicator for changes in parental attention to the child. The odd-numbered columns include the 2009 sample, same as the one in our main regressions, and the even-numbered columns include both 2008 and 2009 waves, to show robust results. Again, we only have test scores for one year (2009) but have most of other variables for both years.

[9] "worry" is an index created by the parents' response on whether they worry about the child's future, in particular, on the aspects of school performance, attending classes, completing homework, spending too much time in internet cafe, watching TV, or playing computer games, being bullied by others, having bad friends, and/or having girlfriend/boyfriend too early.