

The effect of a sibling on the first-born child's health: evidence from two-child families in China

Qundi Feng, Chung-Ping A. Loh, Fancun Meng, Tao Bu & Qinying He

To cite this article: Qundi Feng, Chung-Ping A. Loh, Fancun Meng, Tao Bu & Qinying He (2022) The effect of a sibling on the first-born child's health: evidence from two-child families in China, *Economic Research-Ekonomski Istraživanja*, 35:1, 676-691, DOI: [10.1080/1331677X.2021.1931912](https://doi.org/10.1080/1331677X.2021.1931912)

To link to this article: <https://doi.org/10.1080/1331677X.2021.1931912>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 18 Jun 2021.



Submit your article to this journal [↗](#)



Article views: 1303



View related articles [↗](#)



View Crossmark data [↗](#)

The effect of a sibling on the first-born child's health: evidence from two-child families in China

Qundi Feng^a , Chung-Ping A. Loh^b, Fancun Meng^c, Tao Bu^d and Qinying He^a 

^aCollege of Economics & Management, South China Agricultural University, Guangzhou, P.R. China; ^bDepartment of Economics and Geography, Coggin College of Business, University of North Florida, Jacksonville, FL, USA; ^cSchool of Public Health, Shantou University, Shantou, P. R. China; ^dSchool of Economics and Management, Beijing Jiaotong University, Beijing, P. R. China

ABSTRACT

The first-born child's quality may be affected by a younger sibling in a family based on the quantity-quality trade-off theory. Using data from the China Health and Nutrition Survey, we examine the causal effect of having a younger sibling on the health of the first-born child aged 2–12 in China. We use instrumental variables to address the potential endogeneity of having a younger sibling in the extended regression model. We found that having a sibling significantly decreases the height-for-age z-scores of the first-born child, and the greater age gap may alleviate the effect. Further analysis shows that the effect is particularly strong for the preschool child under 6 years old and the child in a low-income family or the rural area. A sibling influences the first-born child's health by dietary pattern, physical activities, and medical services utilization. The robustness checks, based on individual fixed-effects model and propensity score matching approach, validate our findings, which suggest that future preventive intervention on the deterioration of first-born child's health during the implementation of the universal two-child policy.

ARTICLE HISTORY

Received 24 September 2020
Accepted 13 May 2021

KEYWORDS

Health; children; sibling; two-child policy; extended regression model; China

JEL CLASSIFICATIONS

I12; J13; O15

1. Introduction

Increased attention has been paid to discuss the demographic, health, and social effects of the universal two-child policy in China¹ (Gong et al., 2016; Zeng & Hesketh, 2016; Li, 2016; Zhao, 2017; Liu, 2020). Implementing the two-child policy may help increase China's birth rate, remove the oppressive elements of the one-child policy, and accumulate human capital for economic growth (Zhai, 2014, Zeng & Hesketh, 2016; Li et al., 2019). However, the quantity-quality trade-off (QQ) theory suggests that children with more siblings tend to have worse outcomes in education and health, the two main factors of human capital (Becker & Lewis, 1973; Becker & Tomes, 1976; Li et al., 2008; Angrist et al., 2010). Although we cannot directly

CONTACT Fancun Meng  mfc_105@163.com

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

observe the consequences of implementing the two-child policy so far, analyzing the impact of a younger sibling on the first-born child's health outcome with data of existing two-child families is of interest².

Although the QQ theory suggests that a reduction in child quantity may improve child quality under the resource constraint (Becker, 1960; Becker & Lewis, 1973; Becker & Tomes, 1976), the majority of previous studies focused on the impact of sibling size on child quality (Li et al., 2008; Liu, 2014; Liang & Gibson, 2018). Some empirical research supported a negative association between sibling size and children's academic achievement (e.g., Wolfe & Behrman, 1986; Hanushek, 1992). However, Black and Salvanes (2005) found that the effect of sibling size on the child's quality is not monotonic across the number of children. Qian (2013) even found that an additional child significantly increased school enrollment of first-born children.

Most recent studies isolate the causal effect of sibling size on child quality using a natural experiment or instrument variable (IV) method (Rosenzweig & Zhang, 2009; Angrist et al., 2010; Zhang et al., 2016; Hai, 2017; Argys & Averett, 2019). One important method for tackling endogeneity is to use the exogenous variations in sibling size caused by the natural occurrence of twins or the one-child policy (Li et al., 2008; Qian, 2013; Liu, 2014). The majority of studies use the gender composition of the first two children (Angrist et al., 2010) or the gender of the first child (Lee, 2008) as the instrument for sibling size. The former instrument's idea is that parents with several same-gender offspring are more likely to have an additional child (Angrist & Evans, 1998). The latter instrument is based on the general preference for sons in Asian countries, which indicates that the parents would like to have a second birth, even third birth until they have a son if the first child was female (Lee, 2008).

Encouraged by China's two-child policy, couples tend to have a second child, and then the resource may be diluted for the first child within the family. Some empirical studies test the QQ theory from dietary quality based on the China Health and Nutrition Survey (CHNS). Liang and Gibson (2018) found children's relative energy and protein intakes decrease due to an increase in sibling size, especially for the first-born children. Chen et al. (2019) found that the weight/height and nutrient intake of first-born girls, but not first-born boys, significantly decrease with an exogenous increase in child quantity due to the relaxation of the one-child policy³.

To predict the impact of the two-child policy, this paper will analyze the effect of a younger sibling on a first-born child's health outcomes with data from the 1991–2015 CHNS using an IV. We estimate the impact by using an extended regression model (ERM), which is a more suitable approach for binary endogenous variables⁴. We further employ the individual fixed-effects (FE) panel model to control for time-invariant cross-household heterogeneity and propensity score matching (PSM) method to estimate the causal effects in a sample without random placement. We also test the heterogeneity of this effect across the income distribution, age, gender, and residential household registration (*Hukou*). We find that having a younger sibling significantly reduces height-for-age z-scores of the first-born child, particularly for the child under 6 years old and the child in a low-income family or the rural area. Furthermore, the effect of a younger sibling on the first child's health may be alleviated by the greater age gap. Moreover, we investigate the channel by which having a

younger sibling influence the first-born child's height and find that a child's health outcomes can be explained by changes in dietary patterns, physical activity, and utilization of medical services due to the family resource constraint. Our results provide empirical evidence of the QQ theory in China.

We highlight two contributions to the existing literature on children's health development. First, we predict the health effect of the two-child policy by combining one-child and two-child families with studying the younger child's impact on anthropometric of the first-born child. Second, besides physical activity, and utilization of medical services, we test the response of dietary diversity and nutrient intake to an increase in sibling size due to the dilution of family resources, including family income and parents' accompany time.

The rest of this article is organized as follows. [Section 2](#) describes our data and descriptive analysis. [Section 3](#) illustrates the econometric method. [Section 4](#) presents the empirical results. In [Section 5](#), we do some robustness checks and mechanism analysis. We conclude in [Section 6](#) with a discussion and remarks.

2. Data and descriptive analysis

The data used in this article are from CHNS, an ongoing international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. The survey covers individuals in about 7200 households from communities with diverse socioeconomic backgrounds in 15 provinces and major cities⁵. The Chinese provinces and mega-cities vary substantially in geography and economic development. The CHNS data provide detailed economic, demographic, and health information in ten waves during 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015 so they cover years during the era of rapid economic transition that witnessed considerable health development in children.

The survey uses a multistage random cluster sampling method based on different income levels (high, medium, and low) and weighted sampling. After randomly selecting four counties and two cities with each province, the CHNS randomly identifies villages and towns in each county, urban, and suburban region in each city. It then selects 20 households from each of these communities. In this study, we use the subsample of CHNS that was collected for the children at the time of the survey. For our purposes, we first keep the first-born children aged 2–12 at the time of the survey⁶. Second, we do not include the children from 1989 in our sample because the health and nutritional data in wave 1989 were collected only from preschoolers and adults aged 20–45 and the questionnaire and sampling were substantially different from those used in subsequent waves. Finally, we also exclude the families with three or more children in the main analysis and the sample size is 5879⁷, 33.83% of them have a younger sibling. In the empirical analysis, all of the data are pooled to identify the effect of having a younger sibling.

Importantly, the CHNS records anthropometrical measurements, including height and weight, for each individual⁸. In general, there are three measures of children's health outcomes (Shao, 2007). The first is clinical indicators, such as child mortality,

Table 1. The descriptions of key variables for first-born children with or without a sibling.

Variable name	Without a sibling			Having a sibling			Difference
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
HAZ (height-for-age z-score)	3981	-0.352	1.285	1898	-1.024	1.201	0.672***
Boy (1 = boy; 0 = girl)	3981	0.565	0.496	1898	0.450	0.498	0.115***
Age (year)	3981	7.089	2.891	1898	8.329	2.543	-1.240***
Hukou (1 = rural; 0 = urban)	3981	0.594	0.491	1898	0.858	0.349	-0.265***
Log of household income per capita	3981	8.653	1.056	1898	7.933	1.000	0.720***
Mother's age (year)	3981	31.785	5.000	1898	31.617	3.860	0.168*
Mother's BMI	3981	22.255	3.161	1898	22.071	3.163	0.184**
Mother's educational years (year)	3981	9.282	3.294	1898	6.957	3.279	2.325***
Carbohydrates (g):three-day average	4208	204.093	86.216	1926	263.628	100.882	-59.535***
Fat (g):three-day average	4208	52.158	33.631	1926	45.02	28.414	7.136***
Protein(g):three-day average	4208	48.010	27.919	1926	48.967	18.391	-0.957***
Dietary variety score: three-day average	2912	2.559	2.208	937	1.774	1.898	0.785***
Having medical insurance (1 = yes; 0 = no)	4450	0.500	0.5000	2084	0.309	0.496	0.191***
Children age <6: does physical exercise?	1133	0.553	1.332	173	0.428	1.04	0.125
Children age 6+: does physical exercise?	1987	0.424	1.021	873	0.260	0.778	0.164***

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Having a sibling means that the first-born child has a younger sibling at the time of survey.

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

child morbidity, injuries, etc. the second is self-reported health status. The last is anthropometrical measurements, such as height, weight, waistline, etc. In this paper, we use height as a child health outcome because it is a sensitive measure of health in childhood that reflects the interaction of genetic potential for height with nutrition and their long-term health status, exposure to infectious diseases, and access to medical facilities. To make height comparable among children of different ages and gender, we use height-for-age z-scores (HAZ), which is defined as the number of standard deviations that a person's height is away from the median height of a reference population of healthy children of the same age and sex. We use the information extracted from the growth chart published by World Health Organization (WHO) as the reference height distribution. The negative mean values of HAZ in Table 1 mean that the average height of Chinese children is lower than the average level according to growth standards for children specified by WHO.

We used three-day average carbohydrates, fat, and protein in this paper to measure the children's nutrition intake. The CHNS respondents kept three-day records of their food consumption about the nutrition intake, in terms of meals per person per day and daily food intake. This information was then checked by the research team and reported as three-day average daily intakes for fat, protein, carbohydrates, and total energy from all sources, with energy measured in kilocalories (kcal) and the other three items in grams.

In addition, we computed the children's dietary diversity score based on the food consumption for the same three consecutive days. First, we categorized the food items into eight major food groups based on WHO's Young Children Feeding guidelines (IYCF) and Dietary Guidelines for Chinese Residents (Swindale & Bilinsky, 2005). The eight major food groups are: (i) grains, roots, and tubers; (ii) vegetables; (iii) fruits; (iv) dairy products; (v) legumes and nuts; (vi) flesh foods (meat, poultry); (vii) seafood; (viii) egg. Second, calculate the score for each food group. The score of a child is zero if food intake in a group is below the minimum limit value (the

minimum limit value of soybean food is 5 g, and the others are 25 g); otherwise, the score is one. Third, the group scores are then summed to obtain the dietary diversity score, which ranges from zero to eight. We only got dietary diversity scores of children from 1997 to 2011 because Chinese Food composition before 1997, and the nutrient intake information from the 2015 CHNS data are not available.

In terms of physical activity, it was the indicator of whether a child usually does physical exercises, such as running, playing football, volleyball, badminton, or other sports. As the CHNS data on physical activity apply to two different age groups (<6 or 6+ years), our analysis used these two different age groups.

A two-sample *t*-test is used to formally test whether there is a statistical difference in characteristics between children with or without a younger sibling. The corresponding results are given in the last column of Table 1. It is shown that children with a younger sibling have lower HAZ than those without a more youthful sibling on average. The children having a younger sibling are more likely to come from a family with lower income, lower educational level, and elder mother. The average carbohydrates, protein intake and dietary score are higher for children without a sibling, who are also more likely to have medical insurance and perform well in doing physical activity.

3. Empirical method

To assess the relationship between younger sibling and first child's health outcome, our general estimation approach is given below following Liu (2014) and Chen et al. (2019):

$$Y_{it} = \beta_0 + \beta_1 Ysiblings_{it} + \beta_2 X_{it} + \beta_3 Z_{it} + \varepsilon_{it} \quad (1)$$

where Y_{it} is the HAZ for child i in year t . $Ysiblings_{it}$ is a binary variable, which equals to one if child i has a younger sibling at wave t and zero otherwise. The vector X_{it} has a set of child characteristics, including gender, age, and *Hukou*. The vector Z_{it} includes household per-capita income, age, body mass index, and years of education of the mother. The dummy variables of province and survey year are also controlled. β_1 is the parameter of interest, which measures the change in the first-born child's height-for-age z-score after having a younger sibling. ε_{it} is an error term assumed to be normally distributed.

The fertility decisions are potentially endogenous if some unobservable variables such as parent's preference, genetic, potentially affect both having a second child and the height of the first child. First, to alleviate the endogeneity, we control a rich set of covariates that may affect the decision to have a second child. For instance, the mother's education level affects a child's dietary pattern and physical activities and thus a child's height. At the same time, the mother's education level may also affect the fertility decision. Similarly, the mother's age at the time of the child's birth may be both a determinant of having a second child and a height outcome predictor. Second, we use IV to deal with the possible endogeneity problem. The IV should be correlated with having a younger sibling while not directly affecting the child's height. Specifically, the proportion of the families with two children excluding the family of

child i to the total number of families in the same community is used to serve as the IV for the dummy variable of having a younger sibling. The proportion does not affect a child's height outcome but is correlated with sibling size⁹. Third, we use individual fixed effects panel models to examine how sibling size affects child health outcomes in height because it can avoid the endogeneity problem induced by time-invariant unobservable factors that confound comparisons of outcomes. Finally, we use the PSM method to estimate the causal effects in a sample without random placement.

4. Empirical result

4.1. Impact of having a sibling on first-born child's health

Table 2 represents the estimation results for the effect of having a younger sibling on the first child's HAZ with two models. The OLS regression results are given in Column (1) of Table 2, which shows that the younger sibling has a significant negative effect on a child's HAZ.

Table 2. Effect of having a sibling on the first-born child's HAZ.

Variables	(1) OLS HAZ	(2) ERM		(3) OLS HAZ
		1st stage Having a sibling	2nd stage HAZ	
Having a sibling	-0.178*** (0.0368)		-0.810*** (0.124)	
IV		1.720*** (0.0984)		
age gap				0.0351*** (0.0129)
Boy	0.0592** (0.0294)	-0.513*** (0.0432)	-0.00985 (0.0334)	0.0670 (0.0488)
Age	0.0423*** (0.00716)	0.239*** (0.0117)	0.0744*** (0.0100)	0.0384*** (0.0132)
<i>Hukou</i>	-0.193*** (0.0367)	0.502*** (0.0596)	-0.0876* (0.0448)	-0.261*** (0.0708)
Log of household income per capita	0.0572*** (0.0185)	-0.233*** (0.0277)	0.0135 (0.0206)	0.0900*** (0.0303)
Mother's age	-0.00214 (0.00428)	-0.0614*** (0.00687)	-0.0105** (0.00464)	-0.0150* (0.00839)
Mother's BMI	0.0172*** (0.00498)	0.0133* (0.00722)	0.0175*** (0.00504)	0.0100 (0.00804)
Mother's years of education (year)	0.0517*** (0.00501)	-0.0456*** (0.00753)	0.0420*** (0.00544)	0.0465*** (0.00804)
Constant	-1.994*** (0.260)	-0.00362 (0.549)	-1.331*** (0.290)	-1.154*** (0.387)
Observations	5,879	5,808	5,808	1,898
Control of wave dummies	YES	YES	YES	YES
Control of province dummies	YES	YES	YES	YES
Durbin-Wu-Hausman stat.			16.398***	
F-stat.			499.907***	

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Robust standard errors are in parentheses. Having a sibling means that the first-born has a younger sibling at the time of the survey. The value of Durbin-Wu-Hausman stat. and F-stat. show no presence of weak IV. That the sample size in Column (1) is larger than that in Column (2) was caused by the missing of the instrument variable.

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

Considering that having a younger sibling is a binary endogenous variable, an ERM is used and the corresponding estimation results are reported in Column (2). The Durbin-Wu-Hausman statistic is 16.398 and significant, indicating that there exists an endogeneity problem. The first stage regression, a Probit model, is used to predict the probability of having a second child in a family. After controlling for variables such as the age of the child, gender, *Hukou*, and mother's education, the IV coefficient is significant and positive, indicating no presence of weak IV. Families in a community always have similar fertility concept that is affected by the same culture, such as Confucianism encouraging the patrilineal and patriarchal system, which stresses the importance of continuing the family line through male offspring, receiving the family inheritance, and reinforcing male dominance within a family (Short et al. 2001). In addition, families in a community also experience a similar environment of fertility policy. Therefore, a family living in a community with more two-child families is more likely to have a second child.

The magnitude of the coefficient of having a younger sibling from the ERM is greater than that from OLS, pointing to the stronger negative effect of having a younger sibling on the first-born child's anthropometric outcome in height. The HAZ of first-born children decreased by 0.810 standard deviations, meaning that the height of a 5-year-old child decreased by approximately 3.72 cm¹⁰. The value is much larger than the previous study's estimates that focused on the increase in sibling size (Hai, 2017), which is 0.149 standard deviations of HAZ. The mother's education level is also correlated with the child's HAZ positively. The mother's higher educational level would contribute to child health outcomes through better dietary knowledge and feeding behavior (Variyam et al., 1999).

Since the reference group in Column (1) and (2) is the children who have no siblings, we could not include the age gap between two children. In order to obtain the effect of the age gap on the first-born child's health outcome, we only keep the families with two children. We found that the age gap positively affects the first-born child's HAZ in Column (3). Therefore, a younger sibling's effect on a first-born child's health may be alleviated by the greater age gap between first-born children and second-born children.

4.2. Heterogeneity analysis

Table 3 shows the effects of having a sibling on the first-born child's HAZ across the income distribution, age, gender, and *Hukou*. In panel A, we stratify the sample by quintiles based on income per capita. At the bottom 20th quintile, there are strong significant effects on first-born child's health, but the effect is relatively small from 20th quintile onwards. It is shown that the resource constraints faced by low-income families are more stringent.

In panel B, when splitting the sample into three groups by age, we find that the effect is larger for the pre-school child under 2–6 years old, and the effect is relatively smaller for the child above 6 years old. One possible explanation is that the child at a younger age has less adaptability when the external environment changes. In panel C,

Table 3. Impacts of having a sibling on the first-born child's HAZ across the income distribution, age, gender, or *Hukou*.

Panel A: Stratified by household income per capita (quintiles)					
	Bottom 20th percentile	20th–40th percentile	40th–60th percentile	60th–80th percentile	Top 20th percentile
Having a sibling	–0.933** (0.436)	–0.518 (0.373)	–0.495* (0.254)	–0.764*** (0.264)	–0.721** (0.306)
Observations	1,014	1,195	1,310	1,295	994
Panel B: Stratified by age					
	2 <= age < 6	6 <= age < 9		9 <= age <= 12	
Having a sibling	–1.059*** (0.299)	–0.638*** (0.224)		–0.717*** (0.131)	
Observations	1,910	1,799		2,099	
Panel C: Stratified by gender					
	Boy	Girl			
Having a sibling	–0.796*** (0.157)	–0.760*** (0.175)			
Observations	3,068	2,740			
Panel D: Stratified by Hukou					
	Rural	Urban			
Having a sibling	–0.703*** (0.156)	–0.578** (0.270)			
Observations	3,961	1,847			

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Robust standard errors are in parentheses. The additional control here is the age of children, the age of the mother, the body mass index of the mother, the years of education of the mother, the household per-capita income (inflated 2015), as well as the fixed effects of province and wave. Having a sibling means that the first-born has a younger sibling at the time of the survey. All the results above are based on the IV estimation method. The value of Durbin-Wu-Hausman stat. and F-stat. show no presence of weak IV.

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

we find that the impact of having a sibling is not much different between boys and girls.

In panel D, the result shows that having a sibling has a significantly negative effect on the first-born child's HAZ in rural and urban areas. The negative effects are stronger in rural areas than in urban areas. From the result above, the impacts of having a sibling on the first-born child seem to be greater when the child is in a relatively weak situation, such as low-income family, rural areas, and younger age.

4.3. Mechanism analysis

We test the channel by examining the impact of a younger sibling on dietary patterns, physical activity, and utilization of the first-born child's medical services. The dietary pattern (i.e., dietary diversity and nutrient intake), physical activity, and utilization of medical services are measures of parental investment in children and important inputs into human capital that have their own long-term impacts (Behrman et al., 1988).

Table 4 illustrates the effect of having a younger sibling on the first-born child's dietary patterns. The dietary diversity score, intake of fat, and protein significantly decrease by 0.808, 13.01 g, and 1.855 g, respectively, but carbohydrates significantly increase by 34.83 g. The effects are stronger for children in rural areas. The findings support that the family with more children will face more resource constraints, and will reduce dietary diversity, and consume more food with more carbohydrates

Table 4. Impact of having a sibling on first-born child's dietary pattern.

Variables	(1) ERM Dietary diversity score	(2) ERM Carbohydrates (g)	(3) ERM Fat (g)	(4) ERM Protein (g)
Having a sibling	-0.808*** (0.0853)	34.83*** (4.631)	-13.01*** (2.347)	-1.855*** (0.389)
Boy	-0.0989*** (0.0293)	20.19*** (0.423)	1.260 (1.018)	3.755*** (0.908)
Age	0.0492*** (0.00769)	12.37*** (0.793)	3.445*** (0.158)	3.033*** (0.0933)
<i>Hukou</i>	-0.0782*** (0.0371)	5.689*** (1.344)	-4.439*** (1.558)	-2.532*** (0.486)
Log of household income per capita	0.0539*** (0.0161)	0.0437 (1.929)	3.185*** (0.485)	1.792*** (0.202)
Mother's age	-0.00491 (0.00393)	0.961*** (0.311)	0.00174 (0.134)	0.0663 (0.0622)
Mother's BMI	-0.00387 (0.00445)	-0.548*** (0.197)	-0.0159 (0.161)	0.111*** (0.0220)
Mother's years of education (year)	0.0171*** (0.00536)	-2.225** (0.950)	0.614*** (0.188)	0.371*** (0.134)
Constant	0.158 (0.252)	140.9*** (22.27)	-7.071 (5.743)	7.263*** (2.074)
Observations	3,784	5,888	5,888	5,888
Control of wave dummies	YES	YES	YES	YES
Control of province dummies	YES	YES	YES	YES
Durbin-Wu-Hausman stat.	18.478***	68.909*	85.820*	361.71**
F-stat.	141.673***	55.895*	55.895*	55.895*

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Robust standard errors are in parentheses. Having a sibling means that the first-born has a younger sibling at the time of the survey. All the results above are based on the IV estimation method. The value of Durbin-Wu-Hausman stat. and F-stat. show no presence of weak IV.

Source: The data in column 1 is from 1997 to 2015 CHNS, while the data in columns 2–4 are from 1991 to 2015 CHNS.

but less fat and protein. Diets with more fat and protein are traditionally seen as better nutrition than carbohydrate-rich diets in China (Zhang et al., 2016). The higher income elasticity for protein and fat also reflects this. Moreover, the level of economic development in rural areas is lower than that in urban. Thus the dietary patterns of rural children are relatively single, and they consume more staple food.

In addition, utilization of medical services and physical activity are also important to child health. Children engaged in sports activities show better morphological characteristics, as well as a lower body mass index, skin folds, and body fat, than children who do not play sports (Ak et al., 2020). Table 5 presents the effect of a sibling on the first-born child's utilization of medical services and physical activity. The results are based on IV regressions. Medical insurance and activity are both binary variables, which equal to 1 if the first-born child has medical insurance or does physical exercise.

From Table 5, we find that having a younger sibling significantly reduces the first-born child's probability of having medical insurance and doing physical exercise. An extra child makes families facing greater budget constraints, and then the parent would not buy medical insurance for their children. Moreover, the second child's birth will also reduce the time that parents spend on taking care of the first-born child, so the first child may do less physical exercise without the companion of their parents. It is worth noting that the effect on formal physical activity is insignificant

Table 5. Impact of having a sibling on first-born child's medical insurance and activity.

Variables	(1)	(2)	(3)
	ERM Medical insurance	ERM Activity(Age < 6)	ERM Activity(Age> =6)
Having a sibling	-1.135*** (0.101)	-0.0789 (0.494)	-0.824*** (0.181)
Boy	-0.109*** (0.0380)	0.0395 (0.0806)	0.00823 (0.0595)
Age	0.0615*** (0.0103)	0.302*** (0.0413)	0.123*** (0.0176)
<i>Hukou</i>	0.217*** (0.0496)	-0.305*** (0.0948)	0.0724 (0.0702)
Log of household income per capita	0.126*** (0.0227)	0.146*** (0.0488)	-0.0235 (0.0295)
Mother's age	-0.00858 (0.00565)	0.00137 (0.0122)	-0.0102 (0.00790)
Mother's BMI	0.0224*** (0.00600)	-0.00508 (0.0117)	-0.00753 (0.00826)
Mother's years of education (year)	0.00819 (0.00657)	0.0491*** (0.0164)	0.0167 (0.0107)
Constant	-1.979*** (0.353)	38.00* (19.89)	-93.18*** (10.25)
Observations	6,456	1,288	2,805
Control of wave dummies	YES	YES	YES
Control of province dummies	YES	YES	YES

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Robust standard errors are in parentheses. Having a sibling means that the first-born has a younger sibling at the time of the survey. All the results above are based on the IV estimation method. The value of Durbin-Wu-Hausman stat. and F-stat. show no presence of weak IV. Our physical activity analysis includes two different age groups (<6 or 6+ years) because CHNS questionnaires divide the sample into these two different age groups.

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

for a child below 6-year-old because a child younger than 6-year-old cannot do lots of formal physical activity indeed.

What's more, we also test the channel by examining the impact of dietary pattern, physical activity, and medical services utilization on the first-born child HAZ¹¹. The result indicated that a child with a higher dietary score has a higher HAZ, which is valid for children's nutrition intake of carbohydrates, fat, and protein. Furthermore, a child with medical insurance also benefits his/her HAZ. However, physical exercise to obtain a higher HAZ does not seem to be effective for children, which the limited sample size may bias results.

4.4. Robustness check

In this section, we carry out four stands of robustness check in Tables 6 and 7. The estimation results from the individual fixed-effects models in Column (1) of Table 6 also show a significant negative effect of having a younger sibling, which is consistent with the IV estimation. In addition to HAZ, weight-for-age Z-scores (WAZ) is another important health outcome (Shao, 2007). Thus, we examine the effect of a younger sibling on a child's health outcome in weight. The estimation results are presented in Column (2) of Table 6, which shows that the WAZ of the first-born child significantly decreased by 0.638 standard deviations, meaning that the weight of a 5-year-old child decreased by approximately 1.91 kg.¹²

Since there are more out-of-plan children in reality, where the families with two or more children are very common, particularly in rural China, we expand the sample

Table 6. The estimation results of robustness check.

VARIABLES	(1)	(2)	(3)
	FE HAZ	ERM WAZ	ERM HAZ
Having a sibling	-0.0613* (0.00506)	-0.638*** (0.105)	-
Number of siblings	-	-	-0.566*** (0.0884)
Boy	-	0.0975*** (0.0331)	-0.0137 (0.0327)
Age	-0.438 (0.269)	-0.00937 (0.0109)	0.0673*** (0.00909)
<i>Hukou</i>	-	-0.0865** (0.0437)	-0.115*** (0.0434)
Log of household income per capita	-0.00373** (0.000187)	0.0261 (0.0211)	0.0245 (0.0199)
Mother's age	-0.00661 (0.00776)	-0.00923 (0.00565)	-0.00893** (0.00439)
Mother's BMI	0.0227 (0.0237)	0.0519*** (0.00631)	0.0169*** (0.00490)
Mother's years of education (year)	-0.0204 (0.00671)	0.0309*** (0.00581)	0.0420*** (0.00508)
Constant	-2.582 (1.097)	-1.220*** (0.313)	-1.435*** (0.287)
Observations	6,467	4,548	6,130
Number of idind	3514	-	-
R-squared	0.065	-	-
Control of wave dummies	YES	YES	YES
Control of province dummies	-	YES	YES
Durbin-Wu-Hausman stat.	-	6.168***	25.5189***
F-stat.	-	357.857***	722.743***

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. Robust standard errors are in parentheses. In Column (1), we used a fixed-effects model to estimate the effect of having a younger sibling on a child's HAZ. In Column (2), we estimated the effect of having a younger sibling on a child's WAZ. In Column (3), we estimated the effect of the number of siblings on a child's HAZ.

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

Table 7. Impact of having a sibling on first-born child's health: estimated by PSM.

	HAZ		WAZ	
	kernel matching	Nearest neighbor matching	kernel matching	Nearest neighbor matching
ATT	-0.2578***	-0.2571***	-0.2784***	-0.2170***
Observations				
Treated group		1,898		1,331
Untreated group		3,981		3,256

Notes: *, ** and *** represent significance at the 10%, 5% and 1% levels, respectively. The observed covariates are the age of children, the age of the mother, the body mass index of the mother, the years of education of the mother, the household per-capita income (inflated 2015), as well as province and wave. Two groups are identified: those children with a younger sibling (treated group) and those who do not have a younger sibling (untreated group).

Source: Authors' calculations using data from the China Health and Nutrition Survey (1991–2015).

to families with at least two children. The results are given in Column (3) of Table 6, which indicate that the number of a sibling has a negative and significant effect on the first-born child's height. The magnitude of the effect (0.566) is much smaller than the finding in Table 2. Our findings are consistent with those in Qian (2013) examining sibling size's effect on academic performance when considering birth order.

Furthermore, we use the PSM method to assess potential differences in health outcomes between the first-born child who with a younger sibling and those without a

younger sibling. PSM balances the distributions of observed covariates between a treatment group and a control group based on the similarity of their predicted probabilities of having a given facility (their “propensity scores”). The key to using PSM is to conduct a control group and a treatment group. We therefore use the control variables in [Table 1](#) to construct the control group and treatment group. A balance check is conducted. It is shown in general, that although the systematic difference between the control group and treatment group exists in the unmatched sample, there is no systematic difference in the control variables between the control group and treatment group in the matched sample. The P-value is significant and insignificant in the unmatched sample and in the matched sample, respectively. The results in [Table 7](#) indicate that the first-born child with a younger sibling has significantly lower height and weight than those without a younger sibling, which is also consistent with the IV estimation. Overall, we show that the estimated effect of having a sibling on the first-born child’s health outcome is robust.

5. Conclusions and discussion

This paper estimates the effect of having a younger sibling on the first-born child’s health outcome in height, and it reveals the channels by which the height outcome is affected. We address the endogeneity of having a younger sibling using instrumental variables estimation in ERM based on a sample of 2–12-year-old from the CHNS data.

Our results from the empirical analysis show that having a younger sibling has a significant negative effect on the first-born child’s height-for-age z-scores, and the effect becomes stronger with an increase in sibling size. The effect is particularly stronger for the pre-school child and the child in a low-income family or the rural area. It also provides evidence that having a younger sibling affects the first-born child’s health mainly through decreased dietary patterns, intake of fat and protein, utilization of medical service, physical activity, and increases in carbohydrates. Furthermore, the robustness checks validate our conclusions and strengthen our arguments.

Previous research paid more attention to the overall effect of sibling size, so they ignored the specific effect on individual children in different birth order. Although our study focuses on the first-born children with a younger sibling, our result is consistent with some previous findings on the effect of sibling size from other countries, such as [Lee \(2008\)](#), which found that sibling size had adverse effects on investment in education per child in South Korea. However, [Angrist et al. \(2010\)](#) and [Lafortune and Lee \(2014\)](#) found that a child’s schooling can be increased by having more siblings, indicating that the presence of siblings directly enhances child welfare, which may be explained by that child with siblings benefit socially or take on more responsibility. [Black and Salvanes \(2005\)](#) found a negative correlation between sibling size and children’s educational attainment becomes negligible after controlling for birth orders by using a rich data set on Norway’s entire population. Therefore, it is reasonable to consider birth order when testing the QQ theory.

We consider the effect of the second-born children on first-born children’s health outcome, whose meaningful results have more detailed implications. These findings

extend the literature on child health in developing countries. The advantage of this study provides new evidence that the first-born child's health would be affected by the younger child, which also supports the theory of quantity-quality trade-off in China. Our findings point out the potential effects of an extra child on the first-born child's health and the health inequality between rural and urban areas under the universal two-child policy in China. To avoid the potential negative effect of an increase in sibling size on children's quality, especially rural children, the government should put more effort into future preventive intervention. Specifically, the government should pay more attention to nutrition programs, such as subsidizing food that is important for children's body growth and providing training programs to increase residents' knowledge of nutrition and health.

Our study has some limitations, which provide directions in which further research might develop. We only used data from 1991 to 2015 of CHNS due to the unavailability of more recent data. Since there are few three or more children families in China so far, the same topic can be searched for the effect on the first and second children in three-child studies, and results can be compared in the future and other countries. Although our study's concept is limited to China, these results may produce useful pieces of information, which might help developing countries in the process of creating policies to improve human capital.

Notes

1. China has brought to an end its one-child policy and replaced it with a universal two-child policy that allows all couples to have a second child since 1st June 2016. The one-child policy was introduced in 1979.
2. The CHNS data shows that one-child families account for about 54%, and two-child families accounted for about 40% in 2015. There exist two-child families during the implementation of the one-child policy because of the following reasons. First, a couple could have a second child, but the violation of the one-child policy incurs severe punishments in the form of unaffordable fines and denied bonuses. Second, rural couples in most provinces are allowed to have a second child if their first child is a girl, the so-called 1.5-child policy. Third, two or more children are allowed for ethnic minorities, who account for around 9% of the total population. What's more, in 2013, China put forward a selective two-child policy that a second child is allowed if either parent originates from a one-child family.
3. Chen et al. (2019) defined the IV based on two questions in the community questionnaire of the CHNS: "For couples of your village/neighborhood, are they allowed to have two children?" and "For couples of Han nationality, are they allowed to have one more child if their first child is a girl?". However, we could not apply this IV in our research because we cannot obtain the community-level data of CHNS, which requires applicants to provide the approval of the local institutional review board or ethics committee.
4. An ERM fits a linear regression model that accommodates any kind of endogenous covariates, such as continuous, binary, and ordinal endogenous covariates. However, the two-stage least square (2SLS) regression is only suitable for continuous endogenous variables to tackle endogeneity problem.
5. Data from Beijing, Chongqing, Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, Shanxi, Shandong, Shanghai, Yunnan, and Zhejiang were collected.
6. We do not include older children because puberty starts about age 13, and growth spurts during puberty and after that can make anthropometric measures less reliable. Another

reason is that children older than 13 years of age are likely to have more control over their own food choices, and their health is less affected by household composition (He et al., 2018).

7. We also excluded children whose absolute values of HAZ exceed 5 from the analysis on the ground that those values are implausible. See <http://www.who.int/en/>. This sample size (5879) is for the analysis of HAZ as dependent variables using OLS.
8. Trained health workers who followed standard protocol and techniques collected the height and weight measurements for all survey participants. Height was measured without shoes to the nearest tenth of a centimeter with a portable stadiometer. Weight was measured in light indoor clothing without shoes to the nearest tenth of a kilogram with a balance beam scale. Each of these measurements was taken by at least two health workers. One worker took the measurement, and the other recorded the readings.
9. The community means either an entire village or only a few blocks in cities. The CHNS sample communities are drawn from cities, suburbs, towns, or villages of China, all entities that are legally identified by the National Bureau of Statistics of China. In survey data, we do not know the community's exact name, but we can see the community code for each household. So the households are from the same community if their community codes are the same.
10. The child's HAZ is converted into height, according to the child growth standards of the WHO. https://www.who.int/childgrowth/standards/technical_report/en/
11. The table of result is available on request.
12. The child's weight-for-age z-score is converted into weight, according to the child growth standards of the WHO. https://www.who.int/childgrowth/standards/technical_report/en/

Acknowledgements

We sincerely thank the anonymous reviewers for their thoughtful comments, but we take full responsibility for all errors in this article.

Disclosure statement

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID

Qundi Feng  <http://orcid.org/0000-0002-4328-3210>

Qinying He  <http://orcid.org/0000-0003-1234-5690>

References

- Ak, P. M., Trajkovski, B., & Jovanovi, I. J. (2020). Analysis of the existing condition and a suggestion for obesity prevention in early school-age children. *Economic Research-Ekonomska Istrazivanja*, 33(1), 1194–1199.
- Angrist, J., & Evans, W. N. (1998). Children and their parent's labor supply: Evidence from exogenous variation in sibling size. *American Economic Review*, 88, 450–477.
- Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28(4), 773–824. <https://doi.org/10.1086/653830>

- Argys, L. M., & Averett, S. L. (2019). The effect of sibling size on education: New evidence from China's one-child policy. *Journal of Demographic Economics*, 85(1), 21–42. <https://doi.org/10.1017/dem.2018.22>
- Becker, G. S. (1960). *An economic analysis of fertility, demographic and economic change in developed countries*. Princeton University Press.
- Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2, Part 2), S279–S288. <https://doi.org/10.1086/260166>
- Becker, G. S., & Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84(4, Part 2), S143–S162. <https://doi.org/10.1086/260536>
- Behrman, J. R., Deolalikar, A. B., & Wolfe, B. L. (1988). Nutrients: Impacts and determinants. *The World Bank Economic Review*, 2(3), 299–320. <https://doi.org/10.1093/wber/2.3.299>
- Black, S. E., & Salvanes, D. K. G. (2005). The more the merrier? The effect of sibling size and birth order on children's education. *Quarterly Journal of Economics*, 120(2), 669–700.
- Chen, C., Chou, S. Y., Wang, C., & Zhao, W. (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*. <https://doi.org/10.1515/bejeap-2018-0340>
- Gong, W., Xu, D., & Caine, E. D. (2016). Challenges arising from China's two-child policy. *The Lancet*, 387(10025), 1274–1274. [https://doi.org/10.1016/S0140-6736\(16\)30020-4](https://doi.org/10.1016/S0140-6736(16)30020-4)
- Hai, Z. (2017). The effect of sibling size on children's health and education: Is there a quantity-quality trade-off? *The Journal of Development Studies*, 53(8), 1194–1206. <https://doi.org/10.1080/00220388.2016.1214720>
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 84–117. <https://doi.org/10.1086/261808>
- He, Q., Li, X., & Wang, R. (2018). Childhood obesity in China: Does grandparents' coresidence matter? *Economics and Human Biology*, 29, 56–63. <https://doi.org/10.1016/j.ehb.2018.02.001>
- Lafortune, J., & Lee, S. (2014). All for one? Family size and children's educational distribution under credit constraints. *American Economic Review*, 104(5), 365–369. <https://doi.org/10.1257/aer.104.5.365>
- Lee, J. (2008). Sibling size and investment in children's education: An Asian instrument. *Journal of Population Economics*, 21(4), 855–875. <https://doi.org/10.1007/s00148-006-0124-5>
- Li, H. (2016). Comprehensive second child and economic development. *Economic Research Reference*, 12, 43–44.
- Liang, Y., & Gibson, J. (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. <https://doi.org/10.1016/j.chieco.2017.10.006>
- Liu, H. (2014). The quality–quantity trade-off: evidence from the relaxation of China's one-child policy. *Journal of Population Economics*, 27(2), 565–602. <https://doi.org/10.1007/s00148-013-0478-4>
- Liu, Y. (2020). Aging and economic growth: is there a role for a two-child policy in China? *Economic Research-Ekonomska Istraživanja*, 33(1), 438–455. <https://doi.org/10.1080/1331677X.2019.1699436>
- Li, H., Zhang, J., & Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography*, 45(1), 223–243. <https://doi.org/10.1353/dem.2008.0006>
- Li, H., Zhou, T., & Jia, C. (2019). The influence of the universal two-child policy on China's future population and ageing. *Journal of Population Research*, 36(3), 183–203. <https://doi.org/10.1007/s12546-019-09228-7>
- Qian, N. (2013). Quantity-quality and the one-child policy: The only-child disadvantage in school enrollment in rural China (NBER Working Paper No. 14973). Cambridge, MA: National Bureau of Economic Research.

- Rosenzweig, M. R., & Zhang, J. (2009). Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies*, 76(3), 1149–1174. <https://doi.org/10.1111/j.1467-937X.2009.00563.x>
- Shao, J. (2007). *The social and economic determinants of child health during China's economic growth*. Dalhousie University.
- Short, S. E., Fengying, Z., Siyuan, X., & Mingliang, Y. (2001). China's one-child policy and the care of children: an analysis of qualitative and quantitative data. *Social Forces*, 79(3), 913–943. <https://doi.org/10.1353/sof.2001.0025>
- Swindale, A., & Bilinsky, P. (2005). Household dietary diversity score (hdds) for measurement of household food access: Indicator guide, food and nutrition technical assistance. *Journal of Nutrition*, 138(12), 2448–2453.
- Variyam, J. N., Blaylock, J., Lin, B. -H., Ralston, K., & Smallwood, D. (1999). Mother's nutrition knowledge and children's dietary intakes. *American Journal of Agricultural Economics*, 81(2), 373– 384. <https://doi.org/10.2307/1244588>
- Wolfe, B. L., & Behrman, J. R. (1986). Child quantity and quality in a developing country: Family background, endogenous tastes, and biological supply factors. *Economic Development and Cultural Change*, 34(4), 703–720. <https://doi.org/10.1086/451555>
- Zeng, Y., & Hesketh, T. (2016). The effects of China's universal two-child policy. *The Lancet*, 388(10054), 1930–1938. [https://doi.org/10.1016/S0140-6736\(16\)31405-2](https://doi.org/10.1016/S0140-6736(16)31405-2)
- Zhai, Z., Zhang, X., & Jin, Y. (2014). Demographic consequences of the immediate and comprehensive liberalization of the second-child policy. *Population Studies*, 38 (2), 3–17.
- Zhang, J., Xu, P., & Liu, F. (2016). One-child policy and childhood obesity. *China Economic Review*, 59, 100938. <https://doi.org/10.1016/j.chieco.2016.05.0r03>.
- Zhao, J., Shan, N., Yang, X., Li, Q., Xia, Y., Zhang, H., & Qi, H. (2017). Effect of second child intent on delivery mode after Chinese two child policy implementation: A cross sectional and prospective observational study of nulliparous women in Chongqing . *BMJ Open*, 7(12), e018823. <https://doi.org/10.1136/bmjopen-2017-018823>