Household Level Determinants of Educational Attainment for Young Females: An Empirical Analysis of Chinese Rural Areas

MPP Professional Paper

In Partial Fulfillment of the Master of Public Policy Degree Requirements
The Hubert H. Humphrey School of Public Affairs
The University of Minnesota

Man Xu
Auguest-30-2016

Signature below of Paper Supervisor certifies successful completion of oral presentation and completion of final written version:

Professor Samuel L. Myers, Jr., Paper Supervisor	Date, oral presentation	Date, paper completion

Professor Deborah Levison, Committee Member	Date

Household Level Determinants of Educational Attainment for Young Females: An Empirical Analysis of Chinese Rural Areas

Man Xu

Humphrey School of Public Affairs, University of Minnesota

xuxx0460@umn.edu

Abstract: This paper investigates household level determinants of educational outcomes for young, rural Chinese females, examining whether there is male-preference and one-child advantage in educational outcomes. Results indicate females from one-child families have an advantage for schooling over females from multi-child families. Based on the birth parity in the natal families, within one-child families, females do not have significantly lower years of schooling than males. However, compared to females from two-child families, females from one-child families significantly complete more schooling years (1.94 years more). Based on the gender and birth order within the sibling pairs, having a younger male sibling will not only reduce the older sister's completed schooling years, but also decreases the probability of attending academic high school. Young females in rural China face disadvantage in both completed years of schooling and achieved educational level.

Key words: Young Female Educational Attainment, Rural China, Sibling Effect, Birth Order

Acknowledgement: This paper is for the Master Professional paper of Humphrey School of Public Affairs, University of Minnesota Twin Cities. Acknowledgements to my thesis committee: Dr. Samuel L. Myers, Jr. (chair) and Dr. Deborah Levison. I am really grateful for your advice and support to my data analysis and writing.

1. Introduction and Literature Review

A key indicator of social and economic development is female education. In many developing regions, gender differences in education and literacy are pervasive (Hannum and Xie, 1994). Many scholars view gender inequality in education and the lack of educational attainment for females as a constraint on economic growth (Song, Appleton and Knight, 2005). Studies of market reforms in China show that promoting marketized education creates more barriers for rural children to access educational opportunities than urban children (Hannum, 1999). Using China's Census data, Hannum (2005) showed that gender disparities are exacerbated in educational attainment, even if young females' schooling is obviously improved by market reforms. In order to understand the scope of the social and economic development of rural China, it is necessary to investigate the educational attainment of females.

With the market reformation and economic development, girls are better off living in one child families in the big cities. Tsui and Rich (2002) suggest that for families with one-child, gender equality in education is an unintended consequence for current urban Chinese. In the poor rural areas, situations are worse for girls. Maternal educational expectations for girls are lower than for boys, and girls face greater likelihood of having to do chores. The reason behind the low expectation of some mothers is their view that investing in girls is a waste (Hannum, Kong and Zhang, 2009) because girls will leave the natal family. Mother's educational expectations of children, especially to girls, have a big influence on their completed years of schooling.

This research paper investigates the household structure determinants of young females' educational attainment in rural China. The paper provides an empirical analysis of the **Chinese Household Income Project 2007** data and focuses on young females between the ages of 16 to 20 years old, the typical age of high school to post-secondary education. The primary research question is whether young females from one-child families complete more years of schooling than young females from two-child families[1] in rural

[1] Family and household are interchangeable terms in the paper.

China. The second research question is within the two-child families, what is the effect of the different sex composition and birth order to the educational outcomes to the children.

Male preference exists in China. Females in China are 'less preferred,' especially in the education and job markets (Andors 1990; Hunnum, Kong and Zhang 2009). Generally, for families with limited resources for education, parents give preference to their sons (Hannum, 2005). This preference is associated with the parents' expectation that the sons will support them in old age, because a social pension program[2] does not exist for the elderly in rural China before 2009[3]. This means that rural families are faced with greater uncertainty about their future financial security and living costs than in urban areas. Parents often count on their sons to return home to support them in their old age (Honig and Hershatter, 1988; Wolf, 1985). This explains why the families invest more in their sons' education than their daughters' in rural China.

With the implementation of the one-child policy in China (beginning in the 1970s)[4], many families started having only one child, therefore in CHIPS 2007 everyone older than 26 years old should be born after the one-child policy. However, according to the *Population and Family Planning Law of the Peoples Republic of China*, starting from 1984 many provinces had their own regulations that allowed couples in rural areas to have a second child if the first born was a female. This exemption is called the "one-and-half child policy"[5]. In the following year, the National Health and Family Planning Commission and the State Council expanded one-child exemptions to the following circumstances: 1) if one parent is of a Chinese ethnic minority group; 2) one parent is employed in a dangerous occupation; 3) if the first child is born with a disability; or 4) either of the parents is an only child. It then became legal for households to have a second

[2] In 1992, state council had been trying to promote the social pension program in rural areas. But the program was abolished in 1998.

[3] http://www.gov.cn/zwgk/2009-09/04/content_1409216.htm

[4]Started from 1970s, Chinese Communist Party implemented limited birth policy in urban areas. The initial birth control campaign which the slogan is "Late, Long and Few," cut population growth by half between 1970 and 1976. On September 25, 1980, the Central Committee of the Chinese Communist Party published "*An Open Letter to Members of the Chinese Communist Party and Chinese Communist Youth League on Controlling Population Growth",* as a landmark of launching the One Child Policy nationwide. In 1982, birth control was adopted as a "basic state policy" (jiben guoce). In 1984, the Chinese government legalized that different provinces were subject to different mandated limits. (Greenhalgh 1986).

[5] Not all of the provinces in China have the one-and-half child policy, Tianjin, Beijing, Sichuan, Chongqing, Jiangsu and Shanghai do not have this policy.

child (Chinese Government 2001). It is not rare for households in rural China to have two or more children. **Table 1** displays the distribution of number of children the natal household has in the **CHIP 2007 rural sample**. It is the unweighted count for CHIP 2007 rural sample, the children who left the household cannot be captured in the dataset, only biological siblings are included in the table. Although 46% of households are from one-child households, there are over 40% of household with two children.

[Insert Table 1 here]

The one-child family places their resources into that only child's development; thus, I expect to find no difference between males' and females' completed schooling years in this case. Previous studies have suggested females show more promise in academic achievements (Brown and Park, 2002; Zhang, Kao and Hannum, 2009). The Situation would be different for non-one-child families. In non-western societies, the strong family obligation among the siblings suggests the first born child can improve the education opportunities for the later born. Worldwide, some researchers have investigated why females have lower education attainment than males do, and some of the research examined the impact of siblings' educational attainment. For example, an empirical study in Taiwan shows a child with older sisters completed more schooling years (Parrish and Willis, 1993). When higher birth order children enter higher school grades, the older sisters have already left for work and can support their younger siblings to obtain longer schooling years (Blake, 1989).

Because of the male preference (Chu 2001) and the one-child and one-and-half-child policy in China, parents have higher likelihood of having a son, especially when the first one or two children are daughters. Ebenstein (2010, 2011) and Banister (2004) used China Census data to determine mothers whose previous child were daughters had higher sex-ratio (ratio of males to females) for next birth. Over 70 percent of mothers with two daughters have a son as their third birth. That is to say there is potential endogeneity of having a male sibling for a child in the household. The sex ratio at birth at 2000 and 2010

China Census was 116.86 and 117.94, respectively.[6] The sex ratio at birth worldwide is 107[7]. In order to explain the higher sex-ratio in China, Chu (2011) suggested an underreporting of the female birth rate, high female infant mortality at birth, and increasing prenatal sex detection and sex-selective abortion of female fetuses underreports the real sex ratio at birth (Banister and Hill, 2004; Chu, 2011). The sex ratio for age group 0-5 and 5-9 in 2000 and 2010 China Census were still over 110, and even reached 115; however, the sex ratio fell to 105 for the people over 15 years old.[8] The CHIP 2007 rural sample sex ratio at the survey time (2008) is 109.41, which is within the normal sex ratio range. And the sex ratio of this paper's interest group (16—20 years old) at the survey time is 102.3 (**Table 2**), which also falls in the normal range of sex ratio.

[Insert Table 2 here]

Birth order among siblings also correlated with the completed schooling years. Arleen Leibowitz (1974) argued that additional numbers of siblings reduce the amount of family support for schooling and indicate lower levels of home investments for each child. For the same reason children of lower birth order (older children) can be expected to receive more time inputs from parents. But in non-western social contexts, later birth order children obtain longer completed schooling years (Parish and Willis 1993; Chu, Xie and Ye, 2007).

China has free nine-year compulsory education, which requires all children to finish primary and secondary school.[9] If the analysis includes children between ages of 6 to 15 years, the results cannot reflect their final completed schooling years because they highly like to continue their post-secondary education when they finish the nine-year compulsory education. After the 9th grade, students have two options for furthering their formal education: 1) study at an academic high school (10th – 12th grade); or 2) study at a technical school or vocational senior secondary school. Students determine which route to follow by taking a citywide examination called the high school enrollment

[6] http://www.nhfpc.gov.cn/jtfzs/s3578/201502/ab0ea18da9c34d7789b5957464da51c3.shtml
[7] https://www.cia.gov/library/publications/the-world-factbook/fields/2018.html
[8] Data source: Online Tabulation on the 2000 and 2010 Population Census of the People's Republic of China.
[9] Compulsory Education Law of the People's Republic of China was enacted from July-1-1986, the last revision was done by Apri-24-2015.

examination. If the students do not pass the minimum admission score, they are not eligible for enrolling in an academic high school, but are rather deferred to the technical school or vocational school.

However, in rural China, some students decided to study in the vocational school even if they pass the high school enrollment examination. Hen et.al (2009) summarized that academic high school is costly and has high academic admission requirements, while vocational senior secondary school offers a more accessible and shorter path to entering the labor markets[10]. According to Education Law of the People's Republic of China[11], vocational schools train students in both specific skills, such as computer repair and general academic skills, such as math, English and Chinese. However, some empirical studies show that vocational students are worse off in both specific and general skills compared to academic high school students. Even the worst academic high school students in their school get higher scores on standardized tests than students attending vocational high school. This suggests that vocational school students do not gain an advantage in both specific skills and general skills over their peers in academic schools (FSI-REAP, 2012-). Studying at an academic high school is therefore a necessary part of getting admitted by an academic college under Chinese educational system, even though students from vocational high school have the eligibility to attend the national college entrance examination. Therefore, distinction between academic high school and vocational schools is pertinent for China as a measure of achieved educational level.

This paper identifies the makeup of a household in terms of number of children has to raise and sex composition of sibling pairs, which stress the household structure effect on educational attainment and achieved educational level for children. It does not merely investigate the gender disparities in the educational outcomes, but it looks into the sibling pairs with different sex composition to understand the gender disparities combined with the birth order, which is less well established literature on investments in educational outcomes for young females in rural China. This paper contributes to the literature in the

[10]http://reap.fsi.stanford.edu/research/vocational_education_versus_academic_high_school_a_study_on_relative_value_added

[11] http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_619/200407/1316.html

educational disparities within the sibling pairs by investigating the sex composition and birth order disparities. The paper proceeds as follows: Section 2 presents the theory of optimal investment in human capital. Section 3 describes the CHIP 2007 rural sample. Section 4 specifies the model specification and estimation issues. Section 5 presents the hypotheses. Section 6 provides main empirical results. Section 7 discusses the possible policy implications and conclusions.

2. Theory

In *A Treatise on the Family,* Becker systematically identified the family role of developing children's human capital. He developed a parental investment model that depends on the abilities, handicaps, sex, luck and other characteristics of children (Becker, 1991). When parents are neutral among their children, they favor their less fortunate children, regardless of their sex, birth order, or other characteristics, because the marginal utility of children with fewer advantageous always exceeds the marginal utility from children with more advantages (Becker, 1991).

For poor families whose parents invest only in the human capital of their children, if each child receives the same amount of investment, the marginal rates of return would be higher for better-endowed children, while the marginal utilities would be higher for the worse-endowed (Becker, 1991). More human capital would be invested in the better-endowed only if differences in rates of return exceeded differences in marginal utilities. Thus, poor families have an incentive to invest more in the child they feel will be the most successful, rather than investing in all children equally. With the altruism of family, transferring resources is a compensating mechanism for less endowed children. Even if abler children were not altruistic, a legal system or social norms could force the abler children to agree to look after their siblings. In this case, the poor families would invest more in abler children.

Therefore, according to Becker, in poorer families the amount invested in human capital directly depends on the abilities and the number of siblings (Becker, 1991). In order to

solve the conflict between the equity and the efficiency of parental investments, a less-endowed child would be better off with abler siblings: 1) if the incomes of different children were not close substitutes in his parents' utility function; 2) if abler children were altruistic to siblings or would agree to compensate siblings; or 3) if less able children were to learn more from abler siblings.

This becomes problematic when parents have strong preferences for a particular child; the usual case is the first son or the only son in the family (Becker, 1991). Under these circumstances, the endowments and altruism are outweighed by the preference. This preference is not only based on love, but also partially determined by the later transfers to the parent when they are in their old age. In rural China, parents have traditionally relied on sons for provision of care in old age since daughters often marry into other families, making support unlikely (Croll, 1994). Parish and Willis (1993) argued that the natal family will underinvest in girls because daughters marry into another family, so the investment in their human capital accrues to the family in-law but not the natal one. Becker's assortative matting theory (Becker, 1991) points to similar argument. Becker hypothesizes that individuals with more educational attainment will tend to marry others with similar or higher educational backgrounds. Females are more likely to marry a primary income provider with comparable or better education, but they will leave the natal family after marriage. And it is uncertain that a daughter would provide support for aging parents. So parents will invest more in the son with a greater likelihood of marrying a comparably educated spouse.

Breaking down Becker's theory for households with different numbers of children will explain the theory more explicitly. For one-child households, if the parents would like to invest in their child's education, no matter the gender of the child, this only child can get all the investments. As for wo-child households, a different sex composition in children brings different outcomes. For two-daughter or two-son families, among the neutral parents, children's own endowments are the key influential factors for educational attainment. On the other hand, birth order could be one source of preferences. Among families that have both a daughter and son, if gender preference is more valued than the lower birth child preference, a younger brother with older sisters would preempt in family

investment, and an older brother with younger sisters would have advantage in family investment. Evidence from Taiwan shows that younger brothers have higher educational attainment than their older sisters because once the older sister has reached the legal working age, they can support the younger brother in furthering schooling (Chu, Xie and Ye, 2007).

3. Data and Samples

The data set employed in this analysis is the **Chinese Household Income Project 2007** (CHIP 2007), which is a cross-sectional data. CHIP includes detailed indicators of household socioeconomic status and the household head answerd questions related to all household members. In the rural sample, each household head is required to keep a diary of the living situations and incomes for their household members (Luo, Li, Sicular, Deng and Yue, 2012). The analysis in this paper is based on the data collected from the 2008 questionnaire, which is reflective of the experiences of 2007. According to the questionnaire, the data was collected in December 2008 by local National Bureau of Statistic (NBS). In the original dataset, all missing values are kept to maintain the completion of the records.

The samples in the CHIP were selected from the National Bureau Statistic's Urban Household Survey (UHS) and Rural Household Survey (RHS). The National Bureau of Statistic (NBS) maintains a national pool from the Urban and Rural Household Surveys, so basically CHIP targets households nationwide. The current UHS/RHS is conducted every three years and chooses the households that are registered in the Bureau's list. Using stratified random sampling, cities and counties are divided into three levels: the prefecture-level (and above) city (Dijishi), the county-level city (Xianjishi), and the county (Xian). The number of surveyed households in each city or county is determined by the proportion of the urban/rural population in the province that lives there. Then, the individuals in the urban/rural area are sorted by average annual wage/income, and the surveyed cities and counties are selected by systematic sampling (Appendix I, 2007).

In the CHIP surveys, neither the urban nor rural sample sizes are proportional to their shares in the national population. Also, their regional distributions differ from those in the population. Thus, weights are applied. Weights are calculated by using data provided by the NBS from the 2000 census and the 2005 1% population sample survey (Luo, Li, Sicular, Deng and Yue, 2012).

This paper selected girls and young female adults between 16 and 20 years old from natal family and keeps only respondents who have valid responses on all relevant questions. The questionnaire asked 'how many years have you already studied in formal school'. The household heads answer questions for their household members. Due to the missing values for the dependent variable and independent variables, the sample size for each variable varies. The paper captured 78.05% of original girls and young female in the dataset. As a comparison group for estimating the educational attainment, the paper also keeps the boys and young males from one-child and two-child families between 16 and 20 years old.

To understand the sibling effect on education attainment, I extract all the young females from two-child families and separate them into four sub-samples based on the gender and birth order of their siblings, hereafter called "sibling pairs". The four kind of the sibling pairs are older sister-younger sister (FF), older brother-young brother (MM), older sister-younger brother(FM) and older brother-younger sister(MF). Due to the data design, I cannot obtain the sibling's information if the sibling already left the natal household. Therefore, some children from two-child families are not categorized into the sibling pairs if the sibling no longer lives in the natal household. The **Table A** in the **Appendix** gives detailed number of children were living in the natal household when the data were collected (year of 2007) and the number of children who ever lived in the natal household. In order to keep as many as sibling pairs as possible, I keep all the children between 16—20 years old even though their sibling is out of the age range. For instance, if an 18-years-old female has a 12-years-old brother, this female is in the sibling pair sample, while her brother is not. **Table B** in the **Appendix** shows the loss of number of sibling pairs if I keep both the siblings pairs between 16 to 20 years old. As a remedy of not strictly controlling for the age range of the sibling's age, I control the age difference

between the sibling pairs in the estimates. The mean value of age difference between sibling pairs is 2.97 years, and the standard deviation is 2.90; these give the clue that 16-20-year-old youth are spaced relatively tightly. The definition of "tight birth spacing" is an age difference between siblings of less than 5 years (Chu, Xie and Yu 2007). The rationale of controlling the age difference between sibling pairs is based on the resource dilution model (Blake, 1981 and 1989) and an empirical study of Taiwanese sibling size of Chu, Xie and Yu (2007). **Table C** of the **Appendix** tabulates the detailed age difference between sibling pairs of CHIP 2007 rural sample.

Considering the widespread existence of education delay in rural China, the analysis widens the high school graduation age from 18 or 19 years old to 20 years old (Duan, Lv, Guo and Wang, 2013). Based on the school options students can choose after graduating from middle school, this paper calculates the probability of finishing middle school (9^{th} grade) and the probability of attending academic high school for children and young adults as the indicator of achieved educational level. I am able to test whether second-born children are more likely to go to high school than first-born children from two-child families. The provincial weights are applied; the Taylor Series are used to calculate standard errors.

4. Model Specification and Estimation Issues

The primary interest of this paper is in equation (1), where the dependent variable is the reported completed years of schooling for children from one-child and two-child families and the reference group is females from one-child families. Independent variables include the endowments of parents (measured by completed years of schooling), the family economic situation, and the family structure. The family structure is measured by whether or not the child is the only child in the family and by gender of the sibling. **Table 3** shows the definition of each independent variable used in the estimates; **Table 4** reports the basic descriptive statistics for both dependent and independent variables and **Table 5** presents the means difference test results for some selected independent variables.

Equation (2) and (3) calculate the probability of finishing 9th grade (finishing middle school) and the probability of attending high school. The equations are as the following:

$$Edu=\beta_0+\beta_1 Edu_{father}+\beta_2 Edu_{mother}+\beta_3 HH_{inc}+\beta_4 HH_{onlychild} \quad (1)$$

$$P_{9thgrade}=\beta_0+\beta_1 Edu_{father}+\beta_2 Edu_{mother}+\beta_3 HH_{inc}+\beta_4 Siblingpair+\beta_5 SiblingSpacing \quad (2)$$

$$P_{highschool}=\beta_0+\beta_1 Edu_{father}+\beta_2 Edu_{mother}+\beta_3 HH_{inc}+\beta_4 Siblingpair+\beta_5 SiblingSpacing \quad (3)$$

[Insert Table 3 here]

[Insert Table 4 here]

170 females from one-child families had an average of 9.88 completed years of schooling and 822 females from two-child families had an average of 8.92 completed schooling years. The distributions of parental education attainments reveal an attainment advantage for fathers of two-child families. Generally, the mean values for parental completed schooling years were less than 9 years, which, equivalently speaking, is less than middle school. The CHIP 2007 survey sampled a smaller proportion of minorities than the census. The 2000 Census published that 8.41% of the Chinese population are minorities and the 2010 Census published number is 8.49%.[12] In the CHIP 2007 rural sample only 1.99% of observations are minorities. The sex ratio of siblings is unbalanced among two-child families. Sixty-seven percent of 16-20-year-old females from two-child families have a male sibling, which corresponds with the male-preference in rural China. Except for father's age, no significant means difference exists in the selected household characteristics between one-child and two-child families. Females from one-child families are similar with females with siblings in the controlled independent characteristics.

[Insert Table 5 here]

[12] Data source: Online Tabulation on the 2000 and 2010 Population Census of the People's Republic of China.

5. Hypotheses

Based on the theoretical framework stated in the first and second sections, this empirical analysis aims to test the following hypotheses:

Hypothesis 1	One-child females completed more years of schooling than females from other families.
Hypothesis 2	For males and females from one-child families, the educational attainment gap disappears.
Hypothesis 3	In sibling pairs, the second-born children receive more years of schooling than first-born children.
Hypothesis 4	First-born daughter and first-born son have different directional effects on their younger sister's completed schooling years.
Hypothesis 5	Second-born sons enjoy an advantage in completed schooling years due to both male preference and the late birth order.
Hypothesis 6	Second-born sons have higher probability of attending high school than their older sisters.

Hypothesis 1 addresses the general issue of education disparity between one-child and other families, when households have limited resources to invest in children's education.

Hypothesis 2 emphasizes that young females' deficit in completed schooling years will be eliminated if the household only has one child. Household will allocate all the educational resources to this only child, such that being a female will not be predictably associated with a decrease in an only child's completed schooling years.

Hypothesis 3 highlights the effect of having a sibling. The competition from a sibling deprives the other child of education opportunities when the parents have constrained budgets on children's educational investment. When the younger child in the household is old enough to attend school, the older child faces the potential of being crowded out of school and working to share the family burden of paying for tuition.

Hypothesis 4 is a continuous part of hypothesis 3 and only includes two-child families whose second child is female. Because of possible male-preferences, some households would intentionally save the education investment for males when the other child is a female. Which is to say, I would expect the girls who have an older sister will have longer completed schooling years than girls have an older brother when everything else is same.

The expectation of Hypothesis 5 is that females with a younger brother would have even lower educational attainment than females with an older brother. The purpose of hypothesis 5 is to test the birth order preference controlling for the gender composition of sibling.

The anecdotal evidence shows that first-born daughters are expected to contribute to the education of second-born sons. This leads to first-born daughters having a higher likelihood to attend vocational schools to immediately obtain jobs rather than delay job entry by attending high schools, the entryway to university training. **Hypothesis 6** tests if there is lower investment in both educational quantity and achieved level of education for first-born daughters than for second-born sons.

6. Results

Table 6 reports the educational outcomes for rural males and females between 16 to 20 years old according to their household structure. Generally, sons have higher completed schooling years than daughters, and this difference is significant at 10 percent level. For one-child families, daughters even complete longer schooling years than sons. Panel B and Panel C of **Table 6** show mean probability of finishing middle school and attending high school. Generally, a male between 16 to 20 years old has significantly higher probability of attending high school. Within one-child families, daughters and sons do not significantly differentiate in finishing middle school and attending high school.

[Insert Table 6 here]

Table 7 reports educational outcomes by different sex-composition for rural two-child families of 16 to 20 years old. From the completed schooling years' perspective, second-born sons have advantages in the completed schooling years. From the educational quality perspective, all the second-born children have higher probability of attending high school. Twins are excluded from the analysis because as a special group of two-child families, from the birth parity preference, there is no incentive for parents to prefer a particular child. For the same gender twins, parents cannot prefer either child based on the son-preference. Children's own endowment is the key factor for the educational achievements. For the different gender twins, parents could prefer the son, but the dataset design does not give their birth order information, but there is no way to distinguish birth order of the twins, so they are excluded from the analysis as well. There are only 7 pair of different gender twins between 16-20 years old, which is not enough to draw any conclusive inferences.

[Insert Table 7 here]

OLS estimates of household level determinants for rural female's education attainment are reported in **Tables 8,** the full model is listed as the **Table D in Appendix**. Column 1 shows that compared to daughters from one-child family, daughters from other families have 2.05 years lower completed schooling years on average; thus, females from one-child families have the advantage for schooling compared to the rest of rural females. And this advantage (Column 2) remains when compared to females from two-child families. Therefore, we can reject **Hypothesis 1.** Besides, the ethnic minority status does not significantly decrease the completed schooling years. Column 3 includes daughters and sons from one-child families, when all else being equal, being females do not have significant difference in the years of schooling. This finding corresponds with previous literature that females did not have significantly lower completed schooling years than males within one-child families. So **Hypothesis 2** can also be rejected.

[Insert Table 8 here]

Among the two-child families, the gender of the first child is a plausibly exogenous variable, as some anthropological evidence indicates that "selection for gender occurs

principally in births subsequent to a first-born daughter and that selective abortion prior to the birth of a first child is unusual" (Baochang et al., 2007; Banister, 2004). The evidence in CHIP 2007 rural sample is consistent with the theoretical evidence. Among two-child families, the sex ratio (male/female) for the first child at the year of 2007 is 0.95, while for the second-born child, the sex ratio is 1.67, which is highly imbalanced. Related to intra-household resources allocation, because of sibling rivalry and male preference, female siblings could be intentionally sacrificed when the household have both sons and daughters (Leight, 2014). Therefore, there is potential omitted variable bias if educational outcomes for boy embodies the preference from parents. And if the parental preference is correlated with other parental and familial characteristics then the educational attainment will be overestimated. However, the contamination is difficult to measure directly by any available socio-economic or demographic indicators, and is likely to lead to families with different sex composition in children that are different in both observed and unobserved outcomes, which is to say the male preference might correlate with the error term from the OLS regression estimation of the educational attainment. The sex ratio for the sample (**Table 2**) does not reveal a strong pattern of male preferences except for 4^{th} and 5^{th} parities. I test the endogeneity of the male preference, which in this paper refers to the proportion of having a son of two-child families. I choose year of 1984 as the instrument, which is the beginning year of the one-and-half child policy. The beginning year of the one-and-half child policy is a variable correlated with the likelihood the parent could legally have a son as the second-born children, which satisfy the instrument condition. Moreover, the year of 1984 is not correlated with the error term of the individual's completed schooling years, which satisfies the exclusion restriction as well. Hausman test F-score is 0.758 (p-value =0.3842), which indicates male preference is not endogenous with the error term for the 16 to 20 years old children from natal families when I use the one-and-half child policy as the instrument.

Column 4 and Column 5 in **Table 8** reports the OLS estimates of household level determinants for rural females from two-child families. The reference group for both estimations is still daughters from one-child families. These two estimations give the evidence of negative effect from siblings to completed schooling years. Compared to

females from one-child families, females who have a female sibling have 1.77 less completed schooling years. And the ethnic minority status decreases the completed schooling years by 3.37. As for females having a male sibling, the education completion is 1.70 years' lower than females from one-child families. According to the *Chinese Law of Compulsory Education*, the government must waive the tuition and fees for school-age children. Therefore, the financial burden to support daughters in finishing nine-year compulsory education only exists as an opportunity cost of forfeited labor, which before 15 years old is not high. In contrast, between the age between 16 to 20 years old, which are the right age for high school and college, students needed their families to pay for their tuition and fees. Having a sibling brings competition within the household for the education investments. Moreover, if the age difference is big enough between siblings, the older sister is likely to drop out of school to take care of the younger siblings (Levison and Moe, 1998).

The mixed evidence from male preference and birth parity preference prompts the comparison between first born and second born children in two-child families (**Table 9)**. The coefficients on the sibling pair in four columns are all significant. Comparing with the first child in the family, the second-born child in the family enjoys a significant advantage in the completed schooling years; and **Hypothesis 3** can be rejected. One unexpected finding is the second-born daughters' level of education. Second-born daughters have 1.42 more years of education completed than their older brothers. The higher birth order advantage overwhelming the son-preference. Second-born daughter with have 1.37 more years of education completed than their older sisters. This finding implies that within two-child families, besides son preference, birth order is an important factor correlated with the completed schooling years. And son preference is not always over the birth parity preference within two-child families. The above evidences **fail to reject Hypothesis 4**.

Younger brother jeopardizes the older sister's schooling years by 2.15 years (**Table 9)**. Putting this evidence under the context of **Hypothesis 5**, the younger brothers take an advantage in educational attainment not just because of male preference, but also the later birth parity. Accordingly, **Hypothesis 5** can be rejected. Compared to other sex

compositions two-child families, the younger brother detracts from the older sister's years of schooling most. So when the son preference combines with the later birth order advantage, within the household there is severe disparities on investment to children's education.

[Insert Table 9 here]

Table 10 presents the odds ratio on probability of attending high school for children and young adults from rural two-child families between 16-20 years old. The full models of maximum likelihood estimates are displayed in the **Appendix Table F**. After controlling for household level factors, there still is a distinct and statistically significant penalty for being a first born daughter relative to a younger brother on the probability of attending high school, which confirms the findings of **Table 9**. The odds ratio of second born sons to attend high school is 2.46 times than first born daughters, it is significant at 1% significance level. There are educational level penalty translates to first born son from son-daughter families as well. Second born daughters are 1.72 times more likely to attend high school than their first born brothers, but it only significant at 10% level. There is no difference between same gender sibling pairs on the probability of attending academic high school. Except for the daughter-son sibling pairs, the tightly spaced sibling pairs does not have significantly effect on the educational level. And the positive coefficient is contracted to the theory that the tightly spaced sibling pairs compete for the parental investment. This seemly unexpected estimate may relate to one education subsidy policy for rural China, which starts to provide tuition free high school education from 2005.

An interesting phenomenon is raised for the son-daughter families. According to either the one-child or one-and-half child policy, when parents already have a son, they are not legally to have another child, except for the exempted cases. CHIP 2007 captured more than 300 first-born son and second-born daughters, which is unlikely they are all the exempted families from one-child or one-and-half child policy. There is one possibility that those parents want to have more children than other parents, and the son preference is not their only reason to have more children since their daughters have longer years of schooling and high probability of attending high school.

[Insert Table 10 here]

7. Discussion and Conclusions

Using the weighted CHIP 2007 rural data and focusing on household level determinants, estimates confirms many previous research results that females from one-child families have advantage in the completed schooling years comparing to females from non-one-child families. My findings regarding the negative effect from siblings to completed schooling years also corresponds to previous research that females from one-child families have an advantage in educational attainment over females who have a sibling. In 2015, the State Council started the two-child policy in China. If there are educational subsidies for families with two children, these children's attainments could be promoted. Considering the relatively low resources possessed by households in rural China, financial subsidies for education expenditure might be effective to promote more completed schooling years. Starting in 2005 and expanded in 2011, the Chinese central government has provided free compulsory and high school education to all children in rural areas, which includes free tuition, free textbooks, and a living expenditure subsidy for boarding students. This policy is known as the "Two Exemptions and One Subsidy (TEOS)" policy (China National Center for Student Financial Aid). In the future, rural Chinese females may expect a lessening negative effect of having a sibling on completed schooling years. Policy interventions that aim to increase completed schooling years for children in rural China are typically targeted at the household level, where due to the son-preference, households may allocate these subsidies or benefits more to males than females. Accordingly, in the rural Chinese context, some specific policy interventions are needed to compensate for gender imbalances.

Most of the existing literature addresses gender disparities in educational attainments in rural areas. However, if the one considers birth order within families, more stories can be told about rural areas. Females do not always complete fewer years of schooling; second-born daughters complete more years of schooling compared to their older brothers. Moreover, among the sibling pairs, second-born children have advantage in the

completed schooling years, even though these evidences contrary to Arleen Leibowitz (1974) theory that lower birth order (older children) can be expected to receive more inputs from family. This corresponds to the resource dilution model of parental investment in children.

One more important finding is the relationship between educational level and gender and birth order within sibling pairs. Families invest more in the level of education of second born sons than in first born daughters, which provides empirical evidence for the anecdotal claim that second born sons are more likely to go to high school than first born daughters. The significant educational level penalty for being a first-born daughter relative to a younger brother in the household exposes and exacerbates the gender inequality of the educational resource allocation within household. Females facing a disadvantage in the educational achievement in rural China is a well-studied topic. Without including information about the family structure, the general conclusion would not identify the penalty of first-born daughters. Understanding the way that gender interacts with birth order to create an educational disparity will provide a better basis for targeted policy making.

Due to the dataset design, this paper cannot capture the children who left their natal household, the results are only for children who were still living with their natal families when the data was collected. As per the Hukou registration system[13] (Nationality Law of the People's Republic of China), this research paper could only contain residents or inhabitants registered under Hukou. The population who do not have Hukou and the population who live in places other than their registered Hukou residences are not included in the analysis. However, based on the Hukou registration system, if a child is born outside the one child or one-and-half child policy, the child cannot get the hukou. The unregistered children are called “Heihaizi”[14]. As they grow up, many of the unregistered children will eventually get stable Hukou status, but this delay of registration is more like to happen to girls. Some researchers point out that unregistered

[13] A **Hukou** is a record in the system of household registration required by law in China.
[14] Some research adopts the term “black children” as the translation of the unregistered or undocumented children. Because the term “black children” has culturally loaded connotations, I directly use the Mandarin word.

children and the delay registration is related to the parents' male preferences (Zhou, 2005; Banister, 1987). Without the unregistered and delay registration information, the estimations captured in this paper are likely to underestimate the females' disadvantage in the completed schooling years and probability of getting academic high school education.

In 2015, Chinese government abolished the one child policy, and each couple is allowed to have two children[15] (two child policy). The implementation of two child policy would satisfy the son preference for some parents when the first born child is a female, but reduce the number of unregistered children. It will be interesting to see the impact from this policy to the educational attainment for the Chinese rural females, especially on the education quality.

[15] http://news.china.com/zh_cn/focus/wzqh/11174588/20151029/20655319_all.html

References

Arleen Leibowitz, “Home Investments in Children” in Economics of the Family: Marriage, Children, and Human Capital, ed. Theodore W. Schultz. (Chicago: University of Chicago Press, 1974),432-456.

C.Y. Cyrus Chu, Yu Xie, Ruoh-rong Yu, “Effects of Sibship Structure Revisited: Evidence from Intrafamily Resource Transfer in Taiwan,” Sociology of Education Vol.80 (2007,): 91-113.

Chengrong Duan, Lidan Lv, Jing Guo, Zongping Wang, “Survival and Development of Left-behind Children in Rural China: Based on the Analysis of Sixth Census Data,” *Population Journal* Vol. 3 (2013) :37-49.

Chengfang Liu, Linxiu Zhang, Renfu Luo, Scott Rozelle,Brian Sharbono,Yaojiang Shi, “Development challenges, tuition barriers, and high school education in China,” *Asia Pacific Journal of Education*, 29 (2009): 4, 503 — 520.

Chuliang Luo, Shi Li, Terry Sicular, Quheng, Deng and Ximing Yue, “The 2007 Household Surveys: Sampling Method and Data Description.” In: *Rising Inequality in China: Challenges to a Harmonious Society*, ed. Li Shi et al. (New York: Cambridge Press 2012).

Chuliang Luo, Shi Li, Terry Sicular, Quheng, Deng and Ximing Yue, “The 2002 and 2007 CHIP Surveys: Sampling, Weights, and Combining the Urban, Rural and Migrant samples.” In: *Rising Inequality in China: Challenges to a Harmonious Society*, ed. Li Shi et al. (New York: Cambridge Press 2012).

Deborah Levison, Karine Moe “Household Work as a Deterrent to Schooling: An Analysis of Adolescent Girls in Peru,” *Journal of Developing Areas* 32 (1998): 339-356.

Elisabeth Croll, From Heaven to Earth: Images and Experiences of Development in

China (New York: Routledge,1994).

Emily Hannum, “Political Change and the Urban-Rural Gap in Basic Education in China, 1949-1990,” *Comparative Education Review* Vol. 43 (1999): 193-211.

Emily Hannum, “Market Transition, Education Disparities, and Families Strategies in Rural China: New Evidence on Gender Stratification and Development,” *Demography*, 42 (2005): 275-299.

Emily Hannum Peggy Kong, Yuping Zhang, “Family sources of educational gender inequality in rural China: A critical assessment,” *International Journal of Educational Development 29* (2009): 474–486.

Emily Hannum, Yu Xie, “Trends in educational gender inequality in China: 1949-1985,” *Research in Social Stratification and Mobility*, Vol. 13 (1994): 73-98.

Emily Hannum, Yu Xie, “Ethnic stratification in northwest China: occupation differences between Han Chinese and national minorities in Xinjiang 1982-1990,” *Demography*, Vol. 35-Number 3 (1998): 323-333.

Emily Honig, Gail Hershatter, *Personal Voices: Chinese Women in the 1980’s* (Redwood City: Stanford University Press, 1988).

Gary S. Becker, A Treatise of Family: Enlarged Edition (Cambridge: Harvard University Press, 1991).

Gu Baochang, Wang Feng, Guo Zhigang, and Zhang Erli, “China’s local and national fertility policies at the end of the twentieth century,” *Population and Development Review*, 2007, 33, 129–147.

Jessica Leight, "Sibling Rivalry: Ability and Intra household Allocation in Gansu Province, China,” *Revise and Resubmit, Economic Development and Cultural Change*.

Judith Banister, “Shortage of Girls in China Today,” *Journal of Population Research* 21 (2004): 19–45.

Judith Banister and Kenneth Hill, “Mortality in China 1964-2000”, *Population Studies*, March 2004, Vol. 58:55-75.

Judith Blake, Family Size and Achievement (Berkeley: University of California Press, 1989).

Junhong Chu,” Prenatal Sex Determination and Sex-Selective Abortion in Rural Central China,” *Population and Development Review (2011) Vol.27, No.2:259-281*

Junxia Zhang, Xiaopeng Pang, Linxiu Zhang, Alexis Medina, Scott Rozelle, “Gender inequality in Education in China: A Meta-regression analysis,” working paper 239, May 2012, “reapchina.org/reap.stanford.edu”.

Lina Song, Simon Appleton, John Knight, “Why do girls in total China have lower school enrollment.” *World Development* (2006) Vol.23, No.9: 1639-1652.

Margery Wolf, *Revolution Postponed: Women in Contemporary China* (Redwood City: Stanford University Press, 1985).

Ming Tsui and Lynne Rich, “The only child and educational opportunity for girls in urban China.” *Gender & Society* 16 (2002): 74-92.

Philip H. Brown and Albert Park, “Education and poverty in rural China,” *Economics of Education Review*, 2002 (21), 523-541.

Phyllis Andors, The unfinished liberation of Chinese women, 1949-1980 (Bloomington: Indiana University Press,1990).

Samuel L. Myers Jr., Gao Xiaoyan, Britt Cecconi Cruz “Ethnic Minorities, Race, and Inequality in China: A New Perspective on Racial Dynamics,” *The Review of Black Political Economy,* September 2013, Volume 40, Issue 3, pp 231-244.

Sten Johansson and Ola Nygren “The Missing Girls of China: A New Demographic Account,” *Population and Development Review*, February 1991, Vol. 17, 35-51.

Susan Greenhalgh “Planned births, unplanned persons: ‘Population in the making of Chinese modernity”, *American Ethnologists,* 2003,30(2): 1-20.

William L. Parish and Robert Willis, “Daughters, education and family budgets: Taiwan Experiences,” *The Journal of Human Resources,* Autumn 1993, 28 (24), 863–898.

Xue Lan Rong, Tianjian Shi, “Inequality in Chinese Education,” *Journal of Contemporary China* (2001), 10(26), 107-124.

Tables

Table 1: Number of children in natal households for CHIP rural sample (2007)

Number of children in natal HH	Number of HH	Percent (%)
0	121	1.61
1	3,463	46.1
2	3,105	41.33
3	651	8.67
4	133	1.77
5	28	0.37
6	10	0.13
7	1	0.01
Total	7,512	100

Note: The number of children in natal household in the above table is unweighted count from CHIP 2007 rural sample, the children who left the household are not captured in the table. All children included are biological siblings.

Table 2: Sex ratio by birth parity at year of 2007 for natal children in CHIP rural sample

Whole Sample	1st born	2nd born	3rd born	4th born	5th born	6th born	7th born	8th born	9th born	Total
Male	6,811	4,916	2,394	1,122	562	238	96	45	15	16,199
Female	6,067	4,401	2,417	1,082	478	221	97	32	11	14,806
Ratio(M/F)	112.26	111.70	99.05	103.70	117.57	107.69	98.97	140.63	136.36	109.41
Total	12,878	9,317	4,811	2,204	1,040	459	193	77	26	31,005

16--20 years old	1st born	2nd born	3rd born	4th born	5th born	6th born	Total
Male	795	595	226	87	22	7	1,732
Female	811	568	226	65	14	9	1,693
Ratio(M/F)	98.03	104.75	100.00	133.85	157.14	77.78	102.30
Total	1,606	1,163	452	152	36	16	3,425

Note: The displayed sex ratio is not the sex ratio at birth, but the sex ratio at the year of 2007.

Table 3: Definition of independent variables controlled in the estimates

Variable	Explanation
Age	Interviewee's age.
Father's education	Paternal completed schooling years. Theoretical, positively correlated with children's years of schooling.
Mother's education	Maternal completed schooling years. Theoretical, positively correlated with children's years of schooling.
	An important factor related to children's education in rural areas or developing countries
Father's age	Paternal age. The older father could influenced more by the male preferences in rural China.
Mother's age	Maternal age. The older mother could be influenced more by the male preferences in rural China.
Household income quartile	The highest household income could support children to finish longer years of schooling.
Ethnic minority	Belong to one of the fifty-five ethnic minorities in China. Ethnic minorities face more disadvantage than Han(majority) to obtain longer years of schooling.
Mother's migration status	Maternal migration status in 2007. Many rural Chinese migrate to urban areas for working purposes. On one hand, they can increase the household income by working in the urban areas, but empirical evidences also show with the maternal living with, children have probity of dropping from school .
Father's migration status	Paternal migration status in 2007. Many rural Chinese migrate to urban areas for working purposes. Same as the maternal migration.
Non one-child daughter	If=1, daughter from 1+ child families. The comparison group is daughter from one-child families.
Two-child daughter	If=1, daughter from two-child families. The comparison group is daughter from one-child families.
One-child son	If=1, son from one-child families. The comparison group is daughter from one-child families.
Having a Female sibling	If =1, daughter of two-child families and having one female sibling. The comparison group is daughter from one-child families.
Having a Male sibling	If =1, daughter of two-child families and having one male sibling. The comparison group is daughter from one-child families.
2nd son in two-son family	If =1, second son from two-son families. The comparison group is first son from two-son families.
2nd daughter in two-daughter family	If =1, second daughter from two-daughter families. The comparison group is first daughter from two-daughter families.
2nd daughter in son-daughter family	If =1, second daughter from son-daughter families. The comparison group is first son from son-daughter families.
2nd son in daughter-son family	If =1, second son from son-daughter families. The comparison group is first son from son-daughter families.

Table 4: Selected household level characteristics for rural females between 16--20 years old

	Females from One-child Families		
	Mean	**SD**	**Obs**
Children's education (year)	9.88	2.32	170
Father's education (year)	7.68	2.35	161
Mother's education (year)	7.84	2.70	158
Father's age	43.85	5.14	173
Mother's age	42.57	5.12	170
Household Incomes (quartile)	2.43	1.09	188
Minorities	0.02	0.13	188
Mother's migration status (for working purpose)	0.09	0.28	188
Father's migration status (for working purpose)	0.20	0.40	188
School enrollment status	0.65	0.48	188
	Females from Two-child Families		
	Mean	**SD**	**Obs**
Children's education (year)	8.92	2.36	822
Father's education (year)	7.71	2.37	796
Mother's education (year)	7.76	2.26	796
Father's age	44.80	5.21	849
Mother's age	43.23	4.68	850
Household Incomes (quartile)	2.32	1.07	876
Minorities	0.02	0.13	876
Mother's migration status (for working purpose)	0.06	0.23	876
Father's migration status (for working purpose)	0.20	0.40	876
School enrollment status	0.67	0.47	876
Having a male sibling	0.66	0.47	876
	Females from Families with more than one child		
	Mean	**SD**	**Obs**
Children's education (year)	8.99	2.31	1379
Father's education (year)	7.81	2.31	1315
Mother's education (year)	7.89	2.25	1306
Father's age	45.06	5.28	1392
Mother's age	43.57	4.71	1385
Quartile of Household Incomes	2.40	1.10	1476
Minorities	0.01	0.11	1476
Mother's migration status (for working purpose)	0.06	0.23	1476
Father's migration status (for working purpose)	0.19	0.39	1476
School enrollment status	0.71	0.46	1476

Note: Person Weights applied.

Table 5: Means difference test on selected household level characteristics between rural females from one-child and two-child families, ages 16-20

	Difference	SE	T-score
Father's education (year)	0.029	0.24	0.12
Mother's education (year)	-0.080	0.29	-0.27
Father's age	0.949	0.48	1.98**
Mother's age	0.665	0.49	1.37
Household Incomes (quartile)	-0.103	0.10	-1.02
Minorities	-0.001	0.01	-0.1
Mother's migration status (for working purpose)	-0.033	0.03	-1.26
Father's migration status (for working purpose)	0.007	0.04	0.18
School enrollment status	0.025	0.05	0.54

Note: 1. Females from one-child families is the reference group.

2. Person Weights applied.

* p<0.10, ** p<0.05, *** p<0.01

Table 6: Educational outcomes by sex for CHIP 2007 rural sample and one-child families, ages 16-20

	Female	Male	Difference	T-score
	(N)	(N)		
Panel A : Completed schooling year				
Whole sample	9.10	9.29	-0.19	-1.85*
	(1549)	(1607)		
One-child Families	9.88	9.75	0.12	0.46
	(170)	(279)		
Panel B: Probability of finishing 9th grade				
Whole sample	0.79	0.81	-0.02	1.50
	(1549)	(1607)		
One-child Families	0.88	0.85	0.03	0.98
	(170)	(279)		
Panel C: Probability of attending high school				
Whole sample	0.33	0.38	-0.06	3.30***
	(1549)	(1607)		
One-child Families	0.43	0.48	-0.05	-1.12
	(170)	(279)		

Note: 1. Females is the reference group.

2. Weights applied to the calculation.

Table 7: Educational outcomes by different sex-composition for rural two-child families, ages 16-20

	First-born (N)	Second-born (N)	Difference	T-score
Completed schooling year:				
Two-boy Families	8.75	9.50	-0.74	-2.76***
	(140)	(175)		
Two-girl Families	8.24	8.77	-0.52	-1.09
	(108)	(68)		
Boy-girl Families	8.62	9.17	-0.55	-1.85*
	(156)	(215)		
Girl-boy Families	8.80	9.73	-0.93	-3.99***
	(308)	(221)		
Probability of finishing 9th grade:				
Two-boy Families	0.77	0.83	-0.06	-1.40
	(140)	(175)		
Two-girl Families	0.74	0.78	-0.04	-0.58
	(108)	(68)		
Boy-girl Families	0.78	0.80	-0.02	-0.42
	(156)	(215)		
Girl-boy Families	0.78	0.86	-0.08	-2.4**
	(308)	(221)		
Probability of attending high school:				
Two-boy Families	0.25	0.37	-0.12	-2.31**
	(140)	(175)		
Two-girl Families	0.22	0.38	-0.16	-2.32**
	(108)	(68)		
Boy-girl Families	0.27	0.37	-0.10	-1.99**
	(156)	(215)		
Girl-boy Families	0.27	0.49	-0.23	-5.39***
	(308)	(221)		

Note: 1. The first born children is the reference group.

2. Personal Weights applied.

* p<0.10, ** p<0.05, *** p<0.01

Table 8: OLS estimates on educational attainment for Chinese rural females, CHIP(2007), ages 16-20

	1-child F vs. other F (1)	1-child F vs.2-child F (2)	1-child F vs. M (3)	1-child F vs. F with F sibling (4)	1-child F vs. F with M sibling (5)
	β	β	β	β	β
	(SE)	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score	T-score
Dependent variable: Educational completion(year)					
Independent variables:					
(=1, daughter from 1+ child families)	-2.0512***				
	(0.2571)				
	-7.9789				
(=1, daughter of two-child families)		-1.9423***			
		(0.2730)			
		-7.1148			
(=1, son of one-child families)			-0.2504		
			(0.3128)		
			-0.8004		
(=1, daughter of two-child families and having one female sibling)				-1.7691***	
				(0.3217)	
				-5.4986	
(=1, daughter of two-child families and having one male sibling)					-1.6959***
					(0.2685)
					-6.3159

Note: 1. Females from one-child families is the reference group.

Colum (1) compares females from one-child families to females from other families.

Colum (2) compares females from one-child families to females from two-child families.

Colum (3) compares females from one-child families to males from one-child families.

Colum (4) compares females from one-child families to females having a female sibling.

Colum (5) compares females from one-child families to females having a male sibling.

2. Person Weights applied.

3. Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

Table 9: OLS estimates on educational attainment for Chinese rural females from two-child families, CHIP(2007), ages 16-20

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	β	β	β	β
	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score
Dependent variable: Educational completion(year)				
Independent variables:				
2nd son in two-son family (=1,2nd son)	1.3746***			
	(0.3635)			
	3.7814			
2nd daughter in two-daughter family (=1,2nd daughter)		1.6478***		
		(0.5501)		
		2.9955		
2nd daughter in son-daughter family (=1,2nd daughter)			1.4213***	
			(0.3861)	
			3.6811	
2nd son in daughter-son family (=1,2nd son)				2.1522***
				(0.3374)
				6.3780

Note: 1. First born children is the reference group.

Colum (1) compares first son to second son who are both from two-son families.

Colum (2) compares first daughter to second daughter who are both from two-daughter families.

Colum (3) compares first son to second daughter who are both from son-daughter families.

Colum (4) compares first daughter to second son who are both from daughter-son families

2. Person Weights applied.

3.Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

Table 10: Odds ratio on probability of attending high School for Chinese rural females from two-child families , CHIP(2007), ages 16-20

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	Odds Ratio	**Odds Ratio**	**Odds Ratio**	**Odds Ratio**
Dependent variable: Educational completion(year)				
Independent variables:				
2nd son in two-son family (=1,2nd son)	1.34			
2nd daughter in two-daughter family (=1,2nd daughter)		0.85		
2nd daughter in son-daughter family (=1,2nd daughter)			1.72*	
2nd son in daughter-son family (=1,2nd son)				2.46***

Note: 1. First born children is the reference group.

Colum (1) compares first son to second son who are both from two-son families.

Colum (2) compares first daughter to second daughter who are both from two-daughter families.

Colum (3) compares first son to second daughter who are both from son-daughter families.

Colum (4) compares first daughter to second son who are both from daughter-son families

2. Person Weights applied.

3.Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

Appendix

Table A: Tabulation on number of children in the CHIP 2007 rural sample

Number of Children living in natal households	Number of children ever lived in the natal households									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**
0	82	38	6	5	0	0	3	0	0	134
1	2,508	1,193	348	127	29	13	2	0	0	4,220
2	1,251	4,414	541	162	50	21	2	0	1	6,442
3	119	244	1,503	118	33	11	0	3	0	2,031
4	13	14	23	452	24	16	0	0	0	542
5	3	5	5	5	115	10	0	0	0	143
6	0	6	0	0	0	48	7	0	0	60
7	0	0	0	0	0	0	7	0	0	7
Total	3976	5914	2426	869	251	119	21	3	1	13579

Note: According to the dataset design, the daughters and sons from either one-child or two-child families were living in their natal household when the data was collected. This design captures at least 63.08% of the number of daughters and sons who ever lived in the natal household.

Table B: Tabulation on number of sibling pairs in the CHIP 2007 rural sample

Number of sibling pairs in natal households	Whole sample	Both the pairs are between 16-20 years old
Two-boy Families	635	70
Two-girl Families	352	25
Boy-girl Families	583	84
Girl-boy Families	837	109

Table C: Tabulation on age difference between sibling pairs between 16--20 years old in the CHIP 2007 rural sample

Age difference between sibling pair	Two-boy Families	Two-girl Families	Boy-girl Families	Girl-boy Families
1	54	7	40	69
2	124	55	137	189
3	44	27	80	85
4	28	12	38	52
5	25	16	24	50
6	14	25	20	44
7	10	16	15	24
8	5	9	4	21
9	5	5	10	13
10	9	5	9	9
11	/	1	1	4
12	3	1	2	4
13	2	1	3	/
14	2	3	3	1
15	3	2	1	3
16	2	/	/	1
Total	330	185	162	569

Note: The sibling pair denotes to all children between 16—20 years old even though their sibling is out of the age range.

Table D: OLS estimates on educational attainment for Chinese rural females, CHIP(2007), ages 16-20

	1-child F vs. other F (1)	1-child F vs.2-child F (2)	1-child F vs. M (3)	1-child F vs. F with F sibling (4)	1-child F vs. F with M sibling (5)
	β	β	β	β	β
	(SE)	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score	T-score
Dependent variable: Educational completion(year)					
Independent variables:					
Age	0.2931***	0.2565***	0.2314	0.1826	0.2983***
	(0.0680)	(0.0833)	(0.1456)	(0.1117)	(0.0778)
	4.3119	3.0798	1.5886	1.6346	3.8333
Father's education	0.1085**	0.1339***	0.0148	0.1345*	0.1291***
	(0.0430)	(0.0503)	(0.0688)	(0.0703)	(0.0455)
	2.5236	2.6638	0.2154	1.9131	2.8384
Mother's education	0.0685	0.1090**	-0.0235	0.0733	0.0979**
	(0.0440)	(0.0532)	(0.0791)	(0.0751)	(0.0471)
	1.5584	2.0470	-0.2967	0.9757	2.0796
Father's age	-0.0100	-0.0026	-0.0437	0.0050	-0.0315
	(0.0369)	(0.0484)	(0.0675)	(0.0840)	(0.0423)
	-0.2717	-0.0537	-0.6465	0.0593	-0.7447
Mother's age	0.0634	0.0424	-0.0091	0.0581	0.0830**
	(0.0398)	(0.0521)	(0.0595)	(0.0895)	(0.0418)
	1.5946	0.8141	-0.1522	0.6493	1.9846
4th household income (highest Household income quartile)	0.3155	0.2184	1.1949***	0.4716	0.2817
	(0.2883)	(0.3442)	(0.4575)	(0.4916)	(0.3241)
	1.0944	0.6343	2.6119	0.9595	0.8692
3rd household income	0.2802	0.1964	0.7412*	0.3740	0.0669
	(0.2547)	(0.3085)	(0.3944)	(0.4401)	(0.2901)
	1.1002	0.6366	1.8794	0.8498	0.2305
2nd household income	0.0755	0.1104	0.7864*	0.2874	-0.0454
	(0.2601)	(0.3075)	(0.4759)	(0.4189)	(0.2847)
	0.2903	0.3591	1.6526	0.6860	-0.1595
Ethnic minority	-1.0591	-1.5077	-0.1190	-3.3749***	0.6179
(=1, minority)	(1.2653)	(1.5167)	(0.8768)	(0.8606)	(1.3065)
	-0.8370	-0.9941	-0.1358	-3.9214	0.4729
Mother's migration status	0.3632	0.6071	-0.4755	0.1475	0.4643
(for working purpose)	(0.3790)	(0.4601)	(0.6217)	(0.5649)	(0.3938)
	0.9583	1.3194	-0.7649	0.2611	1.1791
Father's migration status	-0.0915	-0.1135	-0.6569	0.4163	-0.1462
(for working purpose)	(0.2434)	(0.2865)	(0.5300)	(0.3681)	(0.2699)
	-0.3760	-0.3961	-1.2396	1.1311	-0.5417
(=1, daughter from 1+ child families)	-2.0512***				
	(0.2571)				
	-7.9789				

	1-child F vs. other F (1)	1-child F vs.2-child F (2)	1-child F vs. M (3)	1-child F vs. F with F sibling (4)	1-child F vs. F with M sibling (5)
	β	β	β	β	β
	(SE)	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score	T-score
(=1, daughter of two-child families)		-1.9423***			
		(0.2730)			
		-7.1148			
(=1, son of one-child families)			-0.2504		
			(0.3128)		
			-0.8004		
(=1, daughter of two-child families and having one female sibling)				-1.7691***	
				(0.3217)	
				-5.4986	
(=1, daughter of two-child families and having one male sibling)					-1.6959***
					(0.2685)
					-6.3159
Constant	0.8738	1.6026	7.6571***	2.0013	0.5761
	(1.2932)	(1.5092)	(2.1041)	(2.0746)	(1.4940)
	0.6757	1.0619	3.6390	0.9647	0.3856
Observations	1,380	892	371	409	971
Prob > F	0.000	0.000	0.016	0.000	0.000
R-squared	0.0996	0.119	0.0724	0.176	0.111

Note: 1. Females from one-child families is the reference group.

Colum (1) compares females from one-child families to females from other families.

Colum (2) compares females from one-child families to females from two-child families.

Colum (3) compares females from one-child families to males from one-child families.

Colum (4) compares females from one-child families to females having a female sibling.

Colum (5) compares females from one-child families to females having a male sibling.

2. Person Weights applied.

3.Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

Table E: OLS estimates on educational attainment for Chinese rural females from two-child families, CHIP(2007), ages 16-20

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	β	β	β	β
	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score
Dependent variable: Educational completion (year)				
Independent variables:				
Age	0.0335	0.3224**	0.4159***	0.5403***
	(0.1550)	(0.1612)	(0.1388)	(0.1085)
	0.2162	1.9998	2.9963	4.9799
Father's education	0.1684**	0.2831**	0.0348	0.1809***
	(0.0761)	(0.1205)	(0.0676)	(0.0602)
	2.2141	2.3501	0.5143	3.0065
Mother's education	0.1668**	0.1836*	0.1580*	0.0789
	(0.0836)	(0.1087)	(0.0859)	(0.0626)
	1.9950	1.6894	1.8394	1.2618
Father's age	-0.0407	-0.0622	-0.1356*	0.0102
	(0.0790)	(0.1047)	(0.0699)	(0.0539)
	-0.5151	-0.5942	-1.9390	0.1894
Mother's age	0.0256	0.1604	0.0679	-0.0014
	(0.0744)	(0.1141)	(0.0734)	(0.0542)
	0.3435	1.4064	0.9249	-0.0259
4th household income (highest Household income quartile)	1.4679***	1.2586*	0.7431	-0.4433
	(0.4396)	(0.7168)	(0.4819)	(0.4631)
	3.3388	1.7559	1.5419	-0.9573
3rd household income	0.2286	0.9622	-0.4645	0.1724
	(0.4798)	(0.5820)	(0.5575)	(0.3683)
	0.4764	1.6532	-0.8332	0.4681
2nd household income	0.7626	0.7389	-0.1986	-0.4310
	(0.4646)	(0.6467)	(0.4529)	(0.3890)
	1.6414	1.1426	-0.4386	-1.1080
Ethnic minority (=1, minority)	-5.5573***	-2.6950***	-0.3142	0.5455
	(1.1569)	(0.6696)	(3.7170)	(0.8235)
	-4.8036	-4.0248	-0.0845	0.6624
Mother's migration status (for working purpose)	0.5249	0.5216	0.9660*	0.1555
	(0.5847)	(0.7271)	(0.5585)	(0.6706)
	0.8977	0.7174	1.7296	0.2319
Father's migration status	0.1737	2.0493***	-0.0618	-0.0900
	(0.3680)	(0.5826)	(0.4121)	(0.4012)
	0.4720	3.5178	-0.1500	-0.2243

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	β	**β**	**β**	**β**
	(SE)	**(SE)**	**(SE)**	**(SE)**
	T-score	**T-score**	**T-score**	**T-score**
2nd son in two-son family (=1,2nd son)	1.3746***			
	(0.3635)			
	3.7814			
2nd daughter in two-daughter family (=1,2nd daughter)		1.6478***		
		(0.5501)		
		2.9955		
2nd daughter in son-daughter family (=1,2nd daughter)			1.4213***	
			(0.3861)	
			3.6811	
2nd son in daughter-son family (=1,2nd son)				2.1522***
				(0.3374)
				6.3780
Constant	3.4951	-9.7305***	0.2230	-6.7883***
	(2.7488)	(3.3481)	(2.8057)	(2.0192)
	1.2715	-2.9063	0.0795	-3.3619
Observations	283	159	323	479
R-square	0.263	0.373	0.131	0.235
Prob > F	0.000	0.000	0.000	0.000

Note: 1. First born children is the reference group.

2. Person Weights applied.

3.Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01

Table F: Logit estimates on probability of attending high school for Chinese females from two-child families , CHIP(2007), ages 16-20

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	β	β	β	β
	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score
Dependent variable: Probability of attending high school				
Independent variables:				
Age	0.0366	-0.0205	0.2678**	0.2167**
	(0.1114)	(0.1615)	(0.1101)	(0.0872)
	0.3282	-0.1269	2.4332	2.4842
Father's education	-0.0142	0.1401	0.0513	0.0895*
	(0.0604)	(0.0967)	(0.0558)	(0.0464)
	-0.2348	1.4484	0.9201	1.9280
Mother's education	0.0902	0.0793	0.1566***	0.1121**
	(0.0624)	(0.0859)	(0.0605)	(0.0490)
	1.4462	0.9233	2.5898	2.2856
Father's age	-0.0031	0.1520**	-0.0428	0.0169
	(0.0556)	(0.0709)	(0.0456)	(0.0390)
	-0.0551	2.1419	-0.9379	0.4332
Mother's age	0.0296	-0.0027	0.0291	-0.0020
	(0.0588)	(0.0763)	(0.0504)	(0.0430)
	0.5038	-0.0349	0.5767	-0.0462
4th household income (highest Household income quartile)	0.7983**	2.4283***	0.1957	0.4566
	(0.3953)	(0.8820)	(0.3588)	(0.3062)
	2.0194	2.7532	0.5455	1.4914
3rd household income	-0.1187	2.0540**	-0.6457*	0.2038
	(0.4086)	(0.8625)	(0.3828)	(0.2826)
	-0.2905	2.3814	-1.6869	0.7213
2nd household income	0.1123	2.3946***	-0.6638*	0.3725
	(0.3711)	(0.8512)	(0.3502)	(0.2924)
	0.3027	2.8132	-1.8957	1.2737
Ethnic minority (=1, minority)	/	/	-0.1670	0.4323
			(1.2965)	(0.9105)
			-0.1288	0.4748
Disability, not affecting normal life (=1, with disability not affecting normal life)	1.1528	0.1547	0.5286	/
	(1.4834)	(0.9301)	(1.2576)	
	0.7772	0.1663	0.4204	

	1st vs. 2nd son in two-son (1)	1st vs. 2nd daughter in two-daughter (2)	1st son vs. 2nd daughter in son-daughter (3)	1st daughter vs. 2nd son in daughter-son (4)
	β	β	β	β
	(SE)	(SE)	(SE)	(SE)
	T-score	T-score	T-score	T-score
Disability, affecting normal life	/	/	/	0.1785
(=1, with disability affecting normal life)				(1.4402)
				0.1240
Tightly spaced sibling pairs	0.2238	0.1466	0.4653	0.4916**
(=1,age difference between sibling is less than 5)	(0.3330)	(0.4628)	(0.3484)	(0.2433)
	0.6720	0.3168	1.3354	2.0208
2nd son in two-son family	0.2921			
(=1,2nd son)	(0.3087)			
	0.9463			
2nd daughter in two-daughter family		-0.1599		
(=1,2nd daughter)		(0.5190)		
		-0.3081		
2nd daughter in son-daughter family			0.5419*	
(=1,2nd daughter)			(0.3072)	
			1.7641	
2nd son in daughter-son family				0.9017***
(=1,2nd son)				(0.2428)
				3.7145
Constant	-4.0345*	-10.9520***	-7.5640***	-8.5624***
2nd son in two-son family	(2.1138)	(3.3175)	(2.0917)	(1.5958)
(=1,2nd son)	-1.9086	-3.3013	-3.6162	-5.3655
	0.2238	0.1466	0.4653	0.4916**
2nd daughter in two-daughter family	(0.3330)	(0.4628)	(0.3484)	(0.2433)
(=1,2nd daughter)	0.6720	0.3168	1.3354	2.0208
Observations	279	155	322	474
R-square	13.42	27.71	38.08	67.26
Prob > Chi2	0.0384	0.150	0.0940	0.107

Note: 1. First born children is the reference group.

2. Person Weights applied.

3.Robust Standard errors in parentheses.

* p<0.10, ** p<0.05, *** p<0.01