

The Impact of Parental Son Preference on Children's Educational Attainments and  
Its Mechanism

Abstract: This paper studies the impact of parents' preference for sons over daughters on educational attainments of children in China by directly measuring the magnitude of son preference. Based on the data from the Gansu Survey of Children and Families, conducted in year 2000 in Gansu province, China, our study suggests that higher parental preference for sons leads to lower education attainments for children. One standard deviation increase in the level of son preference decreases a child's Chinese and math grades on average by 1.05 and 1.01 points, respectively, out of 100 points, and increases the probability of grade retention by 9.99%. Moreover, we extend our analysis to study the mechanism of the impact of son preference on children's educational attainments. The results show that parental son preference affects children's educational attainments through four channels: child-rearing methods, the mother's expectations for the child's schooling, the child's educational aspirations and the child's self-confidence. After controlling for the above-mentioned channels, son preference no longer has a significant correlation with educational attainments. That is, these four channels constitute the mechanism of son preference on children's educational attainments.

Traditional preference for sons over daughters is a particularly important issue in developing countries (Andersson et al. 2006; Dahl and Moretti 2008; Das Gupta 2007; Kingdon 2005). Son preference leads to a series of social problems, with one of the greatest of these being sex ratio imbalance, which causes increase in sexual crime and decreased marital happiness due to marriage-age imbalance (Gu and Roy 1995; Hull 1990; Park and Cho 1995). In cases in which families demonstrate son preference, family members adopt methods to increase the probability of having boys rather than girls, such as using ultrasound and other diagnostic methods leading to sex selective abortions (Zeng et al. 1993). Infanticide and abandonment are also conducted in some rural areas in developing countries (Gu and Roy 1995).

Most previous works focus on the impact of son preference on sex ratio imbalance (Gu and Roy 1995; Hull 1990), but there is little mention in this area of literature of the impact of son preference on the educational attainments of children. A possible reason for the dearth of studies in this topic may be the unavailability of data on children's educational performance as well as measures of parental son preference.

Our research adopts linear probability, probit and ordered probit models to explore the impact of son preference on children's educational attainments. We use the data from the Gansu Survey of Children and Families (GSCF), conducted in 2000 by researchers at University of Pennsylvania in 20 counties within Gansu province, China, to investigate the impact of son preference on children's educational attainments. GSCF visited 2,000 children aged 9–12 and their schools and families in Gansu province, China. The survey includes questions about parental son preference and the educational attainments of children. We directly measure the

level of son preference using answers to survey questions. We use children's Chinese and math scores from the school year previous to the year the survey was conducted and grade retention data to measure children's educational attainments.

By using the GSCF data, we arrive at two interesting findings. First, we find that parental son preference has a very significant impact on children's educational attainments. On average, one standard deviation increase in the level of son preference decreases a child's Chinese and math grades by 1.05 and 1.01 points, respectively, out of 100 points, and increases the probability of grade retention by 9.99%. Second, we extend our analysis to study the mechanism of the impact of son preference on children's educational attainments. Our evidence suggests that parental son preference mainly works through the four channels: child-rearing methods, the mother's expectations for the child's schooling, the child's educational aspirations and the child's self-confidence.

Our study contributes to this relatively new area of study in the following ways:

First, our paper is the first to investigate how parental son preference affects children's educational outcomes. Documenting the impact of parental son preference on educational attainments is important, since being brought up by parents with strong son preference may have long-lasting effects on children's adult lives (Dahl and Moretti 2008). Educational outcomes have significant impacts on children's futures. For example, worse educational outcomes might lead to worse labor market outcomes for individuals (Dahl and Moretti, 2008). Therefore, the impact of parental son preference on children's educational outcomes is one of the most important aspects to investigate.

Second, this paper directly measures the effects of son preference on children's educational attainments. Since One Child Policy has been implemented in China for more than 20 years by 2000, it is difficult to identify son preference from sibling composition variation as what have been done in previous studies (Parish and Willis 1993; Butcher and Case 1994; Conley 2000). To directly measure son preference, we use from the GSCF data the answers to eight questions asking parents about their thoughts on gender-related issues pertaining to children, such as differences in abilities and opportunities between boys and girls, and several general questions regarding gender equality. To measure children's educational attainments, we look at children's Chinese and mathematics grades from the past year and whether children have experienced grade retention; these data are more detailed and precise than enrollment rate, years of education and dropout rate. Since the enrollment rate of elementary school and junior high school in year 2000 is 99.1% and 94.9% nationwide, respectively<sup>1</sup>, measures such as enrollment rate, years of education and dropout rate will vary little and thus shed little light on educational attainments among children younger than the age of 15.

Third, it is the first paper to discuss the mechanism of parental son preference on children's educational attainments, which could provide important policy implications. Son preference is deeply embedded in cultural and traditional beliefs and cannot be easily altered in the short term. However, if the mechanism of the impact of son preference on children's educational attainments can be understood, it is possible to eliminate the negative impacts to a large extent through policy intervention targeting specific perceptions and practices relevant to son preference, such as

---

<sup>1</sup> China Statistical Yearbook 2008

providing training on child-rearing methods for the parents, promoting the importance of education for girls and providing mental health consultation for children who are not confident or are slightly depressed due to a disharmonious family environment.

## **LITERATURE REVIEW**

The impact of son preference on the educational attainments of children is rarely investigated in the existing literature. Our paper is the first to explore the effects of parents' preference for sons over daughters on children's educational outcomes.

There are many studies on gender differences as they relate to children's educational attainments; there is consensus that these gender differences exist. For example, Wang (2005) , Song and Tan (2004) and Connelly and Zheng (2003) all found that particularly in the rural areas of China, children's educational attainments do differ between boys and girls. Boys are more likely to be sent to schools and to receive more years of education. In addition to these empirical studies, Mechtenberg(2009) provided the first theoretical explanation for the gender differences in education attainment using a cheap talk model of grading between teachers and students. Mechtenberg found that biased grading in schools results in boys having more top and more bottom achievers in math and sciences than girls.

According to some literature, sibling composition is likely to be one of the channels that influence these gender differences. For example, parents may invest less in daughters if they have younger brothers, compared to families with daughters that have either older or no brothers (Parish and Willis 1993). However, there is no consensus on the impact of sibling composition on children's educational outcomes. On the one hand, some studies saw no effect of sibling

composition on educational attainments (Chen, Chen and Liu 2008; Kuo and Hauser 1997). On the other hand, some studies found that the effect exists, but possibly varies in magnitude and cohorts (Butcher and Case 1994; Conley 2000; Parish and Willis 1993). For example, some studies found that the sibship effect is decreasing over time (Kevane and Levine 2001; Post and Pong 1998). There are also a couple of studies that found that the sibship effect is not uniform among different groups or various measures of educational attainments. For example, Kaestner (1997) found that in the United States, there is no sibling composition effect for whites. Among blacks, however, more sisters yield greater educational attainments. Powell and Steelman (1990) found that sibling composition has no effect on test scores in the United States, but there is an effect on grade-point average.

Son preference may be another important factor contributing to gender differences in children's educational attainments. There is evidence that it is the root cause of traditional values that promote gender inequality (OECD, 2009) and thus has strong explanatory power to gender differences in children's educational attainments. It is therefore worthwhile to look into the effects of son preference on children's educational outcomes.

## **DATA AND METHOD**

### **Data Source**

The data we use is from the Gansu Survey of Children and Families (GSCF), conducted by the researchers at University of Pennsylvania. This multilevel survey was designed to increase understanding of rural children's schooling and welfare in the context of poverty. It surveyed 2,000 children aged 9–12, along with their families, teachers, principals, and village leaders in

rural areas of 20 counties in Gansu Province, China<sup>2</sup>. The data is publicly available at CCER China Survey Center<sup>3</sup>. Information about children's educational attainments is obtained from the Target Teacher Form, a set of questions that ask teachers about their student performance. Information about son preference is obtained from the Mother Form, a set of questions that ask mothers about their opinions on issues relating to their children and children in general. Information about family background is provided in the Household Form, a set of questions that ask about background information of family members and household assets and liabilities. Information about mechanism of son preference's impact on educational attainments is provided in Child Form, a set of questions that ask children about their opinions and activities, as well as Mother Form.

## **Variable Definition**

### **Child's Educational Attainments Variables**

Child's Educational Attainments measure the academic performance of children at school. In the Target Teacher Form, there are questions asking for the average Chinese and mathematics grades of children from the school year prior the one during which the survey was conducted, and one question asking whether the child has ever experienced grade retention. In this paper, we

---

<sup>2</sup> Gansu is a province located in the northwest of the People's Republic of China. It has an area of 454,000 km<sup>2</sup> and a population of nearly 31 million (2009). Gansu is one of the poorest provinces in China. Its nominal GDP for 2008 was about 317.6 billion yuan (45 billion USD) and per capita of 12,110 RMB (1,744 USD).

<sup>3</sup> <http://www.chinasurveycenter.org/>

use the three variables, the average Chinese grade, average mathematics grade and occurrence of grade retention, to measure the child's educational attainment levels (Please refer to Table 1 for variable definitions). The Chinese and mathematics grades are measured as scores between 0-100. Grade retention is a dummy variable: "1" denotes the child has repeated a grade; otherwise it takes the value "0."

### **Son Preference Variable**

Section D in Mother Form asks mothers to answer questions about both the mother's and the husband's shared views and opinions on several issues. The first part of this section, Questions 1–8, comprise of questions about son preference. Questions 1, 2, 3, 6 and 7 ask the mother about the views she and her husband share on the differences between sons and daughters in terms of abilities and opportunities, which effectively measure their level of son preference. For example, if parents believe sons are more able and have better opportunities than girls, they are more willing to rear boys than girls by allocating more family resources or parental care for boys. These five questions are worded as statements to be answered based on whether the mother and father "agree," "disagree" or "have no opinion." These questions are as follows: Question 1 is "As long as they work hard, both girls and boys can get good grades"; Question 2 is "Girls should have as many educational opportunities as boys"; Question 3 is "If both are given equal opportunities, women can do as well as men"; Question 6 is "Parents should encourage girls to have their own opinions, just like the boys"; and Question 7 is "It is useless to send girls to school, since girls will inevitably get married."

Questions 4, 5 and 8 address the parent's perceptions of more general gender equality issues.

Since for parents son preference grows from belief in gender inequality, it is necessary to include these three questions in our measurement of son preference. These three questions are as follows: Question 4 is “If both husband and wife have full-time jobs, they should share the housework”; Question 5 is “The important decisions at home should be made by the men”; Question 8 is “Parents should rely on their sons for old-age support.”

As mentioned previously, the mother is asked whether she and her husband “agree” or “disagree” with or “have no opinion” on the above mentioned statements. We code each answer as a multinomial variable of “0”, “1” or “2.” For statements in favor of gender equality (Questions 1, 2, 3, 4 and 6), “0” denotes “agree”, “1” denotes “have no opinion” and “2” denotes “disagree”. For statements in favor of son preference (Question 5, 7 and 8), “0” denotes “disagree,” “1” denotes “have no opinion,” and “2” denotes “agree.” Thus, with regard to the level of son preference, “0” corresponds to “no preference,” “1” corresponds to “some son preference” and “2” corresponds to “strong son preference.” We then use the average value of the eight answers as the measure of son preference. The higher the value means the stronger the son preference. The value is defined on a scale of 0–2.

## **Variables on the Possible Channels of the Impact of Son Preference on Children's**

### **Educational Attainments**

### **Child-rearing methods.**

In the Section K of the Mother Form, the purpose of Question 1 is to help us understand parents’ child-rearing methods. The question is stated as follows: “Last month, did you, your husband or other adults frequently do any of the following things with the child? Activities

include: (1) accompany the child to read storybooks; (2) help the child to do his/her assignments; (3) do family chores with the child, such as washing clothes, dishes, cooking, etc.; (4) do activities that the child likes with the child, such as playing cards, playing hide-and-seek, playing ball, etc.; (5) take the child to bookstores or shops; (6) praise the child; (7) show affection to the child, such as hugging, patting, etc.; (8) scold the child; (9) beat the child; (10) highly praise the child in front of others; and (11) discuss with the child topics of his/her interest.

The mother is asked the frequency (“never,” “sometimes” or “often”) of doing such activities with their children. We code mother’s answers as “0,” “1” or “2”. For Activities 1–7, 10 and 11, which address child-rearing methods that reward the child, “0” denotes “never,” “1” denotes “sometimes” and “2” denotes “often.” For Activities 8 and 9, which address punitive and possibly harmful child-rearing methods, “0” denotes “often,” “1” denotes “sometimes” and “2” denotes “never.” Thus, regarding child-rearing methods, “0” denotes punitive child-rearing methods, “1” denotes moderate child-rearing methods and “2” denotes child-rearing methods that reward the child. Then we use the average value of the 11 answers to measure child-rearing methods. The higher the value means the more rewarding and less punitive the child-rearing methods. The value is defined on a scale of 0–2.

### **Mother’s Expectations for Child’s Schooling.**

In Section J of the Mother Form, Question 12 asks “What is the highest grade that you wish your child could achieve?” We use the answer to this question to measure the mother’s expectations for child’s schooling, which is a multinomial variable. The value “1” denotes graduating from primary school, “2” denotes graduating from junior high school, “3” denotes

graduating from senior high school, “4” denotes graduating from college and “5” denotes graduating from an institution of a level higher than college.

### **Child’s Educational Aspirations.**

In Section E of the Child Form, Question 1 asks “What is the highest level of schooling that you want to complete?” We use the answer to this question to measure child’s educational aspirations, which is a multinomial variable. The value “1” denotes primary school, “2” denotes middle school, “3” denotes high school, “4” denotes vocational high school, “5” denotes vocational college and “6” denotes college or above.

### **Child’s Self-Confidence.**

In Section G of Child Form, Question 1d, phrased as a statement, is “I am confident about my future.” We use the answer to this question to measure the child’s self-confidence, which is a multinomial variable. The value “0” denotes totally disagree, “1” denotes disagree, “2” denotes agree and “3” denotes fully agree.

### **Control Variables**

The control variables in the model are the child’s gender, the number of siblings the child has, parents’ years of education, parents’ marital status and family economic situation (see Table 3 for descriptive statistics). The definition of each variable can be found in Table 1. The number of observations in each regression result can be found in Table 2.

### **Method**

Our study investigates the impact of parental son preference on children’s educational attainments. The baseline model we adopt is

$$OUTCOME_i = \alpha_0 + \alpha_1 SP_i + \alpha_2 SP_i * GIRL_i + \alpha_3 X_i + \varepsilon_i \quad (1)$$

Where  $OUTCOME_i$  denotes the educational attainments of the  $i^{th}$  child.  $SP_i$  measures the parental son preference of the  $i^{th}$  child.  $SP_i * GIRL_i$  is the interaction term of son preference and the gender of the child.

$X_i$  denotes several control variables, which are gender, the number of siblings, the parental educational levels and marital status, the family economic situation and village level fixed effects.

If son preference has negative impact on children's educational attainments, we should expect  $\alpha_1$  to be significantly negative when  $OUTCOME_i$  is Chinese grade and mathematics grade.  $\alpha_1$  should be significantly positive when  $OUTCOME_i$  is grade retention.

Moreover, we add an interaction term  $SP_i * GIRL_i$  to measure differential impacts of son preference on boys and girls. For example, if there are both sons and daughters in the same family, son preference might have larger negative effect on girls than on boys because parents may transfer more of the parental care to boys. However, as mentioned in the literature review, there is no consensus of the effect of sibling composition on educational attainments. Thus we cannot confidently hypothesize that girls would be affected more than boys by parental son preference. If there are larger negative effects of son preference on girls than on boys, we should expect  $\alpha_2$  to be significantly negative when  $OUTCOME_i$  is Chinese grade and mathematics grade.  $\alpha_2$  should be significantly positive when  $OUTCOME_i$  is grade retention.

Our study further investigates the mechanism of son preference on children's educational attainments. To test whether son preference is correlated with our proposed channels of its

impact on children's educational outcomes, we adopt the following model:

$$FACTOR_i = \beta_0 + \beta_1 SP_i + \beta_2 X_i + \varepsilon_i \quad (2)$$

Where  $FACTOR_i$  denotes the mechanism of parental son preference on the  $i^{th}$  child, including scientific child-rearing methods, good mother's expectations for child's schooling, high child's educational aspirations and high child's self-confidence.  $SP_i$  measures the parental son preference of the  $i^{th}$  child.  $X_i$  denotes several control variables, including the gender, the number of siblings, the parental educational levels and marital status, the family economic situation and village level fixed effects.

We expect son preference should have negative effect on the four mechanism variables. For example, higher son preference should be correlated with worse child-rearing methods, lower mother's expectations for child's schooling, lower child's educational aspirations and lower child's self-confidence. Therefore, if the variable  $FACTOR_i$  is a possible channel of son preference on children's educational attainments,  $\beta_1$  should be significantly negative.

After showing that families with strong son preference also have worse child-rearing methods, lower mother's expectations for child's schooling, lower child's educational aspirations and lower child's self-confidence, we add the variables indicating son preference mechanism into equation (1) as control variables. The model is then written as follows:

$$OUTCOME_i = \theta_0 + \theta_1 SP_i + \theta_2 SP_i * GIRL_i + \theta_3 FACTOR_i + \theta_4 X_i + \varepsilon_i \quad (3)$$

The definitions of  $OUTCOME_i$ ,  $SP_i$ ,  $SP_i * GIRL_i$ ,  $FACTOR_i$  and  $X_i$  are the same as model (1) and (2). If the  $FACTOR_i$  variables are the possible channels of the impact of son preference on children's educational attainments, then we should see that in equation (3) the coefficient  $\theta_1$  turns

to be insignificant, and  $\theta_3$  is significantly positive.

## **Estimation Method**

In this article, we use linear probability, probit and ordered probit model as the main estimation techniques. Specifically, for those equations whose dependent variables are continuous (e.g., Chinese grade and mathematics grade of child  $i$ ), we use linear probability model for estimation; for those whose dependent variables are binary (e.g., grade retention experience of child  $i$ ), we use probit model for estimation; for those whose variables are ordinal multinomial (e.g., child-rearing methods, mother's expectations for child's schooling, child's educational aspirations and child's self-confidence), we use linear probability model for estimation, and use ordered probit Model as robustness tests.

## **RESULTS**

### **Summary of Statistics**

Table 3 shows summary statistics. The first column shows the summary of statistics of the whole sample. We define the level of son preference as three categories: “no son preference” if  $SP \in [0, 0.375]$ , “some son preference” if  $SP = 0.5$  and “strong son preference” if  $SP \in [0.625, 2]$ , which is approximately a trisection of the observations. We use the “no son preference” group as the baseline and compare the mean of children's educational attainment levels, the variables indicating the mechanism of son preference and the control variables with the other two groups. For most of the variables, the “some son preference” group and “strong son preference” group differ significantly from the “no son preference” group. For example, the mean Chinese grade in “no son preference” group is 73.09, while the mean is 71.28 and 70.74 in the “some son

preference” group and “strong son preference” group, respectively. The same trend exists with mathematics grades. The probability of grade retention in the “no son preference” group is 15%, which is significantly lower than “some son preference” group (25%) and “strong son preference” group (26%). Also, the means of parental educational levels, family economic situation, child-rearing methods, the mother's expectations for child's schooling, the child's educational aspirations and the child's self-confidence in the “no son preference” group is significantly higher than the mean of “some son preference” group and “strong son preference” group, respectively.

Table 4 shows the regression results of the impact of son preference on children's educational attainments. Three dependent variables are Chinese grade in the last year, mathematics grade in the last year and whether the child repeated a grade. The explanatory variables include son preference (the higher value means the stronger the son preference), the interaction term of son preference and child's gender (Son Preference\*Girls) and other control variables. (Panels A and B show the impact of son preference on children's educational attainments with and without the control variables.) In all the six specifications, the coefficients of son preference are all significant at the 0.05 level, and the signs of coefficients support our hypothesis. The results show that the stronger the parental son preference is, the lower the child's prior Chinese and mathematics grades are, and the higher the probability of grade retention is. We interpret the coefficients using a one-standard-deviation change in son preference, about 26.7 percentage points. One standard deviation increase of son preference decreases child's Chinese and math grades by  $1.05(3.925*26.7\%)$  and  $1.01(3.794*26.7\%)$  points, respectively, and increases the

probability of grade retention by 9.99% ( $37.4\% * 26.7\%$ ). These results prove that son preference has a very significant impact on child's educational attainment that should not be ignored.

Furthermore, the coefficients on the interaction term of son preference and the child's gender (Son Preference\*Girl) are not significant, which means that there is no marginal difference between the impact of son preference on girls and boys. In other words, within the same family, the impact of son preference on the educational attainments of boys and girls should be the same. Moreover, the number of siblings does not have significant effects on educational attainments. Therefore, our results cannot provide evidence to support the hypothesis that sibling composition affects children's educational attainments.

Table 4 Panel B also shows that the parents' years of education has a significant impact on the child's educational outcomes. The more years of education of parents have had, the higher the child's Chinese and mathematics grades are and the lower probability of grade retention. Also, the child's gender, the number of siblings, the family economic situation and parental marital status do not significantly impact on the child's educational outcomes.

### **Mechanism of Son Preference**

Based on the results in Table 4, we come to the conclusion that parental son preference does have a significant impact on the child's educational attainments. The stronger the son preference is, the worse the child's educational attainments are. Since son preference is formed by inherited social norms, it is hard to change in the short term. However, if intervention can be organized to influence the channels that reinforce the preference for sons among parents, the children's educational outcomes in families with strong son preference may be improved.

In this paper, we propose four channels that lead to and reinforce son preference and its impact on education. First, parental child-rearing methods may reinforce son preference and in turn affect the child's educational attainment levels. For example, parents who show a strong preference for sons may not be well educated, and therefore they may not exercise positive, rewarding child-rearing methods; this may affect the child's educational outcomes. Second, son preference may affect the child's educational outcomes through the mother's expectations for the child's schooling. Parents who are very conservative may have strong son preference as well as low expectations for the child's schooling, which may negatively affect the child's educational attainments. Third, son preference may not only affect the child's educational outcomes through the mother's expectations for the child's schooling, but also through influencing the child's educational aspirations; parents with strong son preference may instill in the child the traditional view that education is not important, which can thus lower the child's educational aspirations and thus affects his or her educational outcomes. Finally, the child's self-confidence may be an avenue through which son preference can influence the child's educational attainment levels. For example, parents with strong son preference may be more likely to exert their will on their children, which may affect their self-confidence and thus in turn affect their educational outcomes.

Table 5 examines whether the proposed four channels—child-rearing methods, the mother's expectations for the child's schooling, the child's educational aspirations and the child's self-confidence—are correlated with son preference. In Panel A of Table 5, we use a linear probability model for estimation. Since the mother's expectations for the child's schooling, the

child's educational aspirations and the child's self-confidence are multinomial dependent variables, we use an ordered probit model to re-estimate the three specifications to check for robustness in Panel B of Table 5.

Among the results in Panel B of Table 5, there are some variations among the magnitudes of the coefficients of son preference, but the sign and significance changes very little. After controlling for the child's gender, the number of siblings, parental educational levels, family economic situation and parental marriage status, we find that the four channels are all significantly correlated with son preference at the 0.1 level or above. Higher son preference leads to lower values for the mother's expectations for child's schooling, the child's educational aspirations and the child's self-confidence, and worse child-rearing methods. One standard deviation increase in son preference decreases the mother's expectations for the child's schooling and the child's educational aspirations by a scale of 0.118 ( $0.442*26.7\%$ ) and  $0.109(0.410*26.7\%)^4$ , respectively. One standard deviation increase of son preference also decreases parents' child-rearing methods and child's self-confidence by a unit of 0.033 ( $0.08*26.7\%$ )<sup>5</sup> and 0.039 ( $0.146*26.7\%$ )<sup>6</sup>, respectively.

---

<sup>4</sup> There are five scales for mother's expectations: 1 Graduate from primary school; 2 Graduate from junior high school; 3 Graduate from senior high school; 4 Graduate from college; 5 Graduate from higher than college. There are six scales for child's educational aspirations: 1 Primary school; 2 Middle school; 3 High school; 4 Vocational high school; 5 Vocational college; 6 University or above

<sup>5</sup> Child-rearing method is defined on a scale between 0-2, the higher value means the better

Some control variables also have significant explanatory power. For example, the mother's expectations for girls' schooling are significantly lower than that for boys. Girls' educational aspirations are significantly lower than those of boys. The greater the number of years of education the parents received, the higher the mother's expectations for the child's schooling and educational aspirations, the better the child-rearing methods and the child's self-confidence.

The results in Table 5 show evidence that the four channels we mentioned before are significantly correlated with son preference. If they are the channels that affect son preference and thus indirectly affect the child's educational attainments, we should see that after adding these four channels as control variables, either son preference no longer has a significant correlation with the child's educational attainments, or there is a dramatic decrease in the magnitude of the son preference coefficient. Table 6 shows the regression results with and without the four channels as control variables. Columns 1, 3, 5 are the results without adding these variables and Columns 2, 4, 6 are the results with these variables included.

The results of Table 6 show that after adding the four channels, son preference does not have significant correlation with the child's educational attainments. At the same time, the coefficients of the four channels turn to be significant in the model. For example, in the model which dependent variable is Chinese grade, the coefficient of son preference changes from -3.925 at the

---

child-rearing methods.

<sup>6</sup> Child's self-confidence is defined on a scale between 0-3, the higher value means the child is more self-confident.

0.05 level to -2.759, insignificant. However, the coefficients of the four channels are all significantly correlated with the child's educational attainments at least at the 0.1 level. Better child-rearing methods, higher mother's expectations on child's schooling, higher child's educational aspirations and higher child's self-confidence lead to higher children's educational attainments. For example, in the model which dependent variable is mathematics grade, the coefficient of son preference changes from -3.794 at the 0.05 level to an insignificant -2.341. In the model which dependent variable is grade retention, after adding the four channels, the coefficient of son preference changes from 0.374 at the 0.05 level to 0.308 which is not significant at all.

The same as the results in Table 4, the interaction term of son preference and child's gender is still not significant. In terms of other covariates, the parents' years of education is still significantly correlated with the child's educational outcomes.

From the above results, we can come to the conclusion that child-rearing methods, mother's expectations for child's schooling, child's educational aspirations and child's self-confidence are important channels of son preference on child's educational attainments.

### **Robustness Check**

To test the robustness of our results, we change the definition of son preference into other two measures. Firstly, we define a new variable 'Son Preference' as 1 if  $\text{Son Preference} \geq 0.625$ , which is approximately the 70 percentile of Son Preference, 0 otherwise. Secondly, we define 'Son Preference' as 2 if  $\text{Son Preference} \in [0.625, 2]$ , 1 if  $\text{Son Preference} \in [0.5, 0.625)$ , 0 if  $\text{Son Preference} \in [0, 0.5)$ . We use these two new measurements to replicate Table 6 respectively.

The results are shown in Table 7 and Table 8. We can see that in the models which dependent variables are Chinese grade and mathematics grade, the above results still hold, but there is a slightly decrease in the significance level of son preference coefficients. In the model which dependent variable is grade retention, the coefficients are both insignificant before and after adding the four channels. And the four channels still have strong explanatory power in all the models. Generally, the above results prove that our model is robust to some extent.

## **CONCLUSION**

In this paper, we first show that parental son preference has a significant impact on children's educational attainments. We find that stronger son preference leads to lower educational attainments, measured by Chinese and mathematics grades in the past year and whether a child had to repeat a grade.

In the second part of this study, we look further into the mechanism of son preference effects on children's educational attainments. We find that parental son preference is significantly correlated with child-rearing methods, the mother's expectations for child's schooling, the child's educational aspirations and the child's self-confidence.

In the third part of this paper, we find that after controlling for the four channels indicated above, son preference no longer has a significant correlation with children's educational attainments. Moreover, the four channels are significantly correlated with children's educational outcomes. The results suggest that these four channels are integral to the mechanism of the impact of son preference on children's educational attainments.

Our results have important policy implications. Since son preference is deeply ingrained in some communities and cannot be changed in the short term, it is worthwhile to examine its impact on children's educational outcomes. Furthermore, our results show that the impact of son preference can be eliminated to a large extent if there could be policy intervention in regard to the four aforementioned channels, such as providing training for parents on rearing methods, promoting the importance of education for girls and providing mental health consultation for children may lack confidence or are slightly depressed due to their family environment. This idea has never been proposed in previous literature and is one of the most important contributions of our paper.

There are several limitations in our study. With respect to the data we use, the first problem is the sample representativeness. The GSCF data was collected only in 20 counties in Gansu Province, China, making our results susceptible to generalization. Second, we use the average score of a series of questions to measure parental son preference and child-rearing methods. Since there is not any literature on the measurement of these two variables, the viability of these two variables needs to be further tested.

With respect to the mechanism of the impact of son preference on children's educational attainment levels, we discuss four channels that may comprise this mechanism in this paper: child-rearing methods, the mother's expectations for child's schooling, the child's educational aspirations and the child's self-confidence. After controlling for these channels, son preference is no longer significantly correlated with the child's educational attainment levels, which means that these four channels comprise the mechanism of the effects of son preference on educational

outcomes. However, other mechanisms may exist in addition to the one explored in this paper, as other channels contributing to son preference may be identified in future studies.

Last but not least, there might be selection bias in the measurement of son preference. Those who have strong son preference may use sex-selective abortion to decrease the number of girls in the families, which brings underestimation of the impact of son preference on child's educational attainment levels. Therefore, the coefficients may be downwardly biased. Furthermore, there might be reporting bias on the measurement of son preference. People may not want to express their preference for sons, since the mainstream media supports the notion of gender equality. Thus, the magnitude of son preference may be underestimated, which also leads to downward bias of the coefficient on son preference. Since the main results show that the coefficient of son preference is always significant, the impact of son preference should be more significant in magnitude after taking into consideration the selection and reporting biases.

## **References:**

- Andersson, G., K. Hank, M. Ronsen, and A. Vikat. 2006. "Gendering Family Composition: Sex Preferences for Children and Childbearing Behavior in the Nordic Countries " *Demography* 43(2):255-267.
- Butcher, K.F. and A. Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *The Quarterly Journal of Economics* 109(3):531-563.
- Chen, S.H., Y.C. Chen, and J.T. Liu. 2008. "The Impact of Sibling Sex Composition on Women's Educational Achievements: A Unique Natural Experiment by Twins Gender Shocks."

mimeo, SUNY-Albany, Department of Economics.

- Conley, D. 2000. "Sibship Sex Composition: Effects on Educational Attainment." *Social Science Research* 29(3):441-457.
- Connelly, R. and Z. Zheng. 2003. "Determinants of school enrollment and completion of 10 to 18 year olds in China." *Economics of Education Review* 22(4):379-388.
- Dahl, G.B. and E. Moretti. 2008. "The Demand for Sons: Evidence from Divorce, Fertility, and Shotgun Marriage." *Review of Economic Studies* 75(4):1085-1120.
- Das Gupta, M. 2007. "China's "Missing Girls"—Son Preference or Hepatitis B Infections?" in *World Bank Research Brief (April 2)*: World Bank, Washington D.C.
- Gu, B. and K. Roy. 1995. "Sex Ratio at Birth in China, with Reference to Other Areas in East Asia: What We Know." *Asia Pacific Population Journal* 10(3):17-42.
- Hull, T.H. 1990. "Recent Trends in Sex Ratios at Birth in China." *Population and Development Review* 16(1):63-83.
- Kaestner, R. 1997. "Are Brothers Really Better? Sibling Sex Composition and Educational Achievement Revisited " *Journal of Human Resources* 32(2):250-284.
- Kevane, M. and D. Levine. 2001. "The Changing Status of Daughters in Indonesia." in Institute for Research on Labor and Employment, UC Berkeley, Working Paper Series.
- Kingdon, G.G. 2005. "Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure." *Economic Development and Cultural Change* 53(2):409-451.
- Kuo, H.H.D. and R.M. Hauser. 1997. "How Does Size of Sibship Matter? Family Configuration

- and Family Effects on Educational Attainment." *Social Science Research* 26(1):69-94.
- Mechtenberg, L. 2009. "Cheap Talk in the Classroom: How Biased Grading at School Explains Gender Differences in Achievements, Career Choices and Wages " *Review of Economic Studies* 76(4):1431-1459.
- Parish, W.and R. Willis. 1993. "Daughters, Education, and Family Budgets Taiwan Experiences." *The Journal of Human Resources* 28(4):863-898.
- Park, C.B.and N.-H. Cho. 1995. "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea." *Population and Development Review* 21 (1):59-84.
- Post, D.and S.L. Pong. 1998. "The Waning Effect of Sibship Composition on School Attainment in Hong Kong." *Comparative Education Review* 42(2):99-117.
- Powell, B.and L.C. Steelman. 1990. "Beyond Sibship Size: Sibling Density, Sex Composition, and Educational Outcomes." *Social Forces* 69(1):181-206.
- Song, Y.and L. Tan. 2004. "Lun Wo Guo Ji Chu Jiao Yu De Xing Bie Gong Ping (Chinese)." *Collection of Women's Studies* 58(2):21-27.
- The OECD Social Institutions and Gender Index Results 2009  
[http://www.oecd.org/document/39/0,3343,en\\_2649\\_33935\\_42274663\\_1\\_1\\_1\\_1,00.html#results](http://www.oecd.org/document/39/0,3343,en_2649_33935_42274663_1_1_1_1,00.html#results)
- Wang, W. 2005. "Son Preference and Educational Opportunities of Children in China--" I Wish You Were a Boy!"" *Gender Issues* 22(2):3-30.
- Zeng, Y., P. Tu, B. Gu, Y. Xu, B. Li, and Y. Li. 1993. "Causes and Implications of the Recent

Increase in the Reported Sex Ratio at Birth in China." *Population and Development Review*

19(2):283-302.

**Table 1. Definitions of Variables**

<b>Dependent variables</b>	
Prior Chinese Grade	Average Chinese grade in the last year prior to the survey
Prior Mathematics Grade	Average Mathematics grade in the last year prior to the survey
Grade Retention	1 If the child has ever repeated a grade;0 Otherwise
<b>Independent variables</b>	
<b>Son Preference variables</b>	
Son Preference	<p>The value takes the average of the following eight scores, the bigger the value, the stronger the son preference. The value is between 0-2. Eight questions are as follows:</p> <ol style="list-style-type: none"> <li>1. As long as they work hard, both girls and boys can get good grades 2 Disagree; 1 Have No Opinion; 0 Agree;</li> <li>2. Girls should have as many educational opportunities as boys. 2 Disagree; 1 Have No Opinion; 0 Agree;</li> <li>3. If both are given equal opportunities, women can do as well as men 2 Disagree; 1 Have No Opinion; 0 Agree;</li> <li>4. If both husband and wife have full-time jobs, they should share the housework 2 Disagree; 1 Have No Opinion; 0 Agree;</li> <li>5. The important decisions at home should be made by the men 0 Disagree; 1 Have No Opinion; 2 Agree;</li> <li>6. Parents should encourage girls to have their own opinions just like the boys 2 Disagree; 1 Have No Opinion; 0 Agree;</li> <li>7. It is useless to send girls to school since girls will inevitably get married 0 Disagree; 1 Have No Opinion; 2 Agree;</li> <li>8. Parents should rely on their sons for old-age support 0 Disagree; 1 Have No Opinion; 2 Agree;</li> </ol>
<b>Control Variables</b>	
Girl	1 Girl 0 Boy
Number of Siblings	Number of Siblings of the Child
Father's Education (Years of Schooling)	Father's Years of Schooling
Mother's Education (Years of Schooling)	Mother's Years of Schooling
Economic Situation (3=Very Good, 2= Above Average; 1=Below Average; 0= Very Bad)	How would you rate your family's economic situation in the context of your village? 3 Very Good; 2 Above Average 1 Below Average 0 Very Bad
Intact Family	1 Intact Family; 0 Otherwise
<i>Channels of Son Preference on Children's Educational</i>	
Child-Rearing Methods	<p>The value takes the average of the following eleven scores, the bigger the value, the more rewarding and healthier the child-rearing methods. The value is between 0-2. Eleven questions are as follows:</p> <p>Last month, did you, your husband or other adults frequently do the following things with the child?</p> <ol style="list-style-type: none"> <li>1. Accompany the child to read storybooks 0 Never; 1 Sometimes; 2 Often</li> <li>2. Help the child to do his assignments 0 Never; 1 Sometimes; 2 Often</li> <li>3. Do family chores with the child, such as washing clothes, dishes, cooking etc.; 0 Never; 1 Sometimes; 2 Often</li> <li>4. Do activities that the child likes with the child, such as playing cards, playing hide-and-seek, playing ball etc.; 0 Never; 1 Sometimes; 2 Often</li> <li>5. Take the child to bookstores or shops 0 Never; 1 Sometimes; 2 Often</li> <li>6. Praise the child 0 Never; 1 Sometimes; 2 Often</li> <li>7. Show affection to the child, such as hugging, patting etc. 0 Never; 1 Sometimes; 2 Often</li> <li>8. Scold the child 2 Never; 1 Sometimes; 0 Often</li> <li>9. Beat the child 2 Never; 1 Sometimes; 0 Often</li> <li>10. Highly praise the child in front of others 0 Never; 1 Sometimes; 2 Often</li> <li>11. Discuss with the child on the topic of his/her interests 0 Never; 1 Sometimes; 2 Often</li> </ol>
Mother's Expectations for Child's Schooling	What is the highest grade that you wish your child could achieve? 1 Graduate from primary school;2 Graduate from junior high school; 3 Graduate from senior high school; 4 Graduate from college; 5 Graduate from higher than college
Child's Educational Aspirations	What is the highest level of schooling that you want to complete? 1 Primary school;2 Middle school;3 High school;4 Vocational high school;5 Vocational college;6 College or above
Child's Self-Confidence	I am confident about my future. 0 Totally disagree 1 Disagree 2 Agree 3 Totally agree

**Table 2.****Number of Observations**

<b>Data</b>	<b>Number Excluded</b>	<b>Number Remaining</b>	<b>Used in Table No.</b>
Gansu Survey of Children and Families (GSCF) 2000		2,000	
Exclude if Prior Chinese Grade is missing	20	1,980	
Exclude if Grade Retention is missing	3	1,977	
Exclude if Son Preference is missing	32	1,945	
Exclude if Mother's Education is missing	2	1,943	
Exclude if Economic Situation is missing	4	1,939	
Exclude if Intact Family is missing	23	1,916	
Exclude if Child-Rearing Methods is missing	7	1,909	
Exclude if Mother's Expectations for Child's Schooling is	3	1,906	
Exclude if Child's Educational Aspirations is missing	8	1,898	
Exclude if Girl is missing	3	<u>1,895</u>	Table 4, Panel A, Column 1-2; Table 4, Panel A, Column 1-2; Table 5; Table 6 Column 1-4; Table 7 Column 1-4; Table 8
Loss of observations in Probit Model with Village Fixed Effect <sup>a</sup>	57	<u>1,838</u>	Table 4, Panel A, Column 3; Table 4, Panel B, Column 3; Table 6, Column 5-6; Table 7, Column 5-6; Table 8, Column 5-6;

<sup>a</sup>Probit model with village fixed effect leads to a loss of 57 observations, with county code equals to 12. This is attributed to the "Separation Problem" of Probit model, the main reason of which is sparse data and strong collinearity. (Zorn, 2005)

**Table 3. Descriptive Statistics for Variables Included in the Model**

Variable	All		No Son Preference <sup>a</sup>		Some Son Preference		Strong Son Preference	
	(Mean)	(SD)	(Mean)	(SD)	(Mean)	(SD)	(Mean)	(SD)
<i>Dependent Variables</i>								
Prior Chinese Grade (0-100)	71.71	15.40	73.09	15.33	71.28**	15.75	70.74**	15.10
Prior Mathematics Grade (0-100)	73.40	16.13	74.91	15.66	72.96**	16.54	72.29***	16.15
Grade Retention (0=No, 1=Yes)	0.22	0.41	0.15	0.36	0.25***	0.43	0.26***	0.44
<i>Control Variables</i>								
Girl (0=Boy, 1=Girl)	0.46	0.50	0.46	0.50	0.46	0.50	0.45	0.50
Mother's Education (Years of Schooling)	4.25	3.51	5.37	3.51	3.84***	3.46	3.50***	3.27
Father's Education (Years of Schooling)	7.13	3.78	7.89	3.78	6.91***	3.53	6.56***	3.87
Number of Siblings	1.33	0.78	1.32	0.83	1.35	0.76	1.33	0.73
Economic Situation (3=Very Good, 2=Above Average, 1=Below Average, 0=Very Bad)	1.31	0.77	1.31	0.73	1.29	0.77	1.33	0.82
Intact Family (0=No, 1=Yes)	0.99	0.12	0.99	0.10	0.98	0.13	0.98*	0.14
<i>Channels of Son Preference on Educational Attainments</i>								
Child -Rearing Methods (0=Worst, 2=Best)	0.90	0.28	0.96	0.26	0.89***	0.27	0.86***	0.30
Mother's Expectations for Child's Schooling (1=Finish Primary School, 2=Finish Middle School, 3=Finish High School, 4=Finish University, 5=Higher)	3.58	0.72	3.74	0.55	3.59***	0.69	3.42***	0.85
Child's Educational Aspirations (1=Finish Primary School, 2=Finish Middle School, 3=Finish High School, 4=Finish University, 5=Higher)	3.58	1.64	3.79	1.56	3.49***	1.67	3.45***	1.66
Child's Self-Confidence (0=No Confidence, 3=Very Confident)	2.32	0.77	2.40	0.72	2.32**	0.76	2.25***	0.81
Observations	1,895		658		555		682	

Notes: \*\*\* Significant at the 1% level, \*\* Significant at the 5% level, \* Significant at the 10% level. Standard error in parenthesis. <sup>a</sup>The observations are divided into three categories based on the level of son preference. Since Son Preference is a discrete variable defined between 0-2 with the smallest jump of 0.125, we define the three categories as: No Son Preference [0,0.375], Some Son Preference [0.5,0.5], Strong Son Preference [0.625,2], which is approximately a trisection of the observations. T-tests are conducted between the three categories, taking "No Son Preference" category as the benchmark. Please refer to Table 1 for the definitions of all the variables.

**Table 4. The Impact of Son Preference on Child's Educational Attainments**

<b>Panel A. Without Control Variables</b>			
<b>Variable</b>	<b>Prior Chinese Grade</b>	<b>Prior Mathematics Grade</b>	<b>Grade Retention</b>
Son Preference	-3.631 (1.349)***	-3.279 (1.429)**	0.278 (0.142)**
Village Fixed Effect	Yes	Yes	Yes
Observations	1,895	1,895	1,838
R-squared	0.13	0.11	0.16
<b>Panel B. With Control Variables</b>			
<b>Variable</b>	<b>Prior Chinese Grade</b>	<b>Prior Mathematics Grade</b>	<b>Grade Retention</b>
Son Preference	-3.925 (1.786)**	-3.794 (1.901)**	0.374 (0.190)**
Son Preference*Girl	2.388 (2.513)	2.565 (2.676)	-0.386 (0.269)
Girl	1.413 (1.462)	-0.366 (1.557)	0.172 (0.163)
Number of Siblings	0.009 (0.474)	-0.299 (0.504)	0.020 (0.048)
Father's Education (Years of Schooling)	0.280 (0.097)***	0.278 (0.103)***	-0.024 (0.011)**
Mother's Education (Years of Schooling)	0.218 (0.114)*	0.148 -0.121	-0.029 (0.013)**
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very Bad)	1.248 (0.459)***	1.189 (0.489)**	-0.049 (0.049)
Intact Family	-1.169 (2.792)	0.239 (2.973)	0.399 (0.31)
Village Fixed Effect	Yes	Yes	Yes
Observations	1,895	1,895	1,838
R-squared	0.15	0.12	0.17

*Notes:* \*\*\* Significant at the 1% level.\*\* Significant at the 5% level.\* Significant at the 10% level.Standard error in parenthesis.The data source is GSCF 2000 Survey. There are 2000 observations in the dataset. After deleting all the missing values, there are 1895 observations for OLS Model(Column 1,2 of Panel A and B) and 1838 observations for Probit Model (Column 3 of Panel A and B). The R-squared in Column 3, Panel A and B is Pseudo R-squared. Please refer to Table 1 for the definitions of all the variables.

**Table 5. Possible Channels of the Impact of Son Preference on Children's Educational Attainments**

<b>Panel A. OLS Estimation</b>				
<b>Variable</b>	Mother's Expectations for Child's Schooling	Child-Rearing Methods	Child's Educational Aspirations	Child's Self-Confidence
Son Preference	-0.442 (0.063)***	-0.078 (0.023)***	-0.410 (0.146)***	-0.146 (0.070)**
Girl	-0.163 (0.032)***	0.016 (0.012)	-0.223 (0.074)***	0.018 (0.036)
Number of Siblings	-0.003 (0.022)	-0.001 (0.008)	-0.106 (0.051)**	-0.039 (0.025)
Father's Education (Years of Schooling)	0.013 (0.005)***	0.005 (0.002)***	0.030 (0.011)***	0.009 (0.005)*
Mother's Education (Years of Schooling)	0.01 (0.005)**	0.007 (0.002)***	0.011 (0.012)	-0.006 (0.006)
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very	0.069 (0.021)***	0.043 (0.008)***	0.147 (0.050)***	0.005 (0.024)
Intact Family	-0.116 (0.129)	0.027 (0.047)	0.084 (0.302)	-0.187 (0.145)
Observations	1,895	1,895	1,895	1,895
R-squared	0.17	0.27	0.12	0.07
<b>Panel B. Ordered Probit Model in Column 1,3,4</b>				
<b>Variable</b>	Mother's Expectations for Child's Schooling	Child-Rearing Methods	Child's Educational Aspirations	Child's Self-Confidence
Son Preference	-0.760 (0.115)***	—	-0.295 (0.105)***	-0.196 (0.108)*
Girl	-0.264 (0.058)***	—	-0.153 (0.054)***	0.033 (0.055)
Number of Siblings	0 (0.041)	—	-0.077 (0.037)**	-0.052 (0.037)
Father's Education (Years of Schooling)	0.022 (0.008)***	—	0.023 (0.008)***	0.013 (0.008)*
Mother's Education (Years of Schooling)	0.019 (0.010)**	—	0.009 (0.009)	-0.008 (0.009)
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very	0.118 (0.039)***	—	0.107 (0.036)***	0.005 (0.037)
Intact Family	-0.240 (0.235)	—	0.047 (0.211)	-0.276 (0.227)
Observations	1,895		1,895	1,895
R-squared	0.09		0.04	0.04

*Notes:* \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level. Standard error in parenthesis. The data source is GSCF 2000 Survey. There are 2000 observations in the dataset. After deleting all the missing values, there are 1,895 observations for OLS Model and Ordered Probit Model. Please refer to Table 1 for the definitions of

**Table 6. Impact of Son Preference on Educational Attainments With and Without Mechanism Variables**

Variable	Prior Chinese Grade		Prior Mathematics Grade		Grade Retention	
	Without Control	With Control	Without Control	With Control	Without Control	With Control
Son Preference	-3.925 (1.786)**	-2.759 (1.780)	-3.794 (1.901)**	-2.341 (1.889)	0.374 (0.190)**	0.308 (0.191)
Son Preference*Girl	2.388 (2.513)	2.361 (2.503)	2.565 (2.676)	2.466 (2.655)	-0.386 (0.269)	-0.406 (0.272)
Girl	1.413 (1.462)	1.769 (1.456)	-0.366 (1.557)	0.135 (1.545)	0.172 (0.163)	0.154 (0.165)
Number of Siblings	0.009 (0.474)	0.153 (0.470)	-0.299 (0.504)	-0.122 (0.499)	0.020 (0.048)	0.012 (0.049)
Father's Education (Years of Schooling)	0.280 (0.097)***	0.217 (0.097)**	0.278 (0.103)***	0.202 (0.103)**	-0.024 (0.011)**	-0.020 (0.011)*
Mother's Education (Years of Schooling)	0.218 (0.114)*	0.184 (0.114)	0.148 (0.121)	0.108 (0.121)	-0.029 (0.013)**	-0.028 (0.013)**
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very Bad)	1.248 (0.459)***	0.926 (0.460)**	1.189 (0.489)**	0.804 (0.488)*	-0.049 (0.049)	-0.031 (0.049)
Intact Family	-1.169 (2.792)	-1.053 (2.769)	0.239 (2.973)	0.376 (2.937)	0.399 (0.310)	0.401 (0.315)
<i>Possible Channels of the Impact of Son Preference on Children's Educational Attainments</i>						
Child-Rearing Methods		2.284 (1.380)*		2.301 (1.464)		-0.148 (0.147)
Mother's Expectations for Child's Schooling		1.003 (0.505)**		1.291 (0.536)**		-0.095 (0.052)*
Child's Educational Aspirations		1.020 (0.218)***		1.302 (0.231)***		-0.054 (0.023)**
Child's Self-Confidence		0.789 (0.451)*		0.854 (0.478)*		-0.042 (0.048)
Observations	1,895	1,895	1,895	1,895	1,838	1,838
R-squared	0.15	0.17	0.12	0.15	0.17	0.17

Notes: \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level. Standard error in parenthesis. The data source is GSCF 2000 Survey. There are 2000 observations in the dataset. After deleting all the missing values, there are 1895 observations for OLS Model (Column 1-4) and 1838 observations for Probit Model (Column 5,6). Column 1,3,5 are models before adding the possible channels of the impact of son preference on children's educational attainments, which is identical to Table 4, Panel B. Column 2,4,6 are models after adding those variables. The R-squared in Column 5,6 are Pseudo R-squared. Please refer to Table 1 for the definitions of all the variables.

**Table 7. Impact of Son Preference on Educational Attainments before and after Adding the Mechanism Variables (Son Preference'=0 if Son Preference<0.625; Son Preference'=1 if Son Preference>=0.625)**

Variable	Prior Chinese Grade		Prior Mathematics Grade		Grade Retention	
	Without Control	With Control	Without Control	With Control	Without Control	With Control
Son Preference <sup>a</sup>	-2.008 (0.970)**	-1.538 (0.964)	-1.777 (1.032)*	-1.190 (1.023)	0.104 (0.102)	0.076 (0.102)
Son Preference*Girl	1.507 (1.403)	1.629 (1.394)	1.650 (1.494)	1.766 (1.479)	-0.088 (0.147)	-0.106 (0.148)
Girl	2.099 (0.854)**	2.399 (0.852)**	0.360 (0.910)	0.771 (0.904)	-0.001 (0.094)	-0.024 (0.095)
Number of Siblings	-0.001 (0.474)	0.147 (0.470)	-0.307 (0.505)	-0.126 (0.499)	0.021 (0.048)	0.013 (0.049)
Father's Education (Years of Schooling)	0.286 (0.097)**	0.221 (0.097)**	0.284 (0.103)**	0.206 (0.103)**	-0.024 (0.011)**	-0.020 (0.011)*
Mother's Education (Years of Schooling)	0.229 (0.114)**	0.190 (0.113)*	0.162 (0.121)	0.115 (0.12)	-0.031 (0.012)**	-0.029 (0.013)**
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very Bad)	1.245 (0.459)**	0.917 (0.460)**	1.179 (0.489)**	0.787 (0.489)	-0.049 (0.049)	-0.030 (0.049)
Intact Family	-1.185 (2.793)	-1.049 (2.769)	0.240 (2.975)	0.401 (2.938)	0.395 (0.309)	0.397 (0.314)
<i>Possible Channels of the Impact of Son Preference on Children's Educational Attainments</i>						
Child-Rearing Methods		2.297 (1.379)*		2.332 (1.463)		-0.143 (0.147)
Mother's Expectations for Child's Schooling		1.024 (0.503)**		1.321 (0.534)**		-0.096 (0.052)*
Child's Educational Aspirations		1.026 (0.218)**		1.308 (0.231)**		-0.055 (0.023)**
Child's Self-Confidence		0.805 (0.451)*		0.874 (0.479)*		-0.043 (0.048)
Observations	1,895	1,895	1,895	1,895	1,838	1,838
R-squared	0.15	0.17	0.12	0.15	0.17	0.17

Notes: \*\*\* Significant at the 1% level.\*\* Significant at the 5% level.\* Significant at the 10% level.Standard error in parenthesis. <sup>a</sup>Son Preference' is defined as 1 if Son Preference>=0.625, 0 otherwise.

**Table 8. Impact of Son Preference on Educational Attainments before and after Adding the Mechanism Variables (Son Preference'' =2 if Son Preference ∈ [0.625,2], 1 if Son Preference ∈ [0.5,0.625), 0 if Son Preference ∈ [0,0.5))**

Variable	Prior Chinese Grade		Prior Mathematics Grade		Grade Retention	
	Without Control	With Control	Without Control	With Control	Without Control	With Control
Son Preference <sup>nd</sup>	-1.142 (0.559)**	-0.816 (0.557)	-0.993 (0.595)*	-0.587 (0.591)	0.085 (0.061)	0.064 (0.061)
Son Preference <sup>nd</sup> *Girl	0.771 (0.800)	0.877 (0.795)	0.694 (0.852)	0.805 (0.844)	-0.067 (0.087)	-0.080 (0.088)
Girl	1.858 (1.064)*	2.097 (1.059)**	0.248 (1.133)	0.592 (1.124)	0.038 (0.121)	0.023 (0.122)
Number of Siblings	0.004 (0.474)	0.154 (0.470)	-0.303 (0.505)	-0.119 (0.499)	0.020 (0.048)	0.012 (0.049)
Father's Education (Years of Schooling)	0.282 (0.097)***	0.219 (0.097)**	0.280 (0.104)***	0.204 (0.103)**	-0.024 (0.011)**	-0.020 (0.011)*
Mother's Education (Years of Schooling)	0.222 (0.114)*	0.188 (0.114)*	0.155 (0.121)	0.114 (0.121)	-0.030 (0.012)**	-0.028 (0.013)**
Economic Situation (3=Very Good; 2 Above Average; 1 Below Average; 0 Very Bad)	1.255 (0.459)***	0.925 (0.461)**	1.194 (0.489)**	0.799 (0.489)	-0.050 (0.049)	-0.031 (0.049)
Intact Family	-1.215 (2.793)	-1.075 (2.769)	0.202 (2.975)	0.367 (2.938)	0.397 (0.309)	0.398 (0.314)
<i>Possible Channels of the Impact of Son Preference on Children's Educational Attainments</i>						
Child-Rearing Methods		2.290 (1.380)*		2.313 (1.464)		-0.143 (0.147)
Mother's Expectations for Child's Schooling		1.032 (0.504)**		1.320 (0.535)**		-0.095 (0.052)*
Child's Educational Aspirations		1.023 (0.218)***		1.308 (0.231)***		-0.055 (0.023)**
Child's Self-Confidence		0.807 (0.451)*		0.871 (0.479)*		-0.043 (0.048)
Observations	1,895	1,895	1,895	1,895	1,838	1,838
R-squared	0.15	0.17	0.12	0.15	0.17	0.17

Notes: \*\*\* Significant at the 1% level. \*\* Significant at the 5% level. \* Significant at the 10% level. Standard error in parenthesis.

<sup>a</sup>Son Preference'' is defined as 2 if Son Preference ∈ [0.625,2], 1 if Son Preference ∈ [0.5,0.625), 0 if Son Preference ∈ [0,0.5).