Child Gender, Parental Labor Outcomes, and Investment in Childhood: Evidence from China

Qing Wang*

University of Illinois at Urbana-Champaign

Draft date: February, 2013

Preliminary draft: Please do not circulate beyond this conference.

Abstract

This paper provides new evidence of the effects of child gender on parental labor supply and earnings in a developing country. Using the longitudinal data from the China Health and Nutrition Survey (CHNS), I document that parents with a first-born son have a 10% higher labor supply than parents with a first-born daughter. Cross sectional analysis indicates a 6% earning premia corresponding to having a son than having a daughter. The "son premia" are in line with previous findings in developed countries like the U.S. and Germany. A further analysis by children's age group indicates larger "son premia" for families with preschool-age children. Meanwhile there is almost no difference in parental labor outcomes between families with a son aged 7-19 and those with a daughter of the same age, which seems puzzling. Given the prevalence of son preference and informal childcare in China, I formulate a new hypothesis to explain this type of child-age-dependent "son premia": Paternal grandparents are more willing to provide caregiving for a grandson than granddaughter, which helps parents with a preschool-age son come back to work faster. Both cross sectional analysis and instrumental variable approach find supportive evidence for child-gender-based intergenerational transfer of childcare. The intensification of market work associated with having a son induces differential parental investments of time and money to boys and girls, which may influence children's early development as well as the economic well-beings in the long run.

Key words: Labor supply, Earnings, Child Gender, Child Care, Parental Investment

JEL Classification: J1, J2, D1

* Department of Economics, University of Illinois, 214 David Kinley Hall, 1407 W Gregory Drive, Urbana, IL 61801; Phone: (608) 658-2419; Email: wang155@illinois.edu. I am grateful to Elizabeth T. Powers for extensive advice and continuous support. All errors are my own.

1. Introduction

Both economic theory and empirical work document that gender inequality widely exists in intra-family resources allocation to boys and girls. Previous work has found evidence on the disadvantage of girls on formal education, nutrition intake and health care. However, few theoretical and empirical studies have explored how the gender of offspring influences parental economic well-being like labor outcomes. There are at least two major inadequacies among the existing studies: research is very limited and indirectly documented using data from a developing country where the son preference is strong and prevailing; and the underlying reasons that generate such child gender bias on parent's labor outcomes are not clear and lack of evidence. This paper attempts to fill this gap in the literature using a longitudinal data set of the China Health and Nutrition Survey (CHNS).

Understanding the intergenerational effect of child gender on parental labor outcomes is of substantial interest to researchers and policymakers. The difference in time allocation of parents on market work and resulted earnings gap indicate a distributional effect of child gender. It not only determines the living standard of parents but may intensify potential gender bias on child's development through the human capital production function. The research hence reveals an alarming and unintended consequence of the culture of son preference and compelling evidence on gender inequality even at the very early age of children's life. Proper policies should be implemented to improve gender equality.

To address the heterogeneity of families, I explore the identification strategy based on the random assumption of gender of the first child. The randomness of sex at birth has been used by economists to provide an exogenous variation (Lundberg, 2005). Although, previous findings have shown that the male to female ratios at birth strongly increase with the birth order which leads to the gender of a child at higher birth parities endogenous, gender of the first child is arguably random and leaves cross sectional analysis advantageous in which causality can be established. This strategy has been widely used in, for example, Dahl and Moretti (2008), Bogan (2010), Ebenstein and Leung (2010), Ebenstein (2010) and Li and Wu (2010). I also construct a test and find that the gender of the first child is not significantly correlated with individual and family characteristics in the empirical analysis.

To further address the endogeneity, I examine the differential responses of parents to the birth of a son and the birth of a daughter by tracking the change of parental labor outcomes from the prenatal periods to postnatal periods using fixed effects approach. This method has been proven to correct the omitted variable biases of OLS estimates if the omitted variables are some unobserved individual or family time invariant characteristics like personality, taste and intrinsic ability in the literature of marriage, fertility and labor supply, e.g., Korenman and Neumark (1991, 1992) and Lundberg and Rose (2002).

To preview the findings, both OLS and fixed effects estimates provide evidence of gender bias. First, the cross sectional analysis indicates that having a first born son increases parent's annual wage income by 6%, weekly working time by 0.8 hour and annual working time by 36 hours than having a first born daughter. The direction and magnitude of the estimated effects of child gender on parent's labor supply are similar to the findings in Lundberg and Rose (2002) and Choi *et al* (2005), which suggests the son premia found in developed countries also exist in a developing country like China. The estimated effect of child gender on total yearly wage earnings is higher than their estimates but similar to the findings in Knight et al (2010). In addition, I find the results are driven by parents with young children between 0 and 6 years old. The effects of child gender on parental labor outcomes are not significantly different from zero among families having an older (7-19-year old) child.

Estimates from fixed effect approach are in general consistent with the OLS estimates. They indicate that the birth of a child decreases parental, in particular father's labor supply by near 4 hours per week and 195 hours per year, which is equivalent to a 10-12% "penalty" for people with full-time jobs. However, parent's labor supply tends to increase by 4-5 hours per week and 220 hours per year when the newborn is a son rather than a daughter. This positive effect offsets the negative effect of having any child and lead to no interruption of parent's labor outcomes corresponding to the birth of a son. The estimated coefficient of child gender on parental wage earnings appears large but not significantly different from zero using the standard confidence levels. It may suggest that during the few years after the birth of the child, child gender does not significantly influence wage earnings of parents.

The findings on the dependence of gender bias on child's age are striking and puzzling, since the dependence of gender bias on child's age is not consistent with traditional theories of preference

for sons over daughters, differential costs of raising a son over a daughter and the fatherhood's role model. These theories all predict the effect of child gender should be persistent as the child grows up. I then propose a new hypothesis of child care arrangement to interpret the findings of child gender bias in this paper. This interpretation seems being ignored in previous work to the best knowledge of mine.

Before children reach school age, parents have a high burden of childcare. A great part of work has examined the relationship between female labor force participation and the cost of day care and concluded that the availability and affordability of day care are essential to stimulate the labor participation of females with a young child. But it seems not much research discusses how childcare influences father's labor outcomes and how the childcare choice varies with child gender.

Different from the U.S. and other developed countries, the formal child care provided by a kindergarten or a preschool in China is short of supply due to the economic transition and reforms of state owned enterprises. Du and Dong (2010) estimated that the number of kindergartens decreased by 28.5% between 1997 and 2006. For this reason, the informal child care provided by grandparents is an important alternate to the formal child care in China. Even in areas that childcare facilities are well developed, grandparents can also provide essential help on caregiving when the child is sick or when the daycare is closed during the day time and evening time. Maurer-Fazio et al (2010) and Chen et al (2000) both point to the importance of taking into account the intergenerational kinship ties that extend beyond the household boundary. Their findings emphasize the role of grandparents as caregivers and the strong legacy of a patrilineal culture.

Using cross sectional analysis and instrumental variable approach, I find that parents with a preschool-age son are more likely to obtain help on caregiving from grandparents than parents with a preschool-age daughter. This is probably driven by the stronger preference of male offspring over females from grandparents. All this evidence suggests that the intergenerational transfer of childcare is one of the possible reasons that help parents come back to work sooner after the birth of a son than the birth of a daughter.

An important consequence of the gender bias on parent's time allocation on market labor is parental investment of time and money on their children. Cross sectional estimates indicates parents with a boy spend 2.5 - 3 fewer hours on child care related activities than parents with a daughter of similar age. The gender effect on child care is larger for mothers than fathers. Parents with a boy are both more likely (10%) to choose formal child care and spend more money on formal child care when they have a boy than parents with a girl. Such differential input of time and money from parents may lead to different development between young boys and girls.

In this paper, I contribute to the existing literature in several ways. First, this research extends the previous work on child gender and parental labor outcomes by making a systematic analysis using data from a developing country. Second, by exploring the panel nature of the data set I am able to more consistently estimate the effect of child gender on labor outcomes. Third, from a policy perspective, this paper sheds light on the importance of family ties and culture for economic outcomes. The fact that sons induce parents to work harder and earn more than daughters may suggest a distributional effect of the gender of offspring. Meanwhile, girls may be disadvantaged in both short term and long term development by the gender-based parental investments since previous work has shown that children who received center-based care have higher cognitive ability and social skills than children who receive informal child care (Zhai and Gao, 2010). Proper policies should be implemented to improve the social economic status of girls.

The remainder of the paper is organized as follows. Section 2 presents the theoretical background and literatures. Section 3 introduces the identification strategy. Section 4 describes data, sample and variable construction. Section 5 discusses findings and interprets the results. Section 6 presents additional empirical evidence and interprets. Section 7 concludes with robustness checks and extensions.

2. Background and Literature

The magnitude and direction of any effect of child gender on labor market outcomes is an empirical question (Pabiloni and Ward-Batts, 2007). In the literature, fathers with a son have been found to work longer and harder in some developed countries like the U.S. and Germany. In a leading work in 2002 from Lundberg and Rose, they show that each son increases his

father's labor supply by 40 hours per year and hourly wage by 3% more than each daughter using data from the Panel Study of Income Dynamics. In a similar work, Choi *et al* (2005) uses data from German Socio-economic Panel and found that a first born son increases his father's labor supply by 107 hours per year more than a first born daughter. The premia on paternal working hours and earnings corresponding to having a son is thus called "son premia".

Pabiloni and Ward-Batts (2007) use data from Census and Current Population Survey (CPS) to further examine parental labor supply responses to child gender in the U.S. By allowing for differences in the response to child gender among parents who are native born versus first-generation or second-generation immigrants, or a member of various race and ethnic groups, they find that son preference is persistent among Asia immigrants but they work fewer hours when they have a son rather than a daughter than White natives.

Another branch of literature attempts to find the underlying incentive reason of this gender effect of offspring. First, having a son increases marriage duration (Dahl and Moretti, 2008) which leads couples more likely to make gender specialization. Traditional fathers with boys then are predicted to work harder and longer to get wage premia and promotion than fathers with girls[1]. Lundberg (2005) studies a recent cohort in the National Longitudinal Survey of Youth (NLSY) and finds that a son reduces gender division of labor for more educated couples, but increases specialization of less educated couples in the 3 years following the birth of a child.

Second, a demonstration effect indicates that fathers who express traditional gender ideologies may believe it is more important to model the traditional male role of the breadwinner for their sons than for daughters. Glauber and Gozjolko (2011) use data from the NLSY 1979 and find that fatherhood was associated with an increase of time spent on paid work among White men. The increase was more than twice as strong for traditional White men than for egalitarian White men.

Third, having a son increases the financial distress of a family. Parents may plan to invest more on sons' education than daughters' when they believe the returns for educating their sons are relatively higher. This more likely happens in a developing country where gender discrimination

[1] The assumption is females have the comparative advantage in domestic tasks but disadvantage in market tasks due to their productivity advantage in household activities and the gender wage gap.

may exist in the workplace. I find no work that directly tests this hypothesis, although a great deal of work documents the discrimination on girls of formal education in families with the financial constraint in countries underdeveloped. But Edlund et al (2010) finds an increase of education attainment of males relative to females in areas with high male to female ratios which likely relates to the premarital education investments in China.

Recent evidence from China focuses on another type of financial distress which responses to higher male to female ratios (measured by the number of males per 100 females), which is an importance difference between China and countries like the U.S. Research on China showed this ratio has deviated from the natural level (103-106) and kept growing in recent two decades. The 2000 Census reflected a sex ratio at birth of 119. This number implies near 1 million more males are born than females each year, which is known as the phenomenon of "missing girls".

Work from Wei and Zhang reveals that parents with a son spend more hours on off-farm work in rural areas and are more likely to undertake a dangerous job corresponding to higher male to female ratios (2011b). In a work similar to this paper, Knight et al (2010) provide evidence a positive correlation between having a son and earned family income in rural China and conclude that the incentive effect is likely the reason that generates the gender bias, i.e., parents with a son rather than a daughter have an incentive to increase family income to support their sons. However, all their work only studies people living in rural China and have no direct test of the gender effects on parental labor supply or wage earnings.

3. Methodology

3.1 Cross sectional analysis based on the random assumption of child gender.

Although gender of a child is predetermined and not affected by parental labor outcomes several years after the birth, it could be the outcome of sex selection when the demand for son is strong. Some unobserved individual characteristics, like ability and personality (e.g., self-esteem and ambition) may determine both the likelihood of having a son and labor supply and earnings. I addition, family characteristics that determine fertility choices may also affect parental labor outcomes directly. For example, traditional patrilineal families may be both likely to achieve a

son and make gender division of labor between husband and wife. In this case, the positive effect of having a son and male labor supply would be spurious.

To address the potential endogeneity of sex selection, I use gender of the first child as the main variable on the right-hand-side in empirical analysis. To ensure that I compare apples with apples, I follow the method in Knight et al (2010) and Wei and Zhang (2011) and categorize respondents into families with only one-child and families with at least child (mainly one or two children). Age of the child in one-child families and age of the first child in multiple-child families is restricted to be less than 19.

Although the male to female ratios at birth strongly increase with the birth order which leads to the gender of a child at higher birth parities endogenous, gender of the first child is arguably random in both China and other countries. This leaves cross sectional analysis advantageous in which causality can be established. Figure 1 is from Ebenstein and Leung (2010) who used six waves of China Census from 1982 to 2000 to plot the male fractions at birth by parities with a vertical line indicates the year of initiation of the one-child policy. Male fractions of the second and third birth strictly increase over year but the fraction of the first birth is quite stable and close to the natural level.

By assuming there is no systematic sex selection at the first birth, I use an empirical model that has been previously applied to Lundberg (2005), Dahl and Moretti (2008), Bogan (2009), Ebenstein and Leung (2010), Ebenstein (2010) and Li and Wu (2010).

$$Y_i = \alpha + \beta_1 \times first\ born\ son_i + \gamma X_i + \varepsilon_i \quad (1)$$

In Equation (1) Y_i is individual parent i's labor outcomes at a survey year[2]. I use five variables to measure the labor outcomes in this paper which are yearly wage income, hourly wage, working hours the week prior to the interview, average weekly hours worked last year and total annual hours last year.

$first\ born\ son_i$ is a dummy variable indicating the first child born to a parent is a boy. X_i contains child's demographics (mainly age), parent's demographics including gender, age,

[2] In this part, I pool observations from different waves of survey in one sample and treat it as a repeated cross sectional data set. As longitudinal files, CHNS contains multiple years of data for part of respondents but not all of them due to the problems of attrition and missing values.

education, occupation, province where she lives, year of interview and resident location of city, suburb, town and village. To further capture the local economic conditions which are important determinants of individual labor outcomes, I include variables that measure the local development of transportation facilities, health, social services, population density, and economic activity at the community level. And finally, ε_i is an error term.

A majority of married males (i.e., 89%) are employed. Meanwhile labor force participation rate of females in China is as high as 73% in the sample periods. For this reason, I put my focus on active workers first. The labor force participation, in particular of females is discussed in the section of robustness checks. Alternative specifications of Tobit models for work hours including inactive workers yield similar patterns thus are not reported.

3.2 Fixed effect estimates for the subgroup parents with prenatal information.

A subgroup of couples was interviewed before they had any child and after the child was born. For this group of individuals, I can observe their labor outcomes before and after the birth of a child. I then construct the following empirical model to test how different a birth of a boy is from a birth of a girl to parents.

$$y_{it} = \alpha + \beta_1 \times Having\ any\ child_{it} + \beta_2 \times Having\ any\ boy_{it} + \gamma X_{it} + \theta_i + \varepsilon_{it} \quad (2)$$

In Equation (2), the variable of $Having\ any\ child_{it}$ shows if parent i has any child at the survey time of t, which equals to 0 before the only child was born and 1 after. $Having\ any\ boy_{it}$ is a variable indicating if the newborn was a boy.

Using this method I examine the within-parent change of labor outcomes corresponding to the birth of a child. β_1 captures the effect of the birth of a girl on parent's labor outcomes. $\beta_1 + \beta_2$ shows the effect of the birth of a boy on parent's labor outcomes. The main hypothesis to test is $\beta_2 = 0$, that is in the short period after the child is born, the gender of offspring does not influence the behavior of parents in labor market. θ_i is a vector of individual unobserved factors that are not changing over time. Definitions of y_{it}, X_{it} and ε_{it} are similar to those in equation (1).

The estimates via an individual fixed effects model may still suffer from omitted variable bias when the unobserved factors that correlated with both labor outcomes and fertility change over time. For instance, parents are likely to get married and have a child when they expect to receive

or have received a promotion or have a higher growth rate of wage. In that case, β_1 is biased upwards. However, β_2 is unbiased as long as these time varying disturbances do not determine the propensity of making sex selection and the probability of having a newborn son. In this paper, I use the fixed effect model to further correct the potential heterogeneity of parents and families.

4. Data, Sample and Variables

Data used in this paper is from the China Health and Nutrition Survey (CHNS) conducted by the Population Center at the University of Carolina. It is a longitudinal study that covers approximately 16,000 individuals originating from 4,400 households living in 9 provinces in the years 1989, 1991, 1993, 1997, 2000, 2004, 2006 2009. Nine provinces, Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong participate in this survey. Figure 2 plots the geographic locations of these provinces in China. They come from the coastal, middle, northeastern and western areas and vary substantially in economic development, public resources, and health indicators. Samples were drawn from each province following a multistage, random cluster design. Counties were stratified into three levels of income, and a weighted sampling technique randomly selected four counties in each province (Wang, 2012). The questions in 1989 survey are in general different from questions in the remaining waves so I take it off from the sample of analysis.

I removed parents whose age is younger than 22 or older than 50 at the time of survey to assure the respondents are at the working and marriageable age. I remove families in which the eldest child is beyond 19 because it is not desirable to discuss parental behavior when the child has grown up and begins to contribute to family income. I end up with a sample of over 5,000 observations from fathers and mothers who are between 22 and 50, married, take some economic activity, graduated from school and have the only child younger than 20 years old living with them[3].

For a part of individuals in this cross sectional sample, I can observe the change of their family size over time. I use this information as well as the demographics and labor outcomes to

[3] A potential pitfall is most conclusions in this paper only apply to three person families and may not be generalized to families with multiple children. This may not be a serious concern since a great part of parents in China only have one child due to the limitation of fertility.

construct a subsample for the analysis using the fixed effects approach. I identify the prenatal and postnatal periods of the birth of a child using the information of if any child was born to the family between two consecutive waves of survey. For all new members in old households, there is question asking the reason that the new member joins the household. I define the family has an additional child if the new member is recorded as a newborn. I further identify the gender of a child who newly enters the household. Using this method, I find near 2,200 children who were born to 1,600 mothers and 1,500 fathers[4] between the survey periods from 1991 to 2009. To maximize the use of the prenatal information when it is ever available, I require the individuals no younger than 18 years old (instead of 22) in the prenatal periods. Near 40% of respondents are single when they have no child living with them.

For each work age individual in the sample of fixed effect analysis, I require she has at least one observation in the prenatal period and one observation in the postnatal period. The final sample is then consisted of 1,330 fathers and mothers of whom the demographics and employment status are available[5]. The remainder of this part describes the constructions of variables used in the empirical analysis.

The variables of labor time are constructed according to the information of average working hours per day, working days per week and the number of working months last year. I construct the average weekly working hours by multiplying the working hours per day with the days per week. Since there are no questions asking the weeks worked per month, I assume all respondents work four weeks per month and define the yearly working hours as 4 times of the weekly hours by working month. The income variables are balanced between rural and urban residents. The working time variables, however, generate many missing values for farmers. The reason is farmers may not have a good self-measure of their laboring time. For this reason, I restrict the analysis to individuals who are not mainly taking farming activities when exam their labor supply.

There is direct question on the individual monthly salary from the original survey if the person is employed. To proxy the efficiency of work, I define the hourly wage as the ratio of monthly

[4] For a part of newborns, there is no record of mother's id or father's id when children and parents live in two households.

[5] 1,100 fathers and mothers of whom the individual annual income is available and 600 fathers and mothers of whom all labor outcomes are available in both prenatal and postnatal periods.

salary divided by 4 times of the weekly hours worked. The income variables used in the analysis is CPI-adjusted to remove the trend of inflation. For all income related variables, I take the log transformation after adding 1 to the real value in order to avoid dropping the zeros. I trim the top 1% and bottom 1% respondents in the personal income distribution to remove the influence of outliers.

The gender and the number of offspring used in cross sectional analysis are from the self-reported fertility history. The CHNS contains information of fertility from the ever-married women who are under 52. Variables which are constructed based on this fertility information are more reliable than using the roster file because the roster file usually does not track the children who are not living with their parents.

Information of child care is from the modules of child, adult and family survey. For families with preschool-age children younger than 7, CHNS asked each adult how much time she spends on child care by feeding, bathing, dressing, holding and watching them during the week prior to the interview. Time used to cooking, washing clothes while caring for the child is also included. I use this information to construct father and mother's time spent on caring for young children.

In additional to direct child care from parents, parents may find substitutes of caregivers when they need to go back to the workplace. The alternative caregivers are in general grandparents and day care facilities. I construct a discrete variable indicating the sources of non-parent's child care which are none, paternal grandparents, maternal grandparents and day care. I define the child is not cared by any caregiver other than the parents if her parent has negative answer to the question of "the child was cared by a person who did not live in the household". There are two types of families in this group: nuclear families without grandparents and extended families in which no grandparent's caring for granddaughter is recorded. I define the grandparent offers some help caring for a grandchild if parents confirmed that the child is cared at a grandparent's house or grandparents confirmed they spent some time on child care in co-residing families. I further distinguish grandparent's care from father's side and mother's side. At last, I define a child is cared in a day care if the care took place in a state child care center, a care center run by a work unit, a preschool and a nursery school. Because childcare can take place in multiple locations, these categories are not mutually exclusive.

For families in which grandparents coreside with grandchildren I can further observe the number of hours that grandparent care for grandchildren. For families who have some non-household members as caregivers, I can observe their monthly expenditures spent on childcare.

Variables of age and the family member type are directly recoded from the roster file. Individuals in the final sample are household heads and their spouses as well as children and children-in-law. I do not remove the adults who co-reside with their parents because the extended families are common in China in particular in rural areas and children are likely to co-reside with parents before they are married.

The variable of education is recoded from the question of years of formal education completed in school. The set of occupation dummies are used to proxy the detailed type of individual's primary job which is categorized into administrators and ordinary workers. Province dummies are used to capture the provincial level fixed effect since there is great variation of sex ratios and economic development across provinces in China. Finally, resident location dummies indicate if the individual lives in urban (city and suburb) or rural (town and village) areas.

Variables that measure local transportation, population, health, economics and social services are indexed at the community level in Jones-Smith and Popkin (2010). Transportation facilities are transformed from information of the most common type of road, distance to bus stop and distance to a train station. Population density is recoded from total population of the community divided by community area from official records. Health condition is recoded from the information of number and type of health facilities in or nearby the community and number of pharmacies in community. Economic activity is transformed from the typical daily wage for ordinary male worker and the percentage of population engaged in nonagricultural work reported by community officials. Social services are recoded from provision of preschool for children under 3 years old, availability of commercial medical insurance, free medical insurance, and/or insurance for women and children.

Table 1 presents the summary statistics of the cross sectional sample with the top panel showing demographics and the bottom panel showing labor outcomes. The left panel summarizes families with only one child and the right one describes families with at least one child. Taking the one-child family as an example, 54% respondents in the sample are males. The average age of

individuals in the sample is 35. The sample is consisted of 38% of household head, 32% spouses, 17% children and 13% children-in-law. The relative lower representation of females in the sample is due to the limitation of active workers. 54% of them have a first born son which implies a ratio of 117 between families with a first born son and families with a first born daughter. The age of the only child averages 9.5. 10% of the sample completed secondary school, 49% completed high school and 35% have a college degree. 19% work as administrators, managers and senior professional workers (e.g. doctor, professor and others) and 81% are ordinary workers. Indexes of transportation, social service, health, population and economic activity average 6.4, 1.8, 6.6, 6.4 and 6.2 in the sample. The numbers of sample obtained from different survey years are gradually decreasing over time which might be due to the aging of those initial respondents in earlier waves of the survey. Finally, 32% respondents live in the city, 22% in the suburb, 28% from the town and 19% from the village. The remaining rows in this panel show the distribution of samples drawn from different provinces. Panel B indicates that fathers and mothers average working 45-46 hours per week, which sum up to 2116 hours per year. The log transformation of hourly wage is 1.4 and total annual wage income is 8.6. It is noted that the calculated sex ratio of children among families with at least one child is quite balanced (= 0.5).

Gender of the first child is random as long as parents do not abort girls at their first childbearing. Table 2 presents the results from regression of the likelihood of having a first born son on parental and family characteristics with a similar structure to Table 1. This probit model checks if the gender of the first birth is truly exogenous. Columns (1) – (3) refer to the samples of one-child family with a 0-19, 7-19, and 0-6 years old child and the remaining columns present families with at least one child. Parent's birth cohort appears an important determinant of the probability of having a first born son in the first two columns but not parents with a young child. In none of the cases, the probability of having a first born son is correlated with parent's formal years of schooling.

5. Preliminary Results

5.1 Evidence of son premia in China: OLS findings

Table 3 presents the regression results in which each column is from a separate regression with the name of the left-hand-variable shown as the column title. The reported coefficients and standard deviations from Panel A to Panel D are estimated from four sets of different constraints of the sample. The top panel uses all work age fathers and mothers with the only child younger than 20. Panel B restricts the sample to fathers and mothers with the only child from 7-19. Panel C restricts to fathers and mothers with the only child younger than 7 years old. And Panel D refers to parents with at least one child and the first child younger than 7 years old. I also examine another sample specification by removing the assumptions of non-agricultural workers. Results are essentially the same thus not reported in the table.

The most striking findings are having a boy significantly increase the parent's working hours and earnings. Take the estimates in Panel A as an example, parents whose first child is a son work an additional 0.8 hour per week, 33 hours per year and earn 6% more income than parents whose first child is a daughter. The findings suggest that parents of a boy achieve the "income premia" by increasing their labor supply.

Moreover, it is more striking to see an important age trend of this gender bias when the effect of sex is allowed to change over time. A comparison across different sample specifications indicates that the effect of gender of the first born is most severe among parents with the only child younger than 7. For parents with a relative older child (7-19-year old), the estimated bias of child gender is not different from zero in any of the five measures of labor outcomes according to Panel B. This suggests that parents with an older son do not in general behave differently from parents with an older daughter. Panel C indicates that the effect of child gender on parental labor outcomes is much stronger than the child is younger (0-6-year old). For instance, parents tend to work 1.5-1.6 hours per week and 77 hours per year when the first child is a son rather than a daughter. Meanwhile, parents with a young son earn 16% more for annual wage than parents with a young daughter.

The estimated coefficients and standard deviations of other covariates are presented in the Table Appendix[6] using families with a 0-6-year old child. Individuals increase labor supply over age but not monotonically. Men work longer hours than females (1.6-2.0 weekly hours and 109 annual hours) and earn 37% more earnings. The effect of education on labor outcomes is not linear. Individuals who completed secondary school and high school work more and earn more income than individuals with only elementary diploma. Individuals with college degree earn over 50% more than individuals with little schooling which is likely due to a much higher hourly wage but not more labor supply. Ordinary workers earn a lower hourly wage than administrators but do not have different patterns of hours worked. The role of family types does not significantly influence labor outcomes. The index of economic activity is negatively correlated with individual labor supply since it may proxy the competition within a community. The estimated results of geographic locations and year dummies are in general consistent with prediction.

5.2 Cross sectional estimates by gender

A separate examination of the effect of offspring by the sample of fathers and mothers is reported in Table 4 with the first five columns showing results for mothers and the remaining columns for results for fathers. Table 4 also reveals a strong dependency of the effects of child gender on child's age. For example, none of the estimated coefficients for the five measures of labor outcomes is significantly different from zero in the sample of mothers or fathers from one-child and multiple-child families with the first child older than 6 years old. For mothers from the with a young son under 7, they average 2 more hours on working and have a 20% annual earnings premia than mothers with a young daughter. Similarly, fathers with a young son also earn premia of working hours compared to fathers with a daughter. The estimated earnings premia are large (10%) but not significantly different from zero. The "son premia" associated with a young child are even stronger in multiple-child families with the first child younger than 7. The benchmark result in Table 4 is both mothers and fathers behave according to the gender of their child. Mothers may obtain higher premia when they have a son not a daughter.

[6] The table appendix is available upon request but not included in this paper due to space limitation.

5.3 Cross sectional estimates at the family level

Table 4 also predicts that the total working time of a family as well as earnings should be higher associated with having a son than having a girl. This speculation is supported by the left panel in Table 5. From Column 1 to 5, having a son increases the total family working time by 1.6-2.9 hours per week and 75-170 hours per year. The higher intensity of market work leads to 30-40% annual wage premia. Meanwhile, similar to previous findings, such premia only exit in families with a young child and are stronger in families with multiple children.

The right panel of Table 5 shows little evidence of larger gender specialization within couples. For a great part of the regressions, the estimated differential outcome between husband and wife is not significantly different from zero. In other words, parents with a son are found to be more market-oriented but not more specialized than parents with a daughter.

5.4 Evidence of son premia in China: fixed effects findings

Table 6 presents the estimated results from a fixed effect model. The reported coefficients and standard deviations in Panel A to Panel D are estimated from four sets of different constraints of the sample. The top panel tracks fertility change of fathers and mothers from having no child to one child. Panel B is similar to Panel A but only tracks periods before the child reaches 7 years old. Panel C and D are parallel to A and B except that they include families with multiple children.

This table reveals first that the direction and magnitude of the estimates are reasonably consistent from each of the four sample specifications. Therefore, the following narratives use the estimated number from the Panel A as an example.

The estimated effects of having an additional child and an additional male child are consistent with model predictions. The baseline finding in this table is that the birth of a child decreases parental labor supply by 2 hours the week prior to the interview, by 4 hours per week last year on average and 195 hours in total last year. The influence of having a child on total wage earnings is estimated to be very small and insignificantly. Moreover, parent's labor supply seems increasing by 5.3 hours the week prior to the interview, 3.6 hours per week on average and 221 hours during last year when the newborn is a son rather than a daughter. The estimated effect of the birth of a son is also positive and large but not different from zero. This positive effect offsets the

negative effect of having any child and lead to no change of parent's labor outcomes corresponding to having a son. The estimated effect on hourly earnings of the birth of a child and differential effect of the birth of a son versus a daughter is minimal. The estimated gender bias on parental labor outcomes seems generated by the fact that the birth of a son does not hinder parents from working or makes them come back to the original work faster than the birth of a daughter.

To examine how the findings change between fathers and mothers, I repeat the analysis using fixed effects approach to the subsample of fathers and subsample of mothers separately. The estimated results are shown in Table 7. The effects of the birth of a child to mothers are estimated to be negative but the standard deviations are very large according to the left panel. The estimated coefficients of the birth of a son are positive but not significant, except for the working hours last week in Colum (1). It is noted that the analysis of labor outcomes of mothers may suffer from small sample problem since some of them are not observable before they join the surveyed households by marriage.

The estimates for the sample of fathers are shown in the right panel. Similar to the analysis of mothers, the birth of a child decreases father's weekly working time by 4-5 hours and annual working time by near 200 hours (i.e., 5 weeks). The estimated effects on annual wage earnings are consistent with prediction but not significant at the standard confidence level.

6. Discussions

6.1 Interpretation of the results

The direction and magnitude of the estimated effects of child gender on parent's labor supply are similar to the findings in Lundberg and Rose (2002) and Choi *et al* (2005), which suggests the son premia found in developed countries also exist in a developing country like China. The estimated effect of child gender on total yearly wage earnings from cross sectional analysis is higher than their estimates but similar to the findings in Knight et al (2010). Meanwhile, the fact of no direct difference of labor supply and earnings between parents of a son and parents of a daughter is not contrary to the findings in Wei and Zhang (2011). They explain that families with a daughter catch up to families with a son due to the concern of intrafamily bargaining power which is determined by the relative wealth of men and women.

The finding of different patterns of child gender bias between families with an old child and families with a young child, however, seems puzzling. Traditional theories of preference for sons over daughters, differential costs of raising a son over a daughter, competitive motive and the role model cannot explain this pattern since these theories predict the effect of child gender should be persistent as children grow up. An importance difference between these two types of families is the demand of child care which leads me to test if the child gender bias is generated by demand and supply of child care.

In the remaining part of this section, I provide evidence to link the son premia found in the previous section to parent's choices of child care. In particular, I examine the choice of non-parental child care providers and expenditures on child care and the influence of caregivers on parental labor outcomes.

6.2 Intergenerational transfer of childcare and gender

A summary of parental childcare arrangement is in Table Appendix (available upon request) shows that for parents who have valid answer to the type of child care, 26% indicate no other caregivers take care of their children during the week prior to the interview. 30% show paternal grandparent provides care for a grandchild and less than 6% show maternal grandparent does. 39% of parents send their children to a day care which proves to be the main substitute to parent's care. 72% of parents indicate they spend some time and average 14 hours[7] to look after the child the week prior to the interview. 48% parents have expenses on child care paid to a babysitter, grandparent or day care facility[8]. As to the geographic proximity, 23% households have three generations with one grandparent. 49% households have two grandparents living in the same household. For maternal grandparents, 10% live in the same household, next door or adjacent areas, 12% live in the same neighborhood or village. 63% live outside the neighborhood but within the same city and 16% live in a different city. But for paternal grandparents, near 80 live in the same household or nearby. The information of caregiver and living proximity supports a strong pattern of a patrilineal culture.

[7] Standard deviation is very large which indicates strong heterogeneity among individuals.
[8] The amount averages 196 Chinese yuan inflated to 2009 currency.

As an older generation, grandparents, in particular father's side, may have a stronger preference of boys over girls because only boys could carry the family name and inherit the family patrimony (Li and Wu [2010], Bernhardt [1995]). Then it is natural to speculate that paternal grandparents are willing to provide more caring for grandsons than granddaughters. This directly reduces time spent on child care from parents who can get back to workplace sooner and provide more labor supply to wage work. To test this hypothesis, I first examine the choice of child care using a multinomial model.

$$log\frac{P(other\ care_i)}{P(parent\ care\ only_i)} = \alpha + \beta \times First\ born\ son_i + \gamma X_i + \varepsilon_i \quad (3)$$

In Equation (3), the left-hand-side variable is the relative odds of some non-parent caregiving vs. parent care only which is defined as the base category. $First\ born\ son_i$ indicates if the first born child is a boy. X_i includes covariates of grandchild's demographics, parent's demographics, paternal living proximity, maternal living proximity, province, survey year, locality and local economic conditions. Since there is no need for child care before the child is born I estimate the empirical model using cross sectional methods. It is noted that both child care choice and labor supply are endogenously determined within the system. In this sense, the estimated results should be interpreted with caution.

Table 8 presents the results of how the availability of grandparent's care responses to the gender of a grandchild. Both columns in this table show a quite stable finding that paternal grandparents are more likely to provide care to a grandson than a granddaughter. Both estimates are significant at the regular confidence level. Using the standard interpretation of the relative risk ratios, the relative risk ratio of grandparent's care for a grandson is 1.5 higher than care for a granddaughter relative to care of the child at home by parents only holding other variables in the model constant according to the first column. The relative risk ratios are even higher among parents with only one child (shown in column 2).

Meanwhile, the relative risk ratio of maternal grandparent's care and formal daycare relative to parental care only are not significantly different between families having a boy and families having a girl.

6.3 Intergenerational transfer of childcare and parental labor outcomes

I then examine the magnitude of the effect of the informal grandparent's caring for a grandchild on parental labor outcomes using an instrumental variable approach.

$$Y_i = \alpha + \beta_1 \times Any\ grandparenting_i + \gamma X_i + \varepsilon_i \quad (4)$$

$$Any\ grandparenting_i = \alpha + \delta_1 \times distance_i + \gamma X_i + \epsilon_i \quad (5)$$

where $Any\ grandparenting_i$ indicates if paternal grandparents help to care for the child and $distance_i$ measures the geographic proximity between grandparents and grandchild. I recoded the original geographic distance to four categories which are the same household, next door or adjacent areas, same neighborhood or village, outside neighborhood but within the same city or county, and other city or county.

The validity of the set of distance instruments rely on two assumptions. First, the geographic proximity is predetermined outside the system but labor supply and the fertility is endogenously determined. What may weaken the validity of the instruments is the geographic proximity is an ex post response to parental labor outcomes[9]. Another concern is the living arrangement of grandparents influence parent's labor outcomes via other channels in additional to child care. To address this problem, I then construct a simple falsification test.

$$Y_i = \alpha + \beta_1 \times distance_i + \gamma X_i + \varepsilon_i \quad (6)$$

I test this equation using a subsample of families with a son and subsample of families with a daughter. If the variables of distance influence parental labor outcomes directly, then the estimated value of β_1 should be of similar direction and magnitude between these two types of families. A preliminary test indicates that β_1 is positive and large for families with a son but not significantly different from zero for families with a daughter. Findings in Table 8 also indicate that the distance influences parental labor outcomes via caring for grandchild.

Table 9 presents the results estimated via the IV approach. The estimated coefficient in Column (1) is 2.1 but not significantly different from zero since the standard deviation is large. Column

[9] To address this potential reverse causality, I can construct an analog model with the geographic proximity before the child was born used as the instrument.

(2) and (3) in this table indicate that parents who get help from grandparents on child care average working 7.6 hours per week (near 20% of weekly hours) and over 330 hours per year (16.5%) more than parents without any paternal grandparent's caring for grandchildren. The estimated effects of this intergenerational linkage to earnings shown in the remaining columns are not estimated different from zero.

7. Robustness, Extension and Conclusion

All previous analysis focuses on active workers who are not working in an agricultural sector. This type of sample selection may lead to an underrepresentation of females since women have a lower employment rate than men, although female employment rate in China is very high. In this section, I plan to make more examinations on female labor force participation. A preliminary estimate of the probit model shows that female labor force participation is not significantly influenced by child gender. Tobit regressions which account for non-active workers generate similar findings to OLS. I also make an analog analysis by removing the restriction of non-agricultural workers. The results are in general consistent with previous findings. Table Appendix of these two robustness checks is not reported but available upon request.

The findings that parents correspond to gender of offspring have important implications on children's well-beings and development in the short term and long term.

7.1 Parental investment of money on child

Parents may pay to a caregiver like babysitter, a grandparent or a day care facility for help. I check family monthly expenditure on child care by using the following single equation.

$$Log(Expenditure_i) = \alpha + \beta_1 \times First\ born\ son_i + \gamma X_i + \varepsilon_i \quad (7)$$

where the dependent variable is the log transformation of monthly childcare expenditures. Definitions of the other variables in Equation (7) are exactly the same as those in Equation (1).

The estimated results via cross sectional regression are presented in Table 10. Findings for families with one child shown from column (1) to (3) are quite robust and suggest parents with a young son spend more expenditure on formal childcare. The difference is as large as 23% among parents who report they have such an expense (Column 1) and 50% among all parents who have

a preschool-age child (Column 3). Column (2) also indicates parents with a son are 11% (P-value = 0.15) more likely to choose formal child care than parents with a daughter.

The next three columns in the right panel report an analog analysis for families with one or multiple children. Both the direction and magnitude of the effects are similar to the left panel.

7.2 Parental investment of time on child

Table 11 presents the estimated effect of having a first born on parent's weekly hours spent caring for their children. Results in the left panel show that mothers and fathers who participate in child care activity spend 2.9 fewer hours per week when the only child is a boy than a girl. However, the participation in child care is positively correlated with having a first born son as indicated in Column (2) (P-value = 0.15). The overall gender gap on parental caring time is estimated negative but not significantly different from zero. All this evidence suggests that the decrease in child care time is likely to occur at the internal margin.

The right panel in Table 11 shows the estimated gender gap in families with one or more children. Results are very similar to the findings for one-child families. The magnitude of the effect of child gender is slighter smaller.

In the middle panel, mothers with a first born son are estimated to spend 3.4-3.5 hours per week on child care than mothers with a first born daughter (P-value = 0.15). The participation, however, is higher for mothers with a son than mothers with a daughter. And the overall effect is negative, large but not significant.

The panel in the bottom estimated the effect on fathers. Overall the estimated signs are similar to previous findings in the top and middle panels, except that in none of the cases, the estimated coefficients are different from zero. Therefore, the gender bias on child care is more striking for mothers than fathers.

Findings in Table 10 and 11 provide evidence that young boys obtain more parental investment of money but less investment of time than girls. This gender-based childcare arrangement and parental investments of time and money may lead to differences in children's short and long term developmental outcomes (Zhai and Gao, 2010) which desires more research in the future.

7.3 Conclusion

This paper uses a longitudinal data set from the CHNS to examine the effects of child gender on parental labor supply and earnings in the context of China's prevalence of son preference. Both OLS and fixed effects analysis provide supportive evidence of the existence of son premia on parental labor supply. The premia are 3-5% estimated from OLS and near 10% estimated via a fixed effect model. Cross sectional analysis also indicates a 6-15% earning premia associated with having a son depending on the age of children. A further analysis by children's age group indicates there are almost no difference between families with a 7-19-year old son and a daughter of similar age, which seems a puzzle. I attempt to use the choice of childcare and intergeneration ties between grandparents and grandchildren to explain the puzzle and conclude that parents with a young son are able to increase their labor supply faster than parents with a young daughter because grandparents provide more childcare to grandsons. This theory, however, does not fully rule out other possibilities. More work is thus required to construct in order to better explain the findings in this paper.

References

Bernhardt, Kathryn. 1995. “The Inheritance Rights of Daughters: the Song Anomaly?” *Modern China* 21(3): 269–309.

Bogan, Vicki. 2010. “Household Investment Decisions and Offspring Gender: Parental Accounting.” Available at SSRN: http://ssrn.com/abstract=1434070

Chen, Feinian, Guangya Liu and Christine A. Mair. 2011. “Intergenerational Ties in Context: Grandparents Caring for Grandchildren in China.” *Social Forces* 90 (2): 571-594.

Choi, Hyung-Jai, Shelly J. Lundberg and Jutta M. Joesch. 2005. “Work and Family: Marriage, Children, Child Gender and the Work Hours and Earings of West German Men.” IZA Discussion Paper No. 1761. Available at SSRN: http://ssrn.com/abstract=822878

Compton, Janice and Robert A. Pollak. 2011. “Family Proximity, Childcare, and Women’s Labor force Attachment.” NBER Working Paper 17678. Available at http://www.nber.org/papers/w17678

Dahl, Gordon B. and Enrico Moretti. 2008. “The Demand for Sons.” *Review of Economic Studies* 75(4): 1085-1120.

Du, Fenglian and Xiao-yuan Dong. 2010. “Women’s Labor Force Participation and Childcare Choices in Urban China during the Economic Transition.” Working Paper. Available at http://ideas.repec.org/s/win/winwop.html

Ebenstein, Avraham and Steven Leung. 2010. “Son Preference and Access to Social Insurance: Evidence from China’s Rural Pension Program.” *Population and Development Review* 36(1): 47-70.

Ebenstein, Avraham. 2010. “The “Missing Girls” of China and the Unintended Consequences of the One Child Policy.” *Journal of Human Resources* 45(1): 87-115.

Edlund, Lena, Junjian Yi, Hongbin Li and Junsen Zhang. 2010. “Sex Ratios and Crime: Evidence from China.” IZA Discussion Paper no. 3214.

Glauber, Rebecca and Kristi L. Gozjolko. 2011. “Do Traditional Fathers Always Work More? Gender Ideology, Race, and Parenthood.” *Journal of Marriage and Family* 73 (5): 1133-1148.

Jones-Smith, Jessica C. and Barry M. Popkin. 2010. “Understanding Community Context and Adult Health Changes in China: Development of an Urbanicity Scale.” *Social Science and Medicine* 71(8): 1436-1446.

Knight, John, Li Shi and Deng Quheng. 2010. “Son Preference and Household Income in Rural China.” *Journal of Development Studies* 46(10): 1786-1805.

Korenman, Sanders, and David Neumark, 1991. “Does Marriage Really Make Men More Productive?” *Journal of Human Resources* 26(2): 282-307.

Korenman, Sanders, and David Neumark, 1992. “Marriage, Motherhood ad Wages.” *Journal of Human Resources* 27(2): 233-255.

Li, Lixing and Xiaoyu Wu. 2010. “Gender of Children, Bargaining Power, and Intrahousehold Resource Allocation in China.” *Journal of Human Resources* 46 (2):295-316.

Lundberg, Shelly J. and Elaina Rose. 2002. “The Effect of Sons and Daughters on Men’s Labor Supply and Wages.” *Review of Economics and Statistics* 84(2): 251-68.

Lundberg, Shelly. 2005 “The Division of Labor by New Parents: Does Child Gender Matter?” IZA DP No. 1787.

Maurer-Fazio, Margaret, Rachel Connelly, Chen Lan and Lixin Tang. 2011. “Childcare, Eldercare, and Labor Force Participation of Married Women in Urban China: 1982−2000.” *Journal of Human Resources* 46 (2): 261-294.

Pabiloni, Sabrina Wulff and Jennifer Ward-Batts. 2007. “The Effect of Child Gender on Parents' Labor Supply: An Examination of Natives, Immigrants, and their Children.” *American Economic Review* 97(2): 402-406.

Wang, Shing-Yi. 2012. “Credit Constraints, Job Mobility and Entrepreneurship: Evidence from a Property Reform in China.” Review of Economics and Statistics 94 (2): 523-551.

Wei, Shang-Jin and Xiaobo Zhang. 2011. “The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China.” *Journal of Political Economy* 119(3): 511-564.

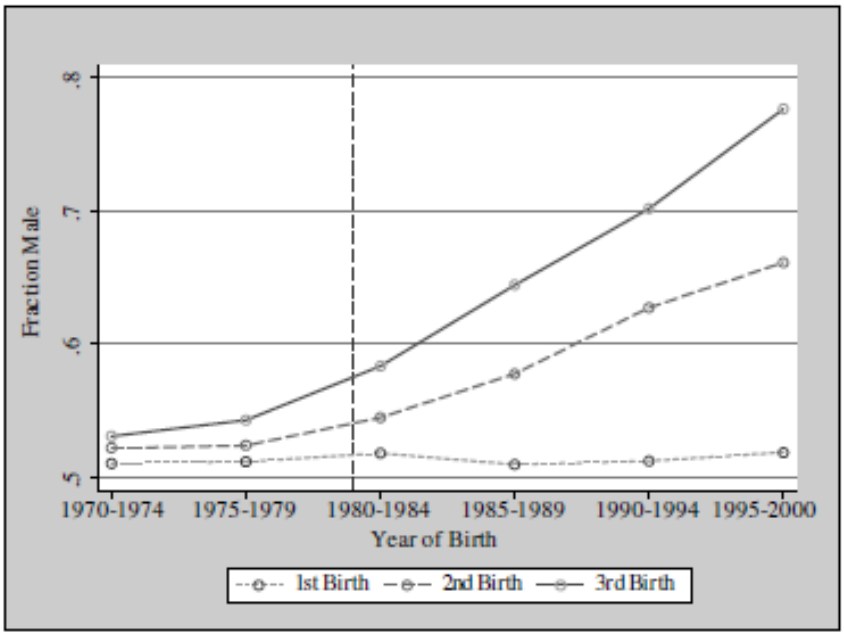
Wei, Shang-Jin and Xiaobo Zhang. 2011. “Sex Ratios, Entrepreneurship, and Economic Growth in the People's Republic of China.” NBER Working Paper No. 16800.

Wei, Shang-Jin, Xiaobo Zhang and Yin Liu. 2012. “Status Competition and Housing Prices.” NBER Working Paper No. 18000.

Zhai, Fuhua, and Qin Gao. 2008. “Center-Based Early Childhood Education and Care in China: Policies, Trends, ad Implications.” *Asian Social Work and Policy Review* 2(3): 127-148.

Zhai, Fuhua and Qin Gao. 2010. “Center-Based Care in the Context of One-Child Policy in China: Do Child Gender and Siblings Matter?” *Population Research and Policy Review* 29(5): 745-774.

Figure 1. Male Fraction of Births following Daughters in China



Source: Ebenstein (2010).

Figure 2. Participating provinces in China Health and Nutrition Survey

Source: CHNS website. The regions of dark green in this map are the provinces in which the survey has been conducted.

Table 1. Summary Statistics of the Variables

	Family with one child			Family with at least one child		
Variable	Mean	Sd	N	Mean	Sd	N
Panel A: Demographics						
Male	0.540	0.498	5,624	0.564	0.496	7,605
Age	35.358	6.035	5,624	36.041	5.951	7,605
Household head	0.375	0.484	5,624	0.417	0.493	7,605
Household spouse	0.327	0.469	5,624	0.326	0.469	7,605
Household child	0.165	0.372	5,624	0.145	0.352	7,605
Household child-in-law	0.132	0.339	5,624	0.113	0.317	7,605
First born son	0.536	0.499	5,624	0.500	0.500	7,605
Age of child	9.521	5.296	5,515	10.414	5.248	7,467
Number of child	1.000	0.000	5,624	1.298	0.543	7,605
Any child care (<=6)	0.734	0.442	1,525	0.719	0.450	1,671
Hours of child care (<=6)	13.629	25.528	1,302	13.452	25.640	1,435
Secondary school	0.102	0.302	5,624	0.129	0.335	7,605
High school	0.492	0.500	5,624	0.485	0.500	7,605
College	0.351	0.477	5,624	0.301	0.459	7,605
Administrator	0.192	0.394	5,624	0.178	0.383	7,605
Ordinary worker	0.808	0.394	5,624	0.822	0.383	7,605
Transportation facility	6.395	2.271	5,624	6.165	2.342	7,605
Social service	1.822	2.735	5,624	1.523	2.518	7,605
Health condition	6.615	1.893	5,624	6.430	1.977	7,605
Population density	6.369	1.459	5,624	6.212	1.508	7,605
Economic activity	6.218	2.830	5,624	5.728	2.867	7,605
Year = 1991	0.192	0.394	5,624	0.217	0.412	7,605
Year = 1993	0.159	0.366	5,624	0.182	0.386	7,605
Year = 1997	0.160	0.366	5,624	0.162	0.368	7,605
Year = 2000	0.140	0.347	5,624	0.136	0.343	7,605
Year = 2004	0.123	0.329	5,624	0.109	0.312	7,605
Year = 2006	0.129	0.336	5,624	0.109	0.311	7,605
Year = 2009	0.097	0.296	5,624	0.085	0.279	7,605
City	0.318	0.466	5,624	0.267	0.442	7,605
Suburb	0.216	0.412	5,624	0.215	0.411	7,605
Town	0.280	0.449	5,624	0.270	0.444	7,605
Village	0.186	0.389	5,624	0.248	0.432	7,605
Liaoning	0.138	0.345	5,624	0.121	0.326	7,605
Heilongjiang	0.094	0.291	5,624	0.074	0.262	7,605
Jiangsu	0.213	0.410	5,624	0.179	0.383	7,605
Shandong	0.121	0.327	5,624	0.130	0.336	7,605
Henan	0.085	0.278	5,624	0.087	0.281	7,605
Hubei	0.106	0.308	5,624	0.126	0.331	7,605
Hunan	0.089	0.285	5,624	0.103	0.304	7,605
Guangxi	0.096	0.295	5,624	0.112	0.316	7,605
Guizhou	0.058	0.234	5,624	0.070	0.255	7,605
Panel B: Labor outcomes						
Working hours last week	44.778	13.616	5,575	44.755	14.538	7,535
Weekly hours last year	45.723	11.758	5,624	45.862	12.934	7,605
Yearly hours	2115.5	588.8	5,624	2107.8	650.6	7,605
Log (hourly wage)	1.436	0.625	5,624	1.386	0.626	7,605

Log (wage income)	8.610	1.314	5,624	8.481	1.529	7,605

Notes: This table reports the summary statistics of the sample used for cross sectional analysis. Sample includes fathers and mothers who are active workers, between 22 and 50 years old and surveyed during the years of 1991-2009.

Table 2. Probit Regression of the Likelihood of Having a First Born Son

	Family with one child			Family with at least one child		
	(1)	(2)	(3)	(4)	(5)	(6)
	0-19 years old	7-19 years old	0-6 years old	0-19 years old	7-19 years old	0-6 years old
Parent birth cohort: 1956-60	0.146**	0.083	0.134	0.050	-0.020	-0.007
	(0.059)	(0.064)	(0.254)	(0.046)	(0.049)	(0.248)
Birth cohort: 61-65	0.064	-0.022	-0.051	0.105**	-0.001	-0.147
	(0.056)	(0.072)	(0.240)	(0.046)	(0.058)	(0.236)
Birth cohort: 66-70	-0.140**	-0.175**	-0.306	-0.054	-0.142*	-0.374
	(0.065)	(0.088)	(0.248)	(0.055)	(0.074)	(0.244)
Birth cohort: 71-75	-0.141*	-0.220**	-0.195	-0.009	-0.105	-0.248
	(0.078)	(0.110)	(0.271)	(0.069)	(0.095)	(0.266)
Birth cohort: 76-80	-0.231**	-0.098	-0.373	-0.041	-0.030	-0.358
	(0.101)	(0.202)	(0.289)	(0.091)	(0.179)	(0.282)
Birth cohort: 81-85	-0.354**		-0.385	-0.136		-0.384
	(0.162)		(0.326)	(0.152)		(0.318)
Secondary school	0.129	0.031	0.305	0.018	-0.028	0.078
	(0.091)	(0.104)	(0.194)	(0.064)	(0.070)	(0.175)
High school	0.010	0.003	-0.078	0.034	0.048	-0.187
	(0.078)	(0.090)	(0.166)	(0.056)	(0.061)	(0.151)
College	-0.040	-0.144	0.046	0.039	-0.017	-0.028
	(0.082)	(0.096)	(0.172)	(0.061)	(0.068)	(0.158)
Administer	-0.075	-0.105**	0.054	-0.064	-0.090**	0.066
	(0.047)	(0.053)	(0.100)	(0.041)	(0.045)	(0.097)
Community Transportation facility	0.0155*	0.012	0.0268*	0.0118*	0.006	0.0309**
	(0.008)	(0.010)	(0.014)	(0.007)	(0.008)	(0.014)
Community quality of health	0.006	-0.010	0.0378*	0.004	-0.001	0.023
	(0.011)	(0.014)	(0.020)	(0.009)	(0.011)	(0.019)
Community population density	-0.0228*	-0.016	-0.029	0.002	0.008	-0.017
	(0.013)	(0.017)	(0.023)	(0.011)	(0.013)	(0.021)
Community economic activity	-0.008	-0.007	-0.012	0.007	0.012	-0.009
	(0.008)	(0.010)	(0.016)	(0.007)	(0.008)	(0.015)
Rural	0.247***	0.281***	0.244***	0.136***	0.135***	0.193***
	(0.041)	(0.050)	(0.074)	(0.034)	(0.040)	(0.070)
Province dummies	X	X	X	X	X	X
Year dummies	X	X	X	X	X	X
Observations	5624	3809	1812	7605	5629	1972
Pseudo R-squared	0.017	0.025	0.042	0.008	0.012	0.036

Notes: Original estimates of probit model are reported. Base birth cohort is 1951-1955. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 3. Cross Sectional Estimates of Labor Outcomes of Parents

	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage
Panel A: One-child family (0-19 years old)					
Having a first born son	0.792**	0.464	32.90**	-0.001	0.0624**
	(0.356)	(0.300)	(15.430)	(0.011)	(0.030)
Observations	5,515	5,563	5,563	5,563	5,563
R-squared	0.087	0.127	0.078	0.601	0.281
Panel B: One-child family (7-19 years old)					
Having a first born son	0.254	-0.0438	10.86	-0.008	0.0171
	(0.430)	(0.365)	(18.450)	(0.013)	(0.035)
Observations	3,751	3,776	3,776	3,776	3,776
R-squared	0.089	0.124	0.082	0.604	0.293
Panel C: One-child family (0-6 years old)					
Having a first born son	1.633**	1.454***	76.74***	0.012	0.164***
	(0.655)	(0.541)	(28.790)	(0.020)	(0.060)
Observations	1,764	1,787	1,787	1,787	1,787
R-squared	0.098	0.142	0.1	0.579	0.259
Panel D: One-or-multiple-child family (first child 0-6 years old)					
Having a first born son	1.392**	1.056*	62.98**	0.006	0.124*
	(0.668)	(0.567)	(29.230)	(0.021)	(0.070)
Observations	1,835	1,861	1,861	1,861	1,861
R-squared	0.19	0.226	0.229	0.560	0.474
Demographics	X	X	X	X	X
Year	X	X	X	X	X
Province	X	X	X	X	X
Locality	X	X	X	X	X
Local economic conditions	X	X	X	X	X

Notes: The estimated coefficients and standard deviations (in parentheses) via OLS are reported in this table. Panel A, B and C use three different samples defined by child's age. See narrative for the explanation of covariates. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 4. Cross Sectional Estimates of Labor Outcomes of Parents by Gender

	Mothers					Fathers				
	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage	(6) Last weekly hour	(7) Weekly hour	(8) Annual hour	(9) Hourly wage	(10) Annual wage
Panel A: One-child family (0-19 years old)										
Having a first born son	0.563	0.254	35.7	-0.011	0.052	1.098**	0.674*	34.20*	0.003	0.0638*
	(0.512)	(0.445)	(23.480)	(0.015)	(0.050)	(0.492)	(0.403)	(20.240)	(0.015)	(0.036)
Observations	2,572	2,587	2,587	2,587	2,587	3,004	3,037	3,037	3,037	3,037
R-squared	0.098	0.115	0.079	0.623	0.257	0.089	0.145	0.088	0.591	0.314
Panel B: One-child family (7-19 years old)										
Having a first born son	-0.386	-0.165	15.06	-0.029	-0.0193	0.926	0.165	14.72	0.006	0.0523
	(0.613)	(0.546)	(27.940)	(0.018)	(0.059)	(0.600)	(0.490)	(24.440)	(0.018)	(0.041)
Observations	1,775	1,782	1,782	1,782	1,782	2,012	2,030	2,030	2,030	2,030
R-squared	0.104	0.124	0.094	0.621	0.268	0.091	0.137	0.087	0.599	0.333
Panel C: One-child family (0-6 years old)										
Having a first born son	2.128**	0.912	68.07	0.021	0.209**	1.202	1.643**	75.30**	-0.001	0.1
	(0.975)	(0.800)	(45.010)	(0.028)	(0.100)	(0.895)	(0.732)	(37.060)	(0.028)	(0.074)
Observations	797	805	805	805	805	992	1,007	1,007	1,007	1,007
R-squared	0.117	0.129	0.102	0.619	0.251	0.105	0.175	0.126	0.565	0.282
Panel D: One-or-multiple child family (first child 0-6 years old)										
Having a first born son	2.274**	1.353*	81.51*	0.015	0.246**	0.562	1.446*	79.62**	-0.001	0.0961
	(0.981)	(0.793)	(44.520)	(0.027)	(0.100)	(0.867)	(0.746)	(37.810)	(0.027)	(0.077)
Observations	846	854	854	854	854	1,100	1,118	1,118	1,118	1,118
R-squared	0.094	0.12	0.099	0.618	0.249	0.091	0.145	0.099	0.550	0.247
Demographics	X	X	X	X	X	X	X	X	X	X
Year	X	X	X	X	X	X	X	X	X	X
Province	X	X	X	X	X	X	X	X	X	X
Locality	X	X	X	X	X	X	X	X	X	X
Local economic conditions	X	X	X	X	X	X	X	X	X	X

Notes: See Table 3 for the notes.

Table 5. Cross Sectional Estimates of Family Aggregation and Gender Difference

	Gender aggregation					Gender difference				
	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage	(6) Last weekly hour	(7) Weekly hour	(8) Annual hour	(9) Hourly wage	(10) Annual wage
Panel A: One-child family (0-19 years old)										
Having a first born son	1.613**	1.100*	74.15**	0.019	0.0969	0.38	0.567	15.71	0.006	0.0762
	(0.731)	(0.637)	(31.700)	(0.026)	(0.069)	(0.551)	(0.477)	(24.900)	(0.019)	(0.051)
Observations	2,939	2,965	2,950	2,778	2,797	2,939	2,965	2,950	2,778	2,797
R-squared	0.656	0.711	0.692	0.753	0.89	0.62	0.693	0.628	0.686	0.882
Panel B: One-child family (7-19 years old)										
Having a first born son	0.902	0.502	46.1	0.012	-0.0123	1.008	0.681	27.37	0.030	0.155**
	(0.909)	(0.782)	(38.390)	(0.031)	(0.085)	(0.649)	(0.577)	(29.550)	(0.023)	(0.061)
Observations	2,023	2,038	2,026	1,909	1,921	2,023	2,038	2,026	1,909	1,921
R-squared	0.643	0.703	0.69	0.772	0.89	0.617	0.677	0.622	0.703	0.887
Panel C: One-child family (0-6 years old)										
Having a first born son	2.978**	2.280**	144.2**	0.010	0.346***	-0.786	0.616	2.88	-0.059*	-0.116
	(1.256)	(1.123)	(57.330)	(0.051)	(0.125)	(1.056)	(0.865)	(46.820)	(0.034)	(0.096)
Observations	916	927	924	869	876	916	927	924	869	876
R-squared	0.692	0.738	0.707	0.718	0.892	0.64	0.732	0.656	0.668	0.876
Panel D: One-or-multiple child family (first child 0-6 years old)										
Having a first born son	2.483**	2.598**	169.4***	0.014	0.409***	-1.124	0.237	3.978	-0.041	-0.129
	(1.242)	(1.129)	(57.620)	(0.050)	(0.129)	(1.035)	(0.887)	(47.560)	(0.033)	(0.094)
Observations	991	1,005	1,002	936	944	991	1,005	1,002	936	944
R-squared	0.678	0.717	0.686	0.707	0.877	0.636	0.719	0.64	0.665	0.876
Demographics	X	X	X	X	X	X	X	X	X	X
Year	X	X	X	X	X	X	X	X	X	X
Province	X	X	X	X	X	X	X	X	X	X
Locality	X	X	X	X	X	X	X	X	X	X
Local economy	X	X	X	X	X	X	X	X	X	X

Notes: See Table 3 for the notes.

Table 6. Fixed Effects Estimates of Labor Outcomes of Parents

	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage
Panel A: Family with one child					
Postnatal	-1.97	-3.796***	-195.0**	0.136**	0.0017
	(1.713)	(1.442)	(80.390)	(0.055)	(0.165)
Postnatal × Son	5.388***	3.606**	221.4**	-0.079	0.071
	(1.862)	(1.566)	(87.320)	(0.060)	(0.179)
Observations	1,749	1,773	1,773	1,773	1,773
R-squared	0.076	0.071	0.064	0.554	0.266
Panel B: Family with one child younger than 7					
Postnatal	-2.848	-4.374***	-201.5**	0.075	-0.188
	(1.829)	(1.567)	(92.640)	(0.062)	(0.206)
Postnatal × Son	3.865*	3.263*	187.6*	-0.078	-0.027
	(1.979)	(1.695)	(100.200)	(0.067)	(0.223)
Observations	1,353	1,374	1,374	1,374	1,374
R-squared	0.114	0.089	0.063	0.428	0.194
Panel C: Family with at least one child					
Postnatal	-2.101	-3.835***	-184.5**	0.0742	-0.078
	(1.582)	(1.325)	(73.290)	(0.049)	(0.152)
Postnatal × Son	5.883***	4.080***	232.4***	-0.072	0.057
	(1.734)	(1.452)	(80.350)	(0.054)	(0.167)
Observations	2,119	2,151	2,151	2,151	2,151
R-squared	0.059	0.057	0.05	0.530	0.228
Panel D: Family with at least one child younger than 7					
Postnatal	-2.555	-4.650***	-213.5**	0.066	-0.203
	(1.746)	(1.458)	(84.920)	(0.056)	(0.182)
Postnatal × Son	4.721**	3.365**	191.6**	-0.067	-0.001
	(1.831)	(1.533)	(89.270)	(0.059)	(0.192)
Observations	1,573	1,600	1,600	1,600	1,600
R-squared	0.082	0.072	0.054	0.406	0.186
Demographics	X	X	X	X	X
Year	X	X	X	X	X
Local economic conditions	X	X	X	X	X

Notes: The estimated coefficients and standard deviations (in parentheses) via fixed effect models are reported in this table. Panels A-E use five different samples. All specifications include covariates as describe in the narrative. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 7. Fixed Effects Estimates of Labor Outcomes of Parents by Gender

	Mothers					Fathers				
	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage	(6) Last weekly hour	(7) Weekly hour	(8) Annual hour	(9) Hourly wage	(10) Annual wage
Panel A: Family with one child										
Postnatal	-1.943	-1.89	-106.5	0.071	0.178	-2.324	-4.305**	-216.8**	0.160**	-0.0345
	(2.985)	(2.621)	(140.3)	(0.093)	(0.282)	(2.145)	(1.776)	(100.6)	(0.070)	(0.210)
Postnatal × Son	7.976**	0.534	38.17	0.092	-0.038	4.483**	4.648**	283.6***	-0.15**	0.076
	(3.526)	(3.112)	(166.5)	(0.110)	(0.335)	(2.232)	(1.841)	(104.3)	(0.072)	(0.218)
Observations	632	639	639	639	639	1,117	1,134	1,134	1,134	1,134
R-squared	0.141	0.096	0.11	0.529	0.247	0.081	0.093	0.08	0.575	0.283
Panel B: Family with one child younger than 7										
Postnatal	-1.861	-0.834	-49.82	-0.004	0.175	-2.348	-4.732**	-215.6*	0.089	-0.256
	(3.835)	(3.308)	(191.9)	(0.128)	(0.408)	(2.153)	(1.844)	(110.4)	(0.074)	(0.252)
Postnatal × Son	8.430**	1.780	83.210	0.044	-0.300	2.288	3.587*	196.6*	-0.110	0.023
	(4.187)	(3.616)	(209.8)	(0.140)	(0.446)	(2.287)	(1.954)	(117.0)	(0.079)	(0.267)
Observations	450	456	456	456	456	903	918	918	918	918
R-squared	0.274	0.17	0.14	0.346	0.124	0.088	0.101	0.07	0.472	0.23
Panel C: Family with at least one child										
Postnatal	-3.189	-2.863	-172.9	0.017	-0.0288	-1.883	-3.987**	-185.2**	0.0947	-0.077
	(2.903)	(2.430)	(131.9)	(0.089)	(0.281)	(1.932)	(1.617)	(90.190)	(0.060)	(0.187)
Postnatal × Son	8.152**	2.108	167.6	0.004	-0.156	5.149**	4.60***	247.9***	-0.098	0.100
	(3.372)	(2.836)	(154.0)	(0.104)	(0.328)	(2.053)	(1.715)	(95.660)	(0.064)	(0.198)
Observations	759	768	768	768	768	1,360	1,383	1,383	1,383	1,383
R-squared	0.118	0.099	0.106	0.491	0.211	0.057	0.063	0.05	0.553	0.244
Panel D: Family with at least one child younger than 7										
Postnatal	-3.581	-2.374	-150.3	0.016	0.0298	-1.866	-5.02***	-225.2**	0.076	-0.255
	(3.750)	(3.008)	(174.0)	(0.118)	(0.372)	(2.008)	(1.711)	(99.710)	(0.066)	(0.217)
Postnatal × Son	10.4***	3.087	227.60	-0.019	-0.222	2.794	3.163*	159.6	-0.068	0.053
	(3.889)	(3.127)	(180.9)	(0.123)	(0.387)	(2.091)	(1.784)	(104.0)	(0.069)	(0.226)
Observations	519	527	527	527	527	1,054	1,073	1,073	1,073	1,073
R-squared	0.222	0.171	0.184	0.331	0.139	0.065	0.073	0.046	0.454	0.217

Notes: *: p<0.10; **: p<0.05; ***: p<0.01. See Table 6 for other notes.

Table 8. Effect of Child Gender on the Choice of Child Care

	(1) Family with a child younger than 7	(2) Family with the only child younger than 7
1: No non-parent caregiver		
Having a newborn son	Base category	
2: Paternal grandparent care		
Having a newborn son	1.487*	1.628*
	(0.353)	(0.434)
3: Maternal grandparent care		
Having a newborn son	1.993	1.733
	(0.928)	(0.91)
4: Day care provider		
Having a newborn son	1.25	1.218
	(0.277)	(0.321)
Observations	806	673
Pseudo R-squared	0.259	0.293
Parent's demographic	X	X
Paternal grandparent's distance	X	X
Maternal grandparent's distance	X	X
Grandchild's demographic	X	X
Province	X	X
Wave	X	X
Locality	X	X
Local economy	X	X

Notes: The estimated relative risk ratios e^{β_2} and the original standard deviations (in parentheses) via multinomial models are reported. All specifications include covariates as described in the narrative. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 9. Intergenerational Childcare Transfer and Parental Labor Outcomes

	(1) Last weekly hour	(2) Weekly hour	(3) Annual hour	(4) Hourly wage	(5) Annual wage
Paternal grandparent's care for a grandchild	2.127	7.582**	336.4*	-0.228	-0.001
	(4.25)	(3.65)	(192.40)	(0.15)	(0.28)
Observations	673	683	683	683	683
F-statistics in first stage	50.10	48.70	48.70	48.70	48.70
Demographic	X	X	X	X	X
Year	X	X	X	X	X
Province	X	X	X	X	X
Locality	X	X	X	X	X
Local economy	X	X	X	X	X

Notes: Sample includes fathers and mothers with the only child younger than 7 years old. Geographic proximity of paternal grandparents is used to instrument the availability of paternal grandparent's care for a grandchild. *: $p<0.10$; **: $p<0.05$; ***: $p<0.01$.

Table 10. Parent's Time Investment on Child Care by Gender

	Family with one child			Family with at least one child		
	(1) Internal	(2) External	(3) Overall	(4) Internal	(5) External	(6) Overall
Panel A: Mothers and fathers						
Having a first born son	-2.944*	0.029	-0.939	-2.465	0.026	-0.716
	(1.744)	(0.020)	(1.727)	(1.697)	(0.019)	(1.677)
Observations	1,156	1,595	1,595	1,266	1,774	1,774
R-squared	0.079	0.213	0.028	0.074	0.224	0.029
Panel B: Mothers						
Having a first born son	-3.423	0.021	-2.034	-3.534	0.011	-2.667
	(2.369)	(0.022)	(2.355)	(2.278)	(0.021)	(2.274)
Observations	711	819	819	785	904	904
R-squared	0.1	0.16	0.017	0.099	0.151	0.017
Panel C: Fathers						
Having a first born son	-2.355	0.035	-0.172	-0.716	0.037	1.260
	(2.451)	(0.034)	(2.462)	(2.398)	(0.032)	(2.401)
Observations	445	776	776	481	870	870
R-squared	0.053	0.138	0.016	0.047	0.142	0.017
Demographics	X	X	X	X	X	X
Year	X	X	X	X	X	X
Local economy	X	X	X	X	X	X

Notes: Internal margin is estimated via OLS; External margin is estimated via linear probability model; and Overall margin is estimated via Tobit model. Original estimated coefficients and standard deviations are reported. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 11. Parental Money Investment on Child Care by Gender

	Family with one child			Family with at least one child		
	(1) Internal margin	(2) External margin	(3) Overall	(4) Internal margin	(5) External margin	(6) Overall
Mothers and fathers						
Having a first born son	0.234***	0.114	0.500**	0.217***	0.125*	0.515**
	(0.061)	(0.079)	(0.222)	(0.060)	(0.074)	(0.217)
Observations	747	1,291	1,291	796	1,406	1,406
R-squared	0.357	0.139	0.052	0.339	0.129	0.049
Demographics	X	X	X	X	X	X
Province	X	X	X	X	X	X
Wave	X	X	X	X	X	X
Locality	X	X	X	X	X	X
Local economy	X	X	X	X	X	X

Notes: Internal margin is estimated via OLS; External margin is estimated via linear probability model; and Overall margin is estimated via Tobit model. Original estimated coefficients and standard deviations are reported. *: p<0.10; **: p<0.05; ***: p<0.01.