UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THREE ESSAYS ON DEVELOPMENT AND LABOR ECONOMICS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CHENG MA
Norman, Oklahoma
2022

THREE ESSAYS ON DEVELOPMENT AND LABOR ECONOMICS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF ECONOMICS

BY THE COMMITTEE CONSISTING OF

Dr. Le Wang, Chair

Dr. Hairong Song

Dr. Alexander Holmes

Dr. Tyler Ransom

Dr. Chunbei Wang

DEDICATION

to:

my parents, Wen and Fulin, and Gabe.

Acknowledgments

I'd like to thank all members of my committee. First I would like to thank Dr. Le Wang for supporting me and encouraging me in my research. It's also worth pointing out Dr. Le Wang's help with logistics that allowed me to stay in the U.S. to pursue my studies. I would like to thank Dr. Holmes for being a great professor to TA for. I'm very grateful for Dr. Ransom for his helpful notes that helped me complete the last chapter of this work. Dr. Chunbein Wang's Labor class inspired me to expand my horizons in my research field. I would also like to thank Dr. Hairong Song for her watchful eyes as the outside member of my committee.

Contents

List of Tables

List of Figures

Abstracts

Chapter 1 examines the causal relationship between family size and upstream intergenerational transfers. Successful family planning policies in China not only led to a dramatic decrease in fertility rate but also raised concerns about whether a smaller family-based elderly care network can provide adequate financial support. The relationship between family size and transfers needs more exploration. However, the endogeneity of family size hindered past studies from establishing a causal effect of family size on transfers. This chapter exploits the variations in the strictness of the One Child Policy implementation to address this endogeneity. Using the data from the China Health and Retirement Longitudinal Study (CHARLS), our estimation suggests that family size does not have an effect on either the probability of receiving transfers or the amount of transfers from children, which can be explained by both quantity-quality trade-off and parental financial need. In addition, our findings suggest that higher parental financial need explains the heterogeneity in effect of family size on transfers between urban households and rural households. Moreover, the inconsistency between our results and the literature is indicative of the heterogeneous nature of the effect of family size on transfers.

Chapter 2 studies the return to spousal education. Although a body of literature has extensively investigated the relationship between spousal education and one's earnings, a causal relationship is rarely established. This occurs because of the failure to address the endogeneity of spousal educational attainment. In this chapter, we use a method that utilizes the nonlinearity of control terms induced by heteroskedasticity to estimate the causal effect

of spousal education on one's own earnings. Using the data from the Chinese Household Income Project (CHIP), our estimation suggests that spousal schooling has a sizable positive effect on one's earnings. We find that spousal education increases one's earnings by raising hourly wages rather than lengthening work hours. Our findings also suggest that there exists substantial heterogeneity in return to spousal education by gender, by spousal education level, by year, by dominant earner status, and by whether spouses share the same occupation.

Chapter 3 is coauthored with Gabe Lebovich. In this chapter, we test the effect of additional children in a family on health and educational outcomes. Estimating this effect is complicated by the endogeneity of family size. We use the variation in the severity of the effect of the one-child-policy in China to extract exogenous variation from the China Health and Nutrition Survey in the same country. After finding a negative effect of family size on health and educational outcomes of children we use a newly developed machine learning approach. Generalized random forests allows us to look at the heterogeneity in treatment effects in the quantity-quality trade-off. We find robust negative treatment effects of additional children on health outcomes but only mild effects on educational attainment. The machine learning algorithm finds mother's age and parent's education level play a large role in the negative quantity-quality trade-off. Pinpointing the factors that exacerbate the negative effect of additional children on child quality can aid future policy decisions.

Chapter 1

Aging With Fewer Children: Estimating the Causal Effect of Family Size on Transfers

1.1 Introduction

It is undeniable that China's family planning policies are successful. Over the past few decades, China's fertility rate has dropped from 5.9 births per woman to today's less than 2 births per woman (Zeng and Hesketh 2016). In the meantime, life expectancy at birth increased from 43.7 years in 1960 to 74.4 years in 2010 [1].These transitions have dramatically changed China's traditional family-based elderly support system. On one hand, families have become smaller as fertility has declined, meaning that senior parents are facing a smaller family-based support network. On the other hand, parents are living longer as life expectancy has increased, which requires children to provide extended care and support. China relies heavily on families as the main source of elderly support (Zimmer and Kwong 2003; Lei et al.

[1] World Bank: `https://data-worldbank-org.ezproxy.lib.ou.edu/indicator/SP.DYN.LE00.IN?locations=CN`.

2015); however, the reality of longer-living parents with fewer children raised concerns about the ability of family networks to support senior parents.

From an economic perspective, there have been two theories put forward to explain the motives underlying children's behavior of providing support to parents: exchange theory (Bernheim et al. 1986, Cox 1987) and altruism theory (Becker 1974). Exchange theory (Bernheim et al. 1986; Cox 1987) views transfers from children as payments for the resources and services provided by parents. This theory is similar to the concept of reciprocity from sociology. Parents consider children as their old-age "support bank". Resources and services parents provide for their children are "deposits" to a "support bank", which can be drawn upon in the future in the form of support from children (Antonucci and Jackson 1990). Parents with more "deposits" in the "support bank" should expect more transfers from children. The altruism theory of (Becker 1974) suggests that children provide transfers to their parents based on parental financial need. Parents with more financial need receive more transfers from their children.

It is theoretically unclear how family size affects the transfers parents receive from their children. On one hand, having more children may lead to more transfers for three reasons. First, a larger family size provides a bigger family support network. With more potential family financial supporters, parents may not only have a higher probability of receiving financial support from at least one of their children, but also receive a larger amount of transfers. Second, parents may provide more services, such as grandchild care, for their children as the demand for these services may increase in family size. In exchange for parents' services, children may provide more transfers for their parents. Third, parents with more children may need more financial support from their children. Raising children can be stressful. Parents with more children may experience worse physical and mental health (Gove and Geerken 1977; Umberson and Gove 1989; Cáceres-Delpiano and Simonsen 2012;Wu and Li 2012; Canning and Schultz 2012) and may be unable to allocate more resources onto themselves. Therefore, they may have less savings compared to their counterparts with fewer children (Banerjee et al. 2010; Ge et al. 2018).

On the other hand, having more children may reduce the transfers parents receive from their children. Previous empirical studies have found that there is a quantity-quality trade-off in the Chinese context (Li et al. 2008; Rosenzweig and Zhang 2009; Liu 2014). Children from larger families are often allocated with fewer resources compared to those from smaller families. Based on the exchange theory (Bernheim et al. 1986, Cox 1987), parents with a larger family size should expect a smaller amount of transfers from each child. Although parents with more children may benefit from having a larger family support network, the total amount of transfers parents receive from their children may be smaller if the quantity-quality trade-off is sizable enough.

The ambiguous relationship between family size and transfers needs more investigation; however, past studies on transfers have almost entirely focused on differentiating the motives behind the transfers (for example, Jensen 2004; Cox et al. 2004; Gibson et al. 2011). Little attention has been paid to the effect of family size on transfers, although family size often enters the analysis as an important control variable in many of these studies. Moreover, difficulty in dealing with the endogeneity in family size hindered past studies from exploring the causal effect of family size on transfers. The source of endogeneity might come from parental heterogeneity. For example, if those parents who put more value on their children's old-age support also prefer to have more children, then the correlation between family size and transfers from children will be driven by parental preferences rather than by family size.

In this paper, we address the endogeneity issue by using the fact that the exposure to the One Child Policy (OCP) varied by prefecture and cohort. Substantial variations existed in the OCP enforcement intensity across prefectures due to the heterogeneity in local OCP implementation. The fertility of individuals who were young enough to be influenced by the OCP during their fertile years should be lower than that of older individuals in all prefectures. However, the difference should be larger in the prefectures where the OCP was more strictly implemented. Under the assumption that in the absence of the OCP, the difference in fertility between young individuals and older individuals would not have been systematically different in strict and relaxed prefectures after controlling prefecture characteristics, the difference in fertility change between the two types of prefectures is due to the difference in the OCP implementation stringency. Therefore, the combination of the OCP enforcement strictness and cohort, a difference-in-difference estimator, can extract the exogenous variation in family size. Similar identification strategy has been used in Duflo (2001) and Li and Zhang (2017).

The main data used in this paper comes from the China Health and Retirement Longitudinal Study (CHARLS) Wave 4. The survey provides information on family size coupled with detailed information data on intergenerational transfers, living arrangement and children's educational attainment. Using this dataset for our estimation we are able to highlight two findings in this paper. First, our results suggest that family size does not have a significant effect on the probability of parents receiving transfers or the amount of transfers parents receive. This could be explained by quantity-quality trade-off and parental financial need, which provides evidence for both exchange theory and altruism theory. Second, we find that there is a heterogeneous effect of family size on transfers between urban households and rural households, which is driven by the greater financial need among rural households.

This paper contributes to the existing literature in three ways. First, it contributes to understanding the causal effect of the number of children on upstream intergenerational transfers. Studies on relationship between family size and upstream intergenerational transfers typically find a positive correlation between family size and the probability of receiving transfers from children (Zimmer and Kwong 2003; Cai et al. 2006; Lei et al. 2012;Wu and Li 2014). The results on the effect of family size on the amount of transfers from children are mixed. Some studies find a positive relationship (Cai et al. 2006; Wu and Li 2014) while others find an insignificant relationship (Cong and Silverstein 2011; Lei et al. 2012). The main limitation of these studies is that they fail to address the endogeneity of family size.

The only recent paper that aims to establish a causal relationship between family size and transfers is Oliveira (2016). She uses the incidence of twins in the first birth as an IV for family size and finds a positive effect of family size on upstream intergenerational transfers. Without the homogeneity assumption, the IV approach delivers the average treatment effect of a group of compliers (Angrist and Imbens 1995). The compliers in Oliveira (2016) are women whose fertility was shifted away from their desired fertility due to the incidence of twins. They could be women who desired only one child, or women whose fertility choice was affected by the incidence of twins in the first birth due to the costly fertility adjustments (Oliveira 2016). Using the complier profiling method proposed in Marbach and Hangartner (2020), we find that compliers generated from using the incidence of twins in the first birth represent an older population with better education and higher income. Estimates derived from this subgroup may not be generalized to the entire population.

Our paper advances the existed literature by using a unique source of exogenous variation in family size provided by the OCP to address the endogeneity. The compliers in our analysis are individuals whose fertility choices were affected by the OCP. Given that the OCP is "one of the most restrictive and large scale family planning policies ever undertaken (Qian 2009)", our compliers should be more representative of the general population. Although compliers in our analysis are older, less educated and have less income compared to the entire sample, the differences of covariate means between compliers and the entire sample are much smaller than those between compliers generated from using the incidence of twins in the first birth and the entire sample. Therefore, our paper provides estimated effects that are closer to the average treatment effects of the general population. In addition, the inconsistency between our results and those in Oliveira (2016) suggest that there is heterogeneity in effect of family size on transfers.

Second, upstream intergenerational transfers are a flow of transfers from younger generation to older generation, which is an important component of the flow of transfers between family generations across the life cycle. The nature of intergenerational transfers is family resource redistribution throughout the life cycle. This paper contributes to understanding the link between fertility and family resource redistribution across the life cycle.

Third, this paper contributes to understanding the long-term consequences of the OCP. There is a large body of literature evaluating the effect of the OCP on various outcomes, including fertility (Ahn 1994; Li et al. 2005; Ebenstein 2010), sex ratios(Ebenstein 2010; Li et al. 2011; Loh and Remick 2015), marriage (Huang and Zhou 2015), children's outcomes (Rosenzweig and Zhang 2009; Qian 2009; Liu 2014; Li and Zhang 2017; Zeng et al. 2020), parental health (Islam et al. 2010; Wu and Li 2012), parental labor market outcomes (Cao 2019), and crime (Edlund et al. 2007). Although understanding how the OCP shaped family-based elderly support has strong policy implications, there has been little study of the effect of the OCP on old-age support.

1.2 Implementation of the One Child Policy

The OCP, one of the most ambitious family planning programs in the world, was introduced in 1979 to curb China's rapid population growth. One of the distinguishing features of the OCP is that this national policy was "interpreted, adapted, and implemented according to local conditions and needs" (Short and Fengying 1998). For instance, all the local OCP polices advocated having only one child and prohibited having a third child. However, local policies towards the birth of a second child varied from "no second child" to "no more than two children". Moreover, in regions where second child was not allowed, the birth of a second child was often permitted under a number of particular conditions and these conditions can be variously interpreted by local officials. One example given by Croll et al. (1985) is that a second child in rural areas were usually allowed "if an individual commune member is having true or real difficulties". Whether an individual is qualified for this condition depends on local official's interpretation on "true/real difficulties".

It has long been recognized that the OCP implementation was more successful in urban areas than that in rural areas. This is because urban residents were more closely and directly affected by government policies as many of them worked in state-owned enterprises and institutions (Zhang 2017). Urban OCP policies usually required couples to have only one child unless they were qualified for an exemption. Couples who did not comply the OCP would face economic penalties for the birth of an additional child. The economic penalties varied across regions. In some regions, the penalized couples needed to pay a one-time fine at the time of birth; in some regions, couples faced a 5% to 10% income deduction for 10 to 16 years after the birth of the addition child and the income deduction could increase to 20% for the birth of a fifth child. Moreover, couples who did not comply the OCP were not eligible for a number of benefits, such as job promotion, wage bonus, hardship subsidies (Croll et al. 1985). Urban couples faced a very high opportunity cost of having an additional child, therefore, they had a strong incentive to comply the OCP. In contrast, rural couples had little incentive as they received much fewer benefits from the government to begin with. In addition, agricultural work required a lot of manpower and this manpower mainly came from family. The common punishment for rural couples having an additional child was a one-time fine, which was difficult for government to collect because many rural couples were too poor to pay for it (Zhang 2017). Therefore, the desire for a larger family was not hindered by the potential penalty as much in rural areas as in urban areas.

Besides the heterogeneity in local OCP policies and local implementation difficulties, the implementation of the OCP was also heavily affected by local leadership. Although the National Family Planning Association was in charge of the implementation of the OCP, the policy actually took place at local level once the national policy was passed down from the administrative chain of command. The success of the OCP relied on whether individuals can be successfully persuaded by local officials to accept the new norm of having only one child. In some places, individuals who resisted to comply the OCP were visited by local officials multiple times until their resistance were worn down. However, in some places, there was no follow-up visit once the policy was expounded and explained to individuals (Croll et al. 1985).

1.3 Data

The data used in this paper comes from two sources. The first source is the 1982 Chinese Population Census. Using the method proposed in Li and Zhang (2017), we use the census data to construct the excess fertility rate (EFR) and prefecture control variables. EFR is calculated as the percentage of Han moms aged between 25-44 who gave a higher order birth in 1981. Since the OCP was officially implemented in 1979 to encourage couples to have only one child. Moms who had higher order birth in 1981 potentially violated the OCP. Therefore, EFR captures the potential violation rate in each prefecture. A low EFR indicates strict local enforcement of the OCP. Prefecture control variables are included in the analysis to control for pre-existing prefecture characteristics such as fertility preferences. The set of prefecture control variables include the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender.

The second source used for the analysis is the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative panel data set of Chinese residents aged 45 and above. The baseline of CHARLS was conducted in 2011-2012, covering about 10,000 households from 150 districts, 450 communities across China (Zhao et al. 2013). Following up surveys are conducted in every two years. In each wave, basic family demographic information, such as family size, education level, employment status, is collected for each household. In addition to that, respondents are also asked to report information on family transfer received in the past year. In this paper, we use the most recent data from CHARLS Wave 4. Since the EFR captures variations in the OCP policy enforcement among households with Han mothers, we also restrict our CHARLS data to households with Han mothers. For concern over measurement error, we exclude households for which the age gap between the mother and the eldest child is less than or equal to 10 as well as those with age gap greater or equal to 80.

Our final household sample consists of 9314 households. Table 1.1 displays summary statistics. On average, households in our sample have 2.62 children, with a relatively large standard deviation of 1.34 children. 21% of households contain only one senior parent. We define household age as the average of two parents' age if the household has two senior parents. If an household has only one senior parent, then his/her age is the household age. The average household age is about 62 years old. The average age of senior mothers in our sample is 61.18. Although only individuals aged 60 and over are officially considered as the elderly population in China, "old age" in labor market is much younger than 60. Typically, men retire at 60 and women retire between 50 - 55. Men aged 50 and women aged 40 are usually considered being old by employers and facing age discrimination in the labor market (Song 2014). Therefore, individuals younger than the official age of elderly might also need support from their children. Household education level is defined in the same way as household age. On average, household education level is primary school level [2]. Among all the households, 75% received financial support from their children in the past year. Financial support consists of money support and in-kind support. It can also be categorized into regular financial support and non-regular financial support based on whether the transfers from children occur according to a predetermined schedule. 53% households received money transfers and 65% households received in-kind transfers. The share of households that received regular transfers is low. 19% households received financial transfers regularly. 15% households received regular money transfers and only 10% received regular in-kind transfers. This indicates that most transfers occurred in a form of non-regular transfers.

The summary statistics also show differences between rural and urban households. We define a household as a rural household if the household is located in a rural area. Rural

[2] Four education levels: 1, No formal education or illiterate; 2, Primary school; 3, Middle school; 4, High school and above.

households have more children than urban households. The average number of living children of rural households is 2.82, compared to 2.34 for urban households. Rural households and urban households are similar with respect to household age and senior mother's age, but rural households tend to have a lower household education level. In addition, rural households are more likely to receive financial support from children. On average, they also received slightly smaller transfer amounts.

1.4 Method

1.4.1 Instrumental variable identification strategy

We use household's exposure to the OCP to extract the exogenous variation in family size. Since local governments adapted and implemented the OCP based on their local conditions and needs, the OCP enforcement intensity varied across prefectures. Therefore, household's exposure to the OCP is heavily influenced by household's prefecture of residence. In our sample, the EFR ranges from 1.7% to 18%. The variation in EFR itself is not sufficient to extract the exogenous variation in family size because more relaxed OCP policies were often implemented in prefectures with a preference of higher fertility. Household's exposure to the OCP is also determined by mother's age. Based on mother's age, we divide our sample into two cohorts, the young cohort and the old cohort. The young cohort were subject to much greater influence of the OCP than the old cohort. The reason is that mothers from the young cohort are younger and their peak reproductive years are more likely to be fully covered by the OCP. In contrast, mothers from the old cohort are more likely to finish their childbearing before the OCP was implemented. For mothers from the old cohort who did not finish their childbearing before the OCP was introduced, their childbearing period is more likely to be only partially covered by the OCP, meaning that their exposure to the OCP should be smaller than that of mothers from the young cohort.

The basic idea behind our instrumental variable (IV) identification strategy is that households from the young cohort have fewer children than those from the old cohort due to the implementation of the OCP and the differences in family size between the young cohort and the old cohort are larger in prefectures where the OCP was more strictly implemented. Under the assumption that differences in family size between the young cohort and the old cohort would not have been systematically different across prefectures if there were no variations in the OCP implementation intensity, the interaction term between the cohort and the OCP enforcement intensity, a difference-in-difference indicator, can be interpreted as the causal effect of the implementation of the OCP on family size. Therefore, to extract the exogenous variation in family size, we instrument family size with the interaction term between a cohort dummy variable and the EFR.

1.4.2 Empirical model

The number of living children is instrumented using the interaction term $EFR \times Young$:

$$nchild_{ijp} = \beta_1(EFR_j \times Young_i) + X_i\gamma_1 + C_j\delta_1 + (C_j \times Young_i)\eta_1 + \phi_p + v_{ijp} \qquad (1.1)$$

where $nchild_{ijp}$ is the number of living children of household i in prefecture j in province p; EFR_j is the excess fertility rate in prefecture j; $Young_i$ equals 1 if household i mother's age is below the median and equals 0 otherwise. X_i contains a set of household controls, including household age, household education level and a dummy variable that equals 1 if household only has one senior parent; C_j contains a set of prefecture controls we construct from the 1982 Chinese Census; $C_j \times Young_i$ is the interaction terms between prefecture controls and the young cohort dummy variable; ϕ_p is province fixed effects.

To estimate the causal relationship between family size and senior parents' outcomes, we estimate the following model on our sample:

$$y_{ijp} = \beta_2 \widehat{nchild}_{ijp} + X_i\gamma_2 + C_j\delta_2 + (C_j \times Young_i)\eta_2 + \phi_p + \epsilon_{ijp} \tag{1.2}$$

where y_{ijp} is the senior parent's outcome variable of interest; $\widehat{nchild}_{ijp}$ is the estimated number of living children using Equation (1.1); the other variables are the same as defined in Equation (1.1).

1.4.3 Validity of IV

The first requirement for a valid instrumental variable is that the variable needs to be highly correlated with the endogenous variable. In our case, $EFR \times Young$ needs to be significantly correlated with family size. We report the first stage estimation results in Table 1.2 column (1). The coefficient on $EFR \times Young$ is 4.01 and it is significant at 1%, indicating that one percentage increase in the EFR significantly increases family size by 0.041. Our estimated coefficient is very similar to that from Li and Zhang (2017). They find the effect of the EFR is 0.04 for households with their first birth being a boy and 0.057 for households with their first birth being a girl. Our first stage results indicate that a more relaxed OCP enforcement does have a significant impact on family size, meaning our instrumental variable meets the requirement of being highly correlated with the endogenous variable.

To deliver meaningful results, it also requires our instrumental variable to be uncorrelated with prefecture-specific characteristics that potentially affect transfers from children to senior parents. Our instrumental variable will be invalid if variations in the EFR are associated with differences in pre-existing fertility preferences and socio-economic characteristics. We address this concern by including the interaction terms between prefecture characteristics and the young cohort dummy variable, $C_j \times Young_i$, to net out effects attributable to differences in pre-existing prefecture characteristics.

In addition, we examine the coefficient on $EFR_j \times Young_i$ among two subgroups, A and B. Subgroup A consists of households with mother aged 69 and over. Subgroup B includes with households with mother aged below 69. If stricter prefectures experiencing larger declines in family size is due to some prefecture characteristics, then the effect of $EFR_j \times Young_i$ should be relatively persistent in both subgroups, since both subgroups are subject to the same prefecture characteristics. However, if stricter prefectures experiencing larger declines in family size is indeed caused by the variation in the OCP enforcement intensity, then we should observe that the effect of $EFR_j \times Young_i$ differs between these two subgroups. A woman's best reproductive years are in her 20s. When the OCP was officially implemented in 1979, mothers in Subgroup A had already stepped into their 30s, meaning that their most fertile years were not covered by the OCP. Therefore, variation in the OCP enforcement strictness should have little influence on the fertility difference between the relatively young households and relatively old households within this subgroup. In contrast, mothers in Subgroup B were much younger when the OCP was introduced. Younger mothers in Subgroup B were entirely exposed to the OCP during their 20s, while older mothers in this subgroup were only partially subject to the OCP. Therefore, variation in the OCP enforcement intensity should have a noticeable effect on the fertility difference between the relatively young households and relatively old households within this subgroup and the effect should be smaller in magnitude compared to the estimated effect in Table 1.2 column (1). Results of estimating Equation (1.1) using Subgroup A and Subgroup B are reported in Table 1.2 column (2) and column (3), respectively. As expected, the coefficient on $EFR_j \times Young_i$ is not significant for Subgroup A at any conventional significance level, but it is significant at 10% in Subgroup B with a smaller magnitude. We take these results as evidence in favor of a causal relationship between the EFR and family size.

Another concern is that variations in stringency of the OCP implementation might be correlated with prefecture-specific shocks, such as the expansion of public pension system. Public pension serves as a substitute for children's support in one's old age, which reduces the importance of children being an old age insurance to their parents, hence decreases the demand for children [3]. Parents might adjust their fertility decision based on the availability of the public pension or the expectation of the availability of the public pension. If the expansion of public pension system is correlated with the OCP implementation stringency, then our instrumental variable will be invalid. We argue that the fertility decisions of households in our sample were unlikely to be influenced by the expansion of public pension system. The first major urban pension system expansion took place in 1999, followed by another big expansion in 2011 (Fang and Feng 2018). In rural area, the National Rural Pension Scheme (NRPS) was launched in 2009 (Ning et al. 2016). In our data, over 90% of the children were born before 1990, almost 10 yeas before the first major urban pension system reform and almost 20 years before the launch of the NRPS. This implies that pension system expansions did not take place during the fertile years of our sampled households. The impact of the pension system expansion on households' fertility is limited. Even if forward looking households might reduce their fertility due to the expectation of the potential pension expansion, given the majority of children in our data were born in the 1980s and earlier, it would be doubtful that parents could foresee the dramatic change in pension system that happened decades after and adjusted their family size accordingly (Cai et al. 2006).

Our identification strategy hinges on the assumption that fertility is the only channel for local OCP enforcement intensity to affect transfers to parents. Empirical studies find that strict OCP implementation also resulted in high sex ratios (Chen et al. 2013, Ebenstein 2010), indicating that the OCP implementation stringency may not only affect family size but may also affect child sex composition. This is concerning to our identification because gender of child has been found to have an impact on the upstream intergenerational transfers (Yang 1996, Lei 2013, Xie and Zhu 2009). If differences in local OCP enforcement intensity not only result in the variations in fertility decline, but also lead to the changes in child sex composition, then our estimates will be biased since child sex composition has its own effect on old-age support. To address this concern, we examine the relationship between child sex

[3] Shen et al. (2020) find a negative effect of China's expansion of New Rural Pension Scheme on the fertility.

composition and $EFR_j \times Young_i$ by using the following regression:

$$\begin{aligned} SexComposition_{ijp} = \beta_3(EFR_j \times Young_i) + X_i\gamma_3 + C_j\delta_3 \\ +(C_j \times Young_i)\eta_3 + \phi_p + u_{ijp} \end{aligned} \tag{1.3}$$

The specification of Equation (1.3) is the same as that of Equation (1.1) except that the dependent variable is a measure of child sex composition. We use the ratio of the number of living daughter to the number of living children as the measure for child sex composition. Results in Table 1.2 column (4) show that the coefficient on $EFR_j \times Young_i$ is not significant, suggesting that our identification will not be confounded by child sex composition.

Finally, we need to be careful when we interpret our findings because the causal effect estimated by the instrumental variable approach is the average treatment effect of compliers if the treatment effects are heterogeneous (Angrist and Imbens 1995). In our context, compliers are households whose fertility decisions would be shifted away from their desired fertility if they were affected by the implementation of the OCP and households who were subject to a stricter OCP experienced a greater decrease from their desired fertility. If the relationship between family size and transfer among these households differs from the relationship among the general population, then our estimated effect will likely be different from the average treatment effect of the general population. To learn more about the compliers that we are making inferences about, we follow Marbach and Hangartner (2020) to compare compliers with the general population. This method imposes two assumptions. The first assumption is that there are no defiers and the second consumption is that the instrument is independently assigned. Both assumptions are suitable in our context because the implementation of the OCP is an exogenous shock to all individuals and individuals would have had more children if they were not subject to the OCP. Since the instrument is independently assigned, covariate means for always-takers and never-takers can be estimated using covariate means for observable always-takers and never-takers [4]. Then covariate mean for compliers can be backed out by subtracting the weighted covariate mean of observable always-takes and weighted covariate mean of observable never-takes from the covariate mean of the sample.

The method proposed in Marbach and Hangartner (2020) requires a binary treatment and a binary instrument. For instrumental variable, we first construct a dummy variable, d_EFR, which takes on value 1 if EFR is above the median. We then create an interaction term between d_EFR and a dummy variable for young cohort, $Young$, to serve as our binary instrumental variable. For treatment variable, we first construct a dummy variable, Two, that takes on value 1 if the household has 2 or more living children. Since our identification strategy relies on a difference-in-difference indicator, we obtain residuals from a simple OLS regression where Two is regressed on d_EFR and $Young$ to control for direct effect of the OCP and cohort effect. Our binary treatment variable, Two', takes on value 1 if the residual is above the median. The characteristic comparison between sample and compliers is presented in Table 1.3. We find that the compliers in our analysis are older, less educated, and have less income compared to the entire sample.

1.5 Results

1.5.1 Effect of family size on transfers

Table 1.4 displays the main results of how family size affects transfers. Panel A reports the effect on whether parents received transfers in the previous year. The first column suggests that there is no evidence that family size affects whether parents received any transfers. Through columns (2) - (6), we decompose transfers into smaller categories: regular transfer, money transfer, regular money transfer, in-kind transfer and regular in-kind transfer. The

[4] Observable always-takes are individuals in control group who take the treatment. Observable never-takes are individuals in treatment group who do not take the treatment.

results suggest that family size do not affect whether parents received these different types of transfers. Overall, Panel A shows that there is no effect of family size on the probability of parents receiving transfers from their children.

Panel B looks into how the amount of transfers parents received in the previous year changes in response to family size. We use IV-Tobit estimation because 25% of parents in our sample did not receive any financial support from children. All IV-Tobit estimates correspond to the marginal effect of number of living children on the expected transfers from children while accounting for left censoring. Column (1) considers the amount of total transfers and columns (2) - (6) investigate the amount of transfers of different smaller categories. Although the coefficients are sizable in magnitude compared to average amount of transfers reported in Table **??**, they are not significant at any conventional level. The general implications derived from panel B are consistent with those from panel A, suggesting that there is no effect of family size on transfers parents received from their children.

1.5.2 Size of family support network

In Table 1.5, we display the mean number of children who provide transfers to their parents by family size. The mean number of children who provide financial support to their parents is 0.643, 1.165, 1.943, 2.656, 3.468, 4.607 for households with 1 child, 2 children, 3 children, 4 children, 5 children, 6 and more children, respectively. This pattern remains for sub-categorical transfers: regular transfers, money transfers, regular money transfers, in-kind transfers and regular in-kind transfers. Results in Table 1.5 suggest that parents with more children do have a larger family support network.

1.5.3 Payment for child care

Table 1.6 columns (1) and (2) show how family size affects parents offering care for their grandchildren. We find that parents with more children are more likely to provide care for their grandchildren and they tend to spend more time on providing child care. Having an additional child increases the probability of parents caring for their grandchildren by 39.4% and increases the time parents spend on looking after grandchildren by 5.5 hours per day. This finding suggests that the effect of family size on parents providing child care is statistical significant and sizable. If child care is the main channel that motivates the transfers from children to their parents, then parents with more children should receive more transfers as they offer more help with child care. However, this is inconsistent with our findings from Table 1.4, implying that parents providing child care is unlikely to be the main drive behind the relationship between family size and transfers from children to their parents.

1.5.4 Financial need

Altruism theory (Becker 1974) suggests that transfers from children to parents are motivated by parental financial need. To explore this potential channel, columns (3) - (4) look into the relationship between family size and parental wellbeing, under the assumption that parents with worse wellbeing have a higher need for transfers from children. It is challenging to accurately measure parental wellbeing. We attempt to use the following four variables to measure parental wellbeing: the value of per capita asset, whether at least one parent is working, whether both parents are working, and the value of per capital consumption. Asset includes cash, checking, savings, stocks, mutual funds, government bonds and other savings such as public housing fund. We choose asset as one of the measures because it captures whether parents are well-prepared for their old age. Parents who are not prepared for their old age might need more financial help from their children. Parental employment status reveals whether parents need other sources of income other than transfers from children. Consumption is another important life quality indicator. It includes consumption on food and non-food goods such as utilities and entertainment.

Results from columns (3) - (4) suggest that family size does not have an effect on parental asset holdings, employment status, or consumption, indicating that parental wellbeing does not significantly change with family size. It is worth mentioning that our parental wellbeing measures might not only capture parental life quality but also capture the effect of transfers from children on parental wellbeing. It is possible that parents save part of the transfers they received from their children and use part of the transfers on daily consumption. It is also possible that parental employment decisions are affected by the amount of transfers they received from their children. Since we do not find a significant effect of family size on transfers in Table 1.4, the potential effect of transfers on parental wellbeing is less worrisome. Our findings show that parents with more children do not face a different amount of financial need than those with fewer children, which provides evidence of altruism theory to explain the results in Table 1.4.

1.5.5 Quantity-quality trade-off

Table 1.7 reports the results of quantity-quality trade-off. To quantify children's quality, we use the following educational measures: upper and lower bounds of child's years of schooling [5], whether graduated from middle school and whether graduated from high shcool. In columns (1) - (4), we use child level data from CHARLS to explore whether a larger family size affects child's educational attainment. In columns (5) - (8), we use household level data from CHARLS to analyze the relationship between family size and children's average educational attainment in a household.

Consistent with previous studies (Rosenzweig and Zhang 2009, Li and Zhang 2017), our findings suggest a negative effect of family size on children's quality. An increase in number of children is associated with about 4-year-decrease in the upper bound and the lower bound of children's years of schooling. Having an additional sibling also decreases one's probability of graduating from middle school by about 40%. The negative effect is even larger on the one's probability of graduating from high school. An increase in family size leads to over 60%

[5] CHARLS only provides children's education level rather than children's years of schooling. We construct the lower and upper bounds of children's years of schooling based on their education level. For example, if a child's education level is middle school, then the lower bound of his years of schooling is 9 years because he needs at least 9 years of schooling to graduate from middle school, which includes 6 years from elementary school and 3 years from middle school. The upper bound of his years of schooling is 11 years because if he studied one more year he would have graduated from high school.

decrease in one's chance of graduating from high school. We now turn to the results from the household level data. Consistent with the results from the child level data, our findings suggest a larger family size is associated with a lower level of children's average educational attainment. Having an additional child decreases the average upper bound and the average lower bound of children's years of schooling by 3.6 years. An increase in family size also lowers the share of children who graduated from middle school by about 30% and lowers the share of children who graduated from high school by 60%. Results in Table 1.7 provides the evidence of the quantity-quality trade-off. Our results suggest that the substantial quantity-quality trade-off attenuates the positive effect of family size on transfers from children to their parents, which explains our findings of a insignificant effect of family size on transfers from children to their parents. This finding also provides support for exchange theory.

1.5.6 Other forms of elderly support

Elderly support takes different forms. Even though we do not find any effect of family size on transfers, family size might play an important role in other forms of elderly support. In this section, we look beyond upstream intergenerational transfers and look into how family size affects other forms of elderly support.

co-residence with parents is a common living arrangement for children to provide daily support for their parents. In our sample, 38.5% of parents co-reside with their children. Parents who do not live with their children often have children living nearby (Zimmer and Kwong 2003; Lei et al. 2015). This kind of living arrangement provides both elderly support and privacy. In our sample, 55% of the households have at least one child who lives in the same city or county. Visits and contacts are also important forms of elderly support because children's companionship is crucial for parental mental health (Bures et al. 2009).

Table 1.8 presents how family size affects living arrangements, frequency of children visiting and contacting their parents. Columns (1) - (2) consider the effect on co-residence. While an increase in family size does not affect the probability of co-residing with children, it increases the time parents co-reside with their children by 5 months. Column (3) suggests that having one more child increases the chance of having at least one child living nearby by 25.9%. Columns (4) - (11) explore the relationship between family size and children's visits and contacts. We highlight two findings from the results. First, family size does not have a significant effect on children's frequency of contacting parents or the number of children who contact parents frequently. Second, An increase in family size increases the probability of parents receiving monthly visit from their children by 29.4%. However, the number of children who pay monthly visit to their parents does not change with family size. Our findings in Table 1.8 are consistent with previous studies (Oliveira 2016; Chen and Fang 2018), which can be potentially explained by quantity-quality trade-off. Children from a larger family tend to have lower educational attainment, hence less likely to migrate (Zhao 1997). Thus, they have lower opportunity cost to pay monthly visit to their parents.

1.5.7 Rural vs. Urban

The coverage of pension system differs dramatically between rural and urban areas. In urban areas, individuals who work in the public sectors and state-owned enterprises (SOEs) have been covered by the pension program since the 1950s. Workers in non-state sectors and migrant workers started being included in the pension program in 1999. This high-benefit Employees' Basic Pension Program combines social pooling with individual account (Zhu and Walker 2018). Employers are required to pay 20% of employees' wage to contribute to the pooled social trust and employees pay 8% of their wage to contribute to their individual accounts. The pension employees receive after retirement consists two parts: general basic pension and individual account pension. General basic pension is determined by average wage of local employees and years of employment. Individual account pension is determined by the total amount of money accumulated over years of employment (Wang et al. 2014). Starting from 2011, urban individuals who are not eligible for enrolling in Employees' Basic Pension Program can voluntarily participate in low-benefit Urban Residents' Basic Pension

Program (Fang and Feng 2018). In this program, the central government contributes a monthly payment of 55 RMB and the local government provides a monthly subsidy of no less than 30 RMB to each individual pension account. Individuals can contribute 100 to 1000 RMB every year to their pension account. The amount of pension an individual receives depends on the total value of their pension account (Wang et al. 2014).

Pension system in rural areas was introduced relatively late. The pilot programs of public pension in rural areas started in the mid 1980s, however, those programs were halted due to low take-up rates and low effectiveness of public financial support (Shen et al. 2020). Then in 2009, the Chinese government launched the National Rural Pension Scheme (NRPS) and it has been expanded to nearly all rural areas in 2012. This pension system is identical to Urban Residents' Basic Pension Program, except individual contributions are set between 100 to 500 Yuan per year (Wang et al. 2014).

In 2011, the average monthly pension for Employees' Basic Pension Program participants, Urban Residents' Basic Pension Program, NRPS participants is 1558 RMB, 78 RMB, and 57.5 RMB, respectively (Wang et al. 2014). Individuals who live in urban areas are more likely to have higher education and work in formal sectors. As a result, they are more likely to participate in the high-benefit pension program. In contrast, rural individuals tend to participate in the low-benefit pension program (Hanewald et al. 2021). Our data lends support to this. The average yearly pension income of urban households is 12052 RMB while the average yearly pension income of rural households is only 913 RMB. The great difference in pension income between urban individuals and rural individuals may affect the importance of children in elderly support and may further influence how transfers from children to parents change in response to family size.

To account for the possible differences in the effect of family size on transfer between urban parents and rural parents, we estimate the following equation:

$$y_{ijp} = \beta_4 \widehat{nchild_{ijp}} + \alpha_4 \widehat{nchild_{ijp} \times rural_i} + X_i\gamma_4 + C_j\delta_4 \\ +(C_j \times Young_i)\eta_4 + \phi_p + \epsilon_{ijp} \tag{1.4}$$

where $rural_i$ is an indicator that takes on value 1 if the household lives in rural area, and 0 if the household lives in urban area. The other variables are the sames as in Equation (1.2). The first stage equations are modified as follows:

$$nchild_{ijp} = \beta_5(EFR_j \times Young_i) + \alpha_5(EFR_j \times Young_i \times rural_i) \\ +X_i\gamma_5 + C_j\delta_5 + (C_j \times Young_i)\eta_5 + \phi_p + v_{ijp} \tag{1.5}$$

$$nchild_{ijp} \times rural_i = \beta_6(EFR_j \times Young_i) + \alpha_6(EFR_j \times Young_i \\ \times rural_i) + X_i\gamma_6 + C_j\delta_6 + (C_j \times Young_i)\eta_6 + \phi_p + w_{ijp} \tag{1.6}$$

Table 1.9 displays the results. Column (1) in panel A suggests that, compared to urban parents, having an additional child increases the probability of parents receiving transfers by 6.3% more. Column (2) - (6) in panel A explore the difference in effect of family size on different types of transfers. We find that an increase in family size increases the probability of parents receiving money transfer, in-kind transfers, regular in-kind transfers by 7.49% more, 5.2% more, 2.65% more, respectively. Panel B estimates the difference in effect of family size on the amount of transfers parents received in the previous year between urban households and rural households. Compared to urban households, family size has a large positive impact on the amount of transfers parents received. Rural parents receive 570 RMB more transfers if there is an increase in their family size. The difference in effect is large in magnitude.

It is more than half of the average pension income of rural households. Columns (2) - (7) looks into the difference in effect on transfers of smaller categories. The results suggest that rural households experience a larger effect of family size on both money transfers and in-kind transfers compared to urban counterparts.

We further examine whether the difference in the effects of family size on transfers between rural and urban households can be explained by the disparity in quantity-quality trade-off and parental financial need. Quantity-quality trade-off is unlikely to be the channel behind the difference in effect among rural and urban households. Previous studies often suggest a stronger quantity-quality trade-off among rural households (Li et al. 2008), which will lead to a negative coefficient on the interaction term between family size and rural status. This is inconsistent with our findings in Table 1.9. Although we do not find evidence of a stronger quantity-quality trade-off among rural households using our sample in Panel A of Table 1.10, the insignificance of coefficients on interaction term between family size and rural status suggests that there is no difference in quantity-quality trade-off between rural and urban households, which also implies that quantity-quality trade-off is unlikely to be the channel. Another possible explanation for the difference in the effects of family size on transfers between rural and urban households is the difference in financial need. As we discussed earlier, rural parents receive a much smaller amount of pension compared to urban parents, implying that rural parents might have a greater financial need. Results in panel B of Table 1.10 lend support to this conjecture. Rural parents are more likely to be working and have a lower consumption level, suggesting they need other sources of income other than pension income. Therefore, rural parents receiving more transfers from children might be because they need more, which is consistent with altruism theory.

1.5.8 Migration

Our identification exploits the local OCP enforcement stringency. Our analysis is conducted under the assumption that parents' residence of prefecture at the time of the interview is the same as their residence of prefecture when they gave birth to their children. Our results will be confounded if there was massive migration taking place among the households in our sample. Some parents might have moved after the births of their children, which introduces measurement error to our analysis. Moreover, some of the migrations might be endogenous. Parents might move to places where they can obtain better elderly support and this will bias our results.

Since the migration information after giving birth is not available in the CHARLS, we use the information on senior parents' birth places to construct a subsample. The subsample consists of households with at least one elderly parent whose birth prefecture is the same as his/her residence of prefecture during the CHARLS Wave 1 and Wave 4 survey interviews. The assumption we impose here is that parents who still live in their birth prefecture during the CHARLS interviews are more likely to stay in their birth prefecture their whole life. We replicate Table 1.4 using this subsample and display the new results in Table 1.11. Despite differences in magnitude, all coefficients remain insignificant, suggesting that there is no significant effect of family size on the probability of parents receiving transfers or the amount of transfer parents receive from their children. The findings using this subsample is consistent with those in Table 1.4, indicating that there is no evidence showing our estimates are driven by potentail migration in our dataset.

1.5.9 External validity comparative analysis

The closest study to this paper is Oliveira (2016). She finds that an increase in family size increases both the probability of receiving transfers from children and the amount of transfers from children. The consistency between her results and ours might be explained by the heterogeneity in effect of family size on transfers. Oliveira (2016) utilizes the incidence of twins in the first birth as an instrumental variable to establish a causal relationship between family size and transfers. Since the IV approach delivers the average effect of a group of compliers if the effect is heterogeneous, it is possible that compliers in Oliveira (2016) are very different compared to those in this paper.

To explore whether differences in compliers can potentially explain the inconsistency of our results with the literature, we attempt to profile the compliers generated by using the incidence of twins in the first birth. We use Two, a dummy variable equals 1 if the household has twins in their first birth, as the binary treatment. The binary treatment variable is a dummy variable, $Twin$, that takes on value 1 if the household has twins in their first birth. Covariant means of the compliers are presented in Table 1.3. The mean household age of the compliers is 71, 9 years above the mean age of the whole sample. The mean parental education level is middle school for compliers but only primary school for the whole sample. The income of the compliers is over twice as much as that of the whole sample. Overall, the compliers represent an older subpopulation with better education and higher income.

There are two consequences of the number of children a couple has on the amount of transfers they may receive in their old age. On one hand, the amount of transfers may grow in the number of children. This effect may be attenuated if there is a significant negative quantity-quality trade-off. One possibility in the divergence of our results and the consensus in the literature is that the instrument they use — the incidence of twins in the first birth, generates a complier group of wealthier couples. Within this group the quantity-quality trade-off is abated by a higher budget constraint. In this group the effect of more children means a higher number of transfers to the parents. On the other hand with an instrument that generates a complier group that is more reflective of the general population the group will face a tighter budget constraint. In this group increasing the number of children does not necessarily generate more transfers in the future since higher order birth children suffer in quality due to lower investments.

1.6 Conclusion

In this paper, we examine the causal relationship between family size and upstream intergenerational transfers using the data from the China Health and Retirement Longitudinal Study (CHARLS). We exploit the variations in local OCP enforcement stringency to address the endogeneity of fertility decisions. We find that family size does not have an effect on either the probability of receiving transfers from children or the amount of transfers from children. Our results suggest two possible explanations. The first one is that the benefit of having a larger elderly support network is diminished by the quantity-quality trade-off. The second one is that the financial need of parents of more children does not significantly differ from that of parents with fewer children. The response of transfer from children to an increase in family size is correlated with household's rural status. Compared to urban parents, having more children has a larger positive effect on transfers among rural parents, implying that children remain as an important source of elderly support for rural households. From a policy standpoint, recent relaxation of the OCP is likely to have a positive impact on rural senor parents' wellbeing. In addition, we explore the effect of family size on other forms of elderly support. We find that parents with more children are more likely to co-reside with children for more months, have a child living nearby, and have children paying monthly visits.

1.7 Tables

Table 1.1: Summary statistics

	Whole sample		Rural		Urban	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Number of living children	2.62	1.34	2.82	1.34	2.34	1.30
Mother's age	61.18	10.67	61.25	10.60	61.07	10.77
Household age	62.03	10.47	62.08	10.40	61.95	10.58
Household education level	2.14	0.80	1.95	0.69	2.41	0.87
Single	0.21		0.20		0.22	
Received financial support	0.75		0.80		0.67	
Amount of financial support received	4726.72	10767.54	4644.20	10746.43	4844.67	10797.94
Received regular financial support	0.19		0.21		0.18	
Amount of regular financial support received	693.15	3244.67	688.82	3075.33	699.34	3472.81
Received money support	0.53		0.59		0.46	
Amount of money support received	3253.70	9206.83	3318.27	9365.09	3161.41	8976.20
Received money support regularly	0.15		0.16		0.14	
Amount of regular money support received	546.15	2906.94	552.46	2714.15	537.12	3162.51
Received in-kind support	0.65		0.70		0.59	
Amount of in-kind support received	1473.02	3713.87	1325.94	3257.71	1683.25	4274.01
Received in-kind support regularly	0.10		0.11		0.09	
Amount of regular in-kind support received	147.00	941.13	136.35	854.30	162.23	1052.81
N	9314		5480		3834	

Notes: Data comes from China Health and Retirement Longitudinal Study (CHARLS) Wave 4. Household age is the average age of respondent and spouse. If a household does not have a spouse, then its household age is equal to the age of the respondent. Household education level is calculated in the same fashion as household age. Four education levels: 1, Nor formal education or illiterate; 2, Primary school; 3, Middle school; 4, High school and above. Single equals 1 if a household does not have a spouse; equals 0 if a household has both respondent and spouse. Financial support consists of money support and in-kind support. Regular money support includes providing living expenses, paying for utilities and other other forms of regular expense. In-kind support includes buying food, clothes and other goods regularly.

Table 1.2: IV validity

	Number of living children			Sex composition
	(1)	(2)	(3)	(4)
EFR × Young	4.010***	3.167	2.328*	0.484
	(1.258)	(3.567)	(1.267)	(0.411)
Control variables:				
Household controls	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
N	9314	2236	7078	9314

Notes: In column (1), the entire sample is used. In column (1), a subsample of households with mother aged greater than or equal to 69 is used. In column (3), a subsample of households with mother aged less than 69 is used. In column (4), the dependent variable, sex composition, is the ratio of the number of living daughter to the number of living children. EFR × Young refers to the interaction between the EFR and the dummy variable for young cohort. A household belongs to the young cohort if mother's age is below median mother's age in the whole sample. Household controls and prefecture controls are the same in all columns. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.3: Complier characteristics

	sample	complier	
		$d\_EFR * Young$	$Twin$
Household age	62	67.23	71.02
Mother's education level	1.95	1.72	2.93
Father's education level	2.46	2.37	3.11
Per capita income	12298.85	7677.54	33455.26
Per capita consumption	18300.35	16403.35	36115.23
Rural	0.59	0.63	0.55
Compliance rate		63.51%	16.64%

Notes: This table reports sample covariates means and complier covariate means. d_EFR takes on value 1 if EFR is above the median. $Young$ is a dummy variable for young cohort. A household belongs to the young cohort if mother's age is below median mother's age in the whole sample. $Twin$ takes on value 1 if the household has twins in their first birth. Household age is the average age of respondent and spouse. If a household does not have a spouse, then its household age is equal to the age of the respondent. Four education levels: 1, Nor formal education or illiterate; 2, Primary school; 3, Middle school; 4, High school and above. Rural is a dummy variable for household rural status.

Table 1.4: Number of living children and transfers

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
	Received transfer	Received regular transfer	Received money transfer	Received regular money transfer	Received in-kind transfer	Received regular in-kind transfer
number of living children	0.0878	0.0212	-0.147	0.127	0.0977	-0.111
	(0.125)	(0.119)	(0.159)	(0.111)	(0.139)	(0.101)
Panel B						
	Total transfer	Total regular transfer	Money transfer	Regular money transfer	In-kind transfer	Regular in-kind transfer
number of living children	752.9196	-377.4116	-472.0277	459.1821	129.7623	-650.3984
	(1740.5145)	(963.0050)	(1653.5686)	(1009.7945)	(609.6669)	(602.5173)
Control variables:						
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	9314	9314	9314	9314	9314	9314

Notes: In all columns of Panel A and Panel B, the independent variable of interest is the number of living children, and it is instrumented by the interaction term EFR × Young. In Panel A, all columns use 2SLS estimation. In Panel B, all columns use IV-Tobit estimation. The IV-Tobit estimates correspond to the marginal effect of regressors on the expected transfers from children while accounting for left censoring. Household controls and prefecture controls are the same in all columns in Panel A and Panel B. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 1.5: Family size and mean number of children who provide transfers to their parents

	Mean number of children who provide transfers					
Number of children	Transfer	Regular transfer	Money transfer	Regular money transfer	In-kind transfer	Regular in-kind transfer
1	0.643	0.120	0.313	0.069	0.556	0.071
2	1.165	0.231	0.707	0.160	0.967	0.119
3	1.942	0.388	1.299	0.282	1.540	0.200
4	2.656	0.605	1.860	0.474	2.021	0.276
5	3.468	0.921	2.611	0.737	2.566	0.439
6+	4.607	1.199	3.429	1.009	3.098	0.497

Notes: In this table, we report the mean number of children who provide transfers by family size. Transfer consists of money transfer and in-kind transfer. Regular money transfer includes providing living expenses, paying for utilities and other other forms of regular expense. In-kind transfer includes buying food, clothes and other goods regularly.

Table 1.6: Number of living children and other outcomes

	Hours of grandchildren care	Caring for grandchildren	Asset	One works	Both work	Consumption
	(1)	(2)	(3)	(4)	(5)	(6)
number of living children	1983.5*	0.394**	-1.573	-0.0576	0.121	-0.380
	(1202.6)	(0.199)	(1.025)	(0.114)	(0.132)	(0.274)
First stage: number of children						
EFR × Young	4.010***	4.010***	4.161***	4.010***	4.010***	4.664***
	(1.258)	(1.258)	(1.358)	(1.258)	(1.258)	(1.351)
Control variables:						
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	9314	9314	8141	9314	9314	7688

Notes: All columns report 2SLS estimates. The independent variable of interest is the number of living children, and it is instrumented by the interaction term EFR × Young. Dependent variable in column (1) is total hours parents spent on caring for grandchildren in the past year. In column (2), it is a dummy variable that equals 1 if parents provided care for grandchildren in the past year. The dependent variable in column (3) is log(household per capita asset + 1). Asset includes cash, checking, savings, stocks, mutual funds, government bonds and other savings such as public housing fund. In column (4), the dependent variable is a dummy variable that equals 1 if at leat one parent in the household works. In column (5), the dependent variable is a dummy variable that equals 1 if both parents work. Dependent variable in column (6) is log(household per capita consumption). Consumption includes consumption on food and non-food goods such as utilities and entertainment. Household controls and prefecture controls are the same in all columns. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Differences in sample size are due to missing data on asset and consumption. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.7: Number of living children and children's human capital

	Child				Household			
	Educ lower bound	Educ upper bound	Finished middel school	Finished high school	Avg. Educ lower bound	Avg. Educ upper bound	Middle school graduation rate	High school graduation rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
number of living children	-4.095***	-4.035***	-0.393***	-0.637***	-3.631***	-3.631***	-0.299***	-0.602***
	(1.188)	(1.148)	(0.124)	(0.169)	(1.181)	(1.152)	(0.113)	(0.180)
First stage: number of children								
EFR × Young	4.213***	4.213***	4.213***	4.213***	4.713***	4.713***	4.713***	4.713***
	(1.019)	(1.019)	(1.019)	(1.019)	(1.313)	(1.313)	(1.313)	(1.313)
Control variables:								
Household controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	22624	22624	22624	22624	8706	8706	8706	8706

Notes: All columns report 2SLS estimates. The independent variable of interest is the number of living children, and it is instrumented by the interaction term EFR × Young. Column (1) - (4) use the child level data from CHARLS. Dependent variables in column (1) and (2) are the lower bound and upper bound of child's years of schooling, respectively. Dependent variables in column (3) and (4) are dummy variables for middle school graduation and high school graduation, respectively. Column (5) - (8) use the household data from CHARLS. The smaller sample size is due to missing data for some children. Dependent variables in column (5) and (6) are the average lower bound and average upper bound of children's years of schooling, respectively. Dependent variables in column (7) and (8) are children's middle school graduation rate and high school graduation rate, respectively. Household controls and prefecture controls are the same in all columns. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.8: Number of living children and other types of old-age support

	Co-reside (months)	Co-reside	Live nearby	Weekly visit	Monthly visit	Number of children visit weekly	Number of children visit monthly	Weekly contact	Monthly contact	Number of children contact weekly	Number of children contact monthly
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
number of living children	5.056*	0.0901	0.259*	0.141	0.294*	0.226	0.367	0.0494	0.0794	0.478	0.465
	(2.580)	(0.149)	(0.153)	(0.154)	(0.163)	(0.272)	(0.308)	(0.183)	(0.128)	(0.384)	(0.382)
First stage: number of living children											
EFR × Young	4.010***	4.010***	4.320***	4.107***	4.107***	4.107***	4.107***	3.691**	3.691**	3.691**	3.691**
	(1.258)	(1.258)	(1.269)	(1.355)	(1.355)	(1.355)	(1.355)	(1.520)	(1.520)	(1.520)	(1.520)
Control variables:											
Household controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9314	9314	9208	8458	8458	8458	8458	7102	7102	7102	7102

Notes: All columns report 2SLS estimates. The independent variable of interest is the number of living children, and it is instrumented by the interaction term EFR × Young. Dependent variable in column (1) is total number of months parents co-resided with their children in the past year. In column (2), it is a dummy variable which equals 1 if parents co-resided with at least one of their children in the past year. The dependent variable in column (3) is a dummy variable equals 1 if parents have at least one child lives nearby. In column (4), the dependent variable is a dummy variables that equals 1 if at least one child paid weekly visit. In column (5), the dependent variable is a dummy variable that equals 1 if at least one child paid monthly visit. In column (6) and (7), dependent variables are the number of children who paid weekly visit to their parents and number of children who paid monthly visit to their children in the past year, respectively. Dependent variables in column (8) - (11) are defined in the same fashion as those in column (4) - (5). Household controls and prefecture controls are the same in all columns. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Differences in sample size are due to missing data on children's visits and contacts. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.9: Rural vs. urban: number of living children and transfers

	(1) Received transfer	(2) Received regular transfer	(3) Received money transfer	(4) Received regular money transfer	(5) Received in-kind transfer	(6) Received regular in-kind transfer
Panel A						
number of living children	0.0168	0.00634	-0.232	0.136	0.0391	-0.141
	(0.141)	(0.134)	(0.186)	(0.125)	(0.157)	(0.116)
number of living children × rural	0.0630***	0.0132	0.0749***	-0.00769	0.0520***	0.0265**
	(0.0158)	(0.0149)	(0.0208)	(0.0139)	(0.0176)	(0.0129)
Panel B						
	Total transfer	Total regular transfer	Money transfer	Regular money transfer	In-kind transfer	Regular in-kind transfer
number of living children	-3.3575	-604.7675	-1310.269	420.3270	-43.0539	-858.3038
	(1968.6732)	(1712.8123)	(1953.7514)	(1127.8470)	(691.2225)	(3082.1191)
number of living children × rural	570.4943***	194.5446	645.6636***	31.7460	137.6892*	165.8996
	(218.5352)	(436.2057)	(229.1505)	(120.1914)	(77.2996)	(585.3003)
Control variables:						
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	9314	9314	9314	9314	9314	9314

Notes: In all columns of Panel A and Panel B, Number of living children and the interaction between the number of living children and an indicator of living in rural area are instrumented using EFR × Young and EFR × Young × rural. In Panel A, all columns use 2SLS estimation. In Panel B, all columns use IV-Tobit estimation. The IV-Tobit estimates correspond to the marginal effect of regressors on the expected transfers from children while accounting for left censoring. Household controls and prefecture controls are the same in all columns in Panel A and Panel B. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 1.10: Rural vs. urban: number of living children and other outcomes

	(1)	(2)	(3)	(4)
Panel A				
	Avg. Educ lower bound	Avg. Educ upper bound	Middle school graduation rate	High school graduation rate
number of living children	-3.590***	-3.601***	-0.309**	-0.611***
	(1.308)	(1.276)	(0.126)	(0.201)
number of living children × rural	-0.0391	-0.0285	0.00920	0.00853
	(0.149)	(0.145)	(0.0143)	(0.0228)
N	8706	8706	8706	8706
Panel B				
	Asset	One works	Both work	Consumption
number of living children	-1.385	-0.107	0.0517	-0.313
	(1.141)	(0.128)	(0.146)	(0.296)
number of living children × rural	-0.156	0.0436***	0.0616***	-0.0706**
	(0.124)	(0.0143)	(0.0163)	(0.0317)
N	8141	9314	9314	7688
Control variables:				
Household controls	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes

Notes: All columns report 2SLS estimates. In all columns of Panel A and Panel B, Number of living children and the interaction between the number of living children and an indicator of living in rural area are instrumented using EFR × Young and EFR × Young × rural. Household controls and prefecture controls are the same in all columns. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Differences in sample size are due to missing data on asset and consumption. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 1.11: Number of living children and transfers using subsample data

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
	Received transfer	Received regular transfer	Received money transfer	Received money transfer	Received in-kind transfer	Received regular in-kind transfer
number of living children	0.113	-0.0211	0.0148	0.0906	0.137	-0.0793
	(0.129)	(0.138)	(0.157)	(0.125)	(0.152)	(0.113)
Panel B						
	Total transfer	Total regular transfer	Money transfer	Regular money transfer	In-kind transfer	Regular in-kind transfer
number of living children	-304.9721	-669.0623	-847.5503	-98.2500	378.8962	-248.2741
	(1913.3250)	(1101.8032)	(1828.8312)	(1024.0360)	(658.7499)	(450.8316)
Control variables:						
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture controls	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	6237	6237	6237	6237	6237	6237

Notes: The subsample used in this table consists of households with at least one elderly parent whose birth prefecture is the same as his/her residence of prefecture during the CHARLS Wave 1 and 4 survey interviews. In all columns of Panel A and Panel B, the independent variable of interest is the number of living children, and it is instrumented by the interaction term EFR × Young. In Panel A, all columns use 2SLS estimation. In Panel B, all columns use IV-Tobit estimation. The IV-Tobit estimates correspond to the marginal effect of regressors on the expected transfers from children while accounting for left censoring. Household controls and prefecture controls are the same in all columns in Panel A and Panel B. Household controls are: household age, household education level, a dummy variable equals 1 if household only has one parent. Prefecture controls are: the average total number of births for females aged between 45 and 54; the shares of females aged between 25 and 44 with 1, 2, 3, and 4+ births, respectively; the share of females aged between 25 to 19; the share of females aged between 30-34; the share of each education level category among among adults aged between 25 and 49 by gender; the agricultural sector employment share among adults between 24 to 49 by gender. In all specifications, interactions between prefecture controls and young cohort dummy variable are included. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Chapter 2

Positive Assortative Mating Effect vs. Cross-Productivity Effect? Estimating the Causal Effect of Spousal Education on Individuals' Earnings

2.1 Introduction

China's income inequality is among the highest in the world (Jain-Chandra et al. 2018), and it has been linked to the positive assortative mating pattern in the marriage market (Nie and Xing 2019). A well-educated individual is more likely to marry another well-educated individual because they share similar traits (Benham 1974). This mating behavior not only leads to a larger human capital dispersion among households but also increases income inequality because well-educated couples are more likely to earn higher wages than ill-educated couples.

In addition to positive assortative mating, income inequality might be even more exacerbated by the cross-productivity effect within marriage among well-educated couples. A better-educated spouse can increase one's earnings by being more effective in specialization, sharing more productive knowledge, giving better advice (Benham 1974), making better decisions (referred as "allocative effect" in Welch 1970), and providing higher-quality networks. Therefore, a better-educated spouse can help an individual accumulate more human capital and increase earnings.

Understanding the two effects — mating effect and cross-productivity effect — not only sheds light on the process of human capital accumulation but also has important implications for income inequality in China. However, disentangling these two effects is not trivial. The cross-productivity effect is the causal effect of spousal education on wages, while the mating effect is merely a correlation. In the estimation of return to spousal education, the mating effect introduces endogeneity to the estimation.

The traditional way to filter out endogeneity is to use instrumental variables (IVs). However, finding good IVs is often challenging. Several approaches have been employed in the literature to address endogeneity. Welch (1974) attempted to use the husband's IQ and background variables to partial out the possible mating effect. The limitation of this approach is that it can not address endogeneity that is uncorrelated with IQ and background variables.

Benham (1974)), Wong (1986), and Neuman and Ziderman (1992)) dealt with the challenge by including the length of marriage in the regression. Since the mating effect is merely a trait match, it does not change over time. In contrast, the cross-productivity effect needs time to realize because it is a process of human capital accumulation. A longer marriage may not only provides more time for an individual to accumulate human capital from spousal education but also makes the process more efficient as communication costs decrease in marriage length. If the positive correlation changes with marriage length, one can conclude that the correlation is not only a mating effect. However, this approach can not examine whether the positive correlation is the causal cross-productivity effect or a mixture of both the cross-productivity effect and the mating effect.

Another approach that has been employed in the literature is to use between-twins variations to isolate the causal cross-productivity effect from the mating effect (Huang et al. 2009). Mono-zygotic twins are typically considered identical in terms of family background and unobservable characteristics such as ability since they come from the same family and share the same set of genes. Causal effects can be obtained by comparing spousal education and one's own earnings between twins. However, researchers have addressed their concerns about using a twin-based approach. First, between-twins estimation might not be able to fully eliminate the omitted variable bias if the productivity traits consist of more than just genes. Hence, estimators should be regarded as the upper bound of the return to spousal education (Huang et al. 2009; Neumark 1999; Bound and Solon 1999; Isacsson 2007). Second, comparing twins can eliminate much of the endogenous variation in education, but much of the exogenous variation in education are differenced out at the same time. Twin-based estimates are subject to as large an endogeneity inconsistency as cross-sectional estimates if endogenous variation accounts for much of the remaining between-twin variation (Griliches 1979; Neumark 1999; Bound and Solon 1999). Third, this method may deliver potentially imprecise because twins data is often small and contains limited variations.

In this paper, we use a control function approach developed by Klein and Vella (2010) to estimate the causal effect of spousal schooling on wages in China using three waves of data from the Chinese Household Income Project (CHIP). Our approach advances other approaches in the literature in the following ways. First, our approach does not require IVs. Second, our approach relies on the nonlinearity of error terms induced by heteroskedasticity to address endogeneity. The key assumption underlying our approach is the presence of heteroskedasticity, which can be statistically tested. Third, our approach allows us to utilize a sizable dataset with rich variations to deliver more precise estimates.

Our results from the control function approach suggest that spousal education has a positive effect on wages for both husbands and wives. An additional year of schooling received by the wife increases her husband's annual wage by 4.79%. An additional year of schooling received by the husband increases his wife's annual wage by 2.51%. Compared with our OLS results, we detect that there exists negative assortative mating on unobservables. Our also results suggest that there is substantial heterogeneity in return to spousal education by gender, spousal education level, year, dominant earner status, as well as whether spouses share the same occupation. Our findings suggest strong evidence of specialization and network effect being the potential channel underlying the positive relationship between spousal education and wages.

This paper contributes to the literature in the following ways. First, this paper joins the large body of the literature on spousal education and confirms the existence of the cross-productivity effect. Some of our results are comparable with the literature. For example, our estimate of the return to spousal education for wives (2.51%) is very similar to the estimate (3.3%) in Huang et al. (2009). In addition, we provide evidence that there also exists the cross-productivity effect for husbands, which has not been detected before in the Chinese context. Moreover, our estimates are potentially more precise. In our results, the estimates for return to own schooling remain significant, which alleviates the concern that our estimates for return to spousal education merely capture the return of own schooling.

Second, our findings that spousal education has a cross-productivity effect for both husbands and wives shed light on our understanding of the process of human capital accumulation. From the close association in marriage, the husband and wife can benefit from each other's education. On one hand, these positive externalities of education within marriage provide additional justifications for government intervention to improve education. On the other hand, the presence of the cross-productivity effect suggests that the positive assortative mating behavior plays an even larger role in income inequality. Government should provide more educational and vocational services to ill-educated couples to reduce income inequality.

Third, while our paper focuses on estimating the return to spousal schooling, our study also contributes to the literature on estimating the return to own schooling. Spousal education is often used as an instrumental variable in estimating the effect of own education on earnings (for example, Guifu and Hamori 2009, Wang 2013). This approach relies on two assumptions about spousal education. The first assumption is that spousal education is positively correlated with one's own education due to the positive mating effect. The second assumption is that spousal education does not have a direct impact on one's earnings, meaning there is no significant cross-productivity effect. Our results of the existence of the cross-productivity effect between husbands and wives indicate that spousal education violates the exclusion restriction of being an instrumental variable. Therefore, using spousal education as an instrumental variable to estimate the return to own education is problematic, and the relevant literature should be revisited.

2.2 Empirical Methodology

2.2.1 Basic Set-up

We begin by considering a simple model of education and wages as follows:

$$W_i = \alpha + \beta_1 S_i^{self} + \beta_2 S_i^{spouse} + X_{0i}\delta + u_i \tag{2.1}$$

$$S_i^{self} = X_{1i}\omega_1 + v_{1i} \tag{2.2}$$

$$S_i^{spouse} = X_{2i}\omega_2 + v_{2i} \tag{2.3}$$

where W_i measures labor market outcome. S_i^{self} is one's own years of schooling. S_i^{spouse} stands for spousal years of schooling. Vector X is a set of individual exogenous characteristics. β_2 is the parameter of interest, measuring the returns to spousal education. We refer to equation (2.1) as the wage equation, and refer to equation (2.2) and equation (2.3) as the education equations.

The source of endogeneity arises when u_i and v_{ci}, $c = 1, 2$ are correlated. u_i captures unobservable wage determinants such as ability. The correlation between u_i and v_{1i} comes from the concern that unobservable factors such as ability may not only influence an individual's wages but also affect their choices of education level. The correlation between u_i and v_{2i} emanates from the assortative mating effect. Abler individuals who have higher wages may marry better-educated individuals because they share similar traits. A typical way to address the endogeneity is to employ an instrumental variable (IV) approach. The traditional IV approach relies on finding variables that determine educational attainment but do not influence wages. However, detecting such variables is challenging.

To circumvent employing IVs, we adopt a control function approach proposed by Klein and Vella (2010). This method utilizes the nonlinearity in the control terms induced by heteroskedasticity to address the endogeneity. In particular, we consider the following error term heteroskedasticity structures for equations (2.1) - (2.3):

$$u_i = H_u(X_{0i})u_i^* \tag{2.4}$$

$$v_{1i} = H_{1v}(X_{1i})v_{1i}^* \tag{2.5}$$

$$v_{2i} = H_{2v}(X_{2i})v_{2i}^* \tag{2.6}$$

where u_i^*, v_{1i}^*, and v_{2i}^* are homoskedastic error terms. Functions $H_u(X_{0i})$, $H_{1v}(X_{1i})$, and $H_{2v}(X_{2i})$ capture heteroskedasticity. Klein and Vella (2010) show that equation (2.1) can be estimated via equation (2.7) if $\rho_1 \frac{H_{ui}(X_{0i})}{\hat{H}_{1vi}(X_{1i})}$ and $\rho_2 \frac{H_{ui}(X_{0i})}{\hat{H}_{2vi}(X_{2i})}$ are not constant. $\hat{v}_{1i}$, $\hat{H}_{1vi}(X_{1i})$, $\hat{v}_{2i}$, and $\hat{H}_{2vi}(X_{2i})$ are consistent estimates from equations (2.2) and (2.3). ρ stands for the correlation between homoskedastic errors from the wage equation and the education equation.

$$W_i = \alpha + \beta_1 S_i^{self} + \beta_2 S_i^{spouse} + X_{0i}\delta + \rho_1 \frac{H_{ui}(X_{0i})}{\hat{H}_{1vi}(X_{1i})}\hat{v}_{1i} + \rho_2 \frac{H_{ui}(X_{0i})}{\hat{H}_{2vi}(X_{2i})}\hat{v}_{2i} + e_i \qquad (2.7)$$

The identification of equation (2.7) requires two assumptions to be satisfied. First, at least one of $H_{ui}(X_{0i})$ and $H_{cvi}(X_{ci})$, $c = 1, 2$ varies across X. Second, correlation between u_i* and v_{ci}^*, $c = 1, 2$ is constant, i.e., $\mathbb{E}[u_i^* v_{1i}^*] = \rho_1$ and $\mathbb{E}[u_i^* v_{2i}^*] = \rho_2$. The first assumption is statistically testable, therefore we provide more discussion in the validity test section. Although the second assumption is not testable, we argue that it is a reasonable assumption in our context. u_i^*, v_{1i}^*, and v_{2i}^* can be considered as abilities associated with genes. Then ρ_1 measures the return to own education through ability that is determined by genes. Similarly, ρ_2 measures the return to spousal education through matching genetic traits in the marriage market. The constancy of ρ_1 and ρ_2 assumes that the return to education via the genetic channel is not affected by individual non-genetic characteristics. The presence of heteroskedasticity in the model allows the return to education via genes to scale up based on a heteroskedasticity function of individual characteristics.

2.2.2 Practical Implementation

Variables for X_{0i}, X_{1i}, and X_{2i}

Unlike the traditional IV approach, Klein and Vella (2010) method does not impose exclusion assumptions on X_{1i} and X_{2i}, which means that X_{1i} and X_{2i} do not have to be different from X_{0i}. In the practice of choosing variables for X_{1i} and X_{2i}, we try to include variables that are pre-determined before individuals choose their education level. These variables include demographic variables (such as birth cohort and ethnicity) and family background variables (such as parental education and parental occupation). For the choice of X_{0i}, we include fixed demographic variables that are crucial for determining both wages and education. We also include variables (such as work experience) that are commonly used in wage equations in the literature. However, we do not include variables that are potentially determined by education. Examples of this type of variables include individual's occupation and firm characteristics. On one hand, these variables are potentially endogenous. Including them may introduce bias to our estimates. On the other hand, these variables serve as channels through which education affects wages. The inclusion of this type of variables may absorb part of the effect of education, hence preventing our parameter of interest, β_2, from capturing the *total* effect of spousal education on wages.

Implementation of the Klein and Vella (2010) Approach

To reduce the computational cost, we parameterize the unknown heteroskedasticity function $H(X_i)$ as $\sqrt{exp(X_i\theta)}$. The exponential heteroskedasticity specification has been commonly used in econometrics since Harvey (1976) and has been employed in similar contexts as our study (for example, Farré et al. 2013, Millimet and Roy 2016, Chen et al. 2018). The exponential form of heteroskedasticity not only guarantees the non-negativity of the variance of the error term for all possible values of θ, but also provides a simple way to examine what variables in X contribute to the the heteroskedasticity. We use the error term from equation (2.3) as an example. From equation (2.6), we can write v_{2i}^2 as $exp(X_{2i}\theta_2)v_{2i}^{*2}$. Taking logarithm on both sides, we have $ln(v_{2i}^2) = X_{2i}\theta_2 + ln(v_{2i}^{*2}) = X_{2i}\theta_2 + \omega_{2i}^*$, meaning that we can use a simple OLS regression to explore how the variance of v_{2i} is affected by covariates X_{2i}. The Appendix provides detailed steps for estimating the return to spousal education using the Klein and Vella (2010) approach.

2.3 Data

We use data from the Chinese Household Income Project (CHIP), a repeated cross-sectional study focusing on collecting household income and expenditure information. Five waves of surveys have been conducted since 1989. In each wave, the CHIP collects detailed data for each household member on income, employment, education, and demographic characteristics. Starting from wave 2002, parental information (for example, education and occupation) is also collected. In this analysis, we use data from the three most recent waves, wave 2002, 2007, and 2013.

As rural household income is generally indivisible (Millimet and Wang 2006), we focus on urban households. If there are multiple husbands and wives in the same household, we do not have information from the CHIP to link each husband to his wife. We thus restrict our sample to households with one husband and one wife. We further restrict our sample to households for which both husbands and wives are aged between 18 and 60 because individuals younger than 18 are more likely to be still in school, and individuals older than 60 are more likely to be retired. We also drop households with potential measurement errors.[1].

We now discuss the variables used for wage equation (2.1) and education equations (2.2), (2.3). In the wage equation, we use the log of annual wage as our dependent variable.[2] Independent variable of interest is spousal years of schooling. As in a standard Mincer earnings function, we include work experience and work experience squared. Besides these variables, we also include an indicator for whether the individual is Han, an indicator for whether the individual has urban Hukou, birth cohort dummies (born before 1960, born between 1961 and 1970, and born after 1971), province fixed effects, and wave fixed effects.

In education equations (2.2) and (2.3), the dependent variable is the number of years of schooling. Control variables include an indicator for whether the individual is Han, an

[1]In wave 2007, respondents were asked to report monthly income and number of hours worked per week at their primary job. They were also asked to report total monthly income and the number of hours per week worked at all jobs if they had some side jobs. We drop observations whose total monthly income at all jobs is less than that at the primary job. Similarly, we also drop observations whose number of hours worked at all jobs is smaller than that at the primary job. It is also unrealistic for an individual to work more than $24 \times 7 = 168$ hours per week. We thus exclude individuals who worked more than 168 hours per week in wave 2007.

[2]Following the CHIP instruction, annual wage for wave 2002 is constructed as a sum of the following: total income, subsidy for minimum living standard, living hardship subsidies from work unit, second job, and sideline income, and monetary value of in-kind income. In wave 2007, respondents were asked to report monthly income from the primary job and total monthly income from all jobs if they had some other jobs besides the primary job. The annual wage for wave 2007 is calculated as (monthly primary job income + monthly other job income) $\times$ 12. In wave 2013, annual wage is computed as the sum of annual income from the primary job and annual income from side jobs.

indicator for whether the individual has urban Hukou, birth cohort dummies (born before 1960, born between 1961 and 1970, and born after 1971), father's years of schooling, mother's years of schooling, an indicator for whether father has/had a white collar job, an indicator for whether mother has/had a white collar job, and province fixed effects.[3] [4]

We present summary statistics in Table 2.1. Our final sample contains 7081 households. Table 2.1 indicates that there are substantial differences in labor market outcomes between husbands and wives. Husbands tend to have higher wages. The average annual wage for husbands is 28577 RMB, while the average annual wage for wives is 21499 RMB. The wage gap between husbands and wives can be partially explained by husbands' higher hourly wages and slightly longer working hours. The average hourly wage is 13.24 RMB for husbands and 10.50 RMB for wives. Compared with wives, husbands work about 71 hours more per year, approximately 11 minutes more per day. In addition to labor market outcomes, husbands tend to be older and have ampler work experience.

2.4 Empirical Results

2.4.1 Validity Test: Existence of Heteroskedasticity

The validity of the Klein and Vella (2010) approach hinges on the assumption of the presence of heteroskedasticity in the wage-education model described by equations (2.1) - (2.3). As we discussed earlier, the assumption of heteroskedasticity requires at least one of u_i and v_{1i}, and at least one of u_i and v_{2i} to be heteroskedastic. In practice, we employ the Breusch-Pagan/Cook-Weisber test to test the presence of heteroscedasticity in the education

[3]The CHIP provides each respondent's parental education level instead of parental years of schooling. However, CHIP collects both education level and years of schooling for each respondent. To convert parental education level to schooling years, in each wave, we obtain the average schooling years for each education level using data on respondent educational achievement. We use the average years of schooling associated with parental education level as the estimate for parental years of schooling.

[4]An individual is considered having a white collar job if he/she has an occupation that falls into the following categories: 1. Jobs at state agencies, party organizations, enterprises, and public institutions; 2. Professional technicians; 3. Clerk and relevant personnel. These occupations usually require formal education and are more likely to build long-lasting careers.

equations (2.2) and (2.3). The results are reported in Table 2.3. We can reject the null of constant variance for all variables together at 0.1% confidence level for both education equations, confirming the existence of heteroskedasticity in the wage-education model.

To further examine the sources of heteroskedasticity, we run OLS regressions using the same specifications in the education equations (2.2) and (2.3) but with the dependent variable being $ln(\hat{v}_{ci}^2)$, $c = 1, 2$. $\hat{v}_{ci}$ is estimated from the education equations using OLS. Results in Table 2.2 columns (3) and (4) show that sources of heteroskedasticity differ between husbands and wives. For husbands, demographic characteristics such as ethnicity and urban Hukou status, as well as family background such as mother's education and whether mother has/had a white collar job contribute to the heteroskedasticity. However, these factors are not sources of heteroskedasticity for wives. We also find that birth cohort and province of residence contribute to the heteroskedasticity for both husbands and wives. These results provide direct statistical evidence of the existence of heteroskedasticity.

In comparison of results in Table 2.2 columns (1) - (2) and (3) - (4), we find that some of the variables that contribute to the variance of schooling associated with ability (heteroskedasticity) also affect the educational achievement. For example, results in Table 2.2 indicate that average level of educational attainment, as well as the variance of education vary across different age cohorts and provinces of residence.

Birth cohort is an important determinant for both the average education level and the variance of education. On one hand, younger cohorts have a higher average education level because they had a wider access to education. They experienced the re-introduction of key schools and private schools after the Cultural Revolution (1966 - 1976). In addition, some of them may also have benefited from the establishment of nine years of compulsory education in 1985 (Qian and Smyth 2008). Besides the education reform, younger cohorts could also derive benefit from the One Child Policy carried out in 1979 due to the quantity-quantity trade-off (Zhang 2017). On the other hand, younger cohorts also experience larger education dispersion. Prior to the education reform, individuals had relatively uniform access to education. After the education reform during China's economic and social reforms, local governments take on a greater role in educational investment. and households are responsible for their own educational cost. The resulting differences in school qualities across different regions and unequal educational opportunities among families with different characteristics (Chunling, 2006) amplify the variance of education among younger cohorts.

With respect to province of residence, we find that provinces with a lower average educational attainment often have a larger variance in education. The literature has generally found significant education disparities among provinces (Golley and Kong 2018, Qian and Smyth 2008). More developed provinces are likely to be equipped with better schools with higher-quality teachers. If the goal of all schools is to increase their students' education, then we should expect the average educational achievement among better schools to be higher. Since better schools are more effective and successful in terms of directing a majority of their students to a certain educational level (Kim and Choi 2008), we should also expect the variation in students' educational achievement to be smaller. The unequal distribution of educational resources across provinces can potentially explain why province of residence serves as an important source of heteroskedasticity.

2.4.2 Effects of Spousal Education on Wages

OLS Results

Columns (1) and (3) of Table 2.4 present the estimated effects of spousal education on the log of annual wages using OLS. The OLS results show that an additional year of spousal years of schooling is associated with a 2.37% increase in wages for husbands and a 1.71% increase in wages for wives. These results are statistically significant at the 1 percent level. The OLS estimates are potentially imprecise as the OLS estimates could be biased due to the endogeneity. The direction of bias depends on the mating pattern. If there exists positive assortative mating (i.e., well-educated individuals may have higher wages and tend to marry a well-educated spouse), the OLS estimates are biased upward. If there exists negative assortative mating (i.e., well-educated individuals may have higher wages and tend to marry an ill-educated spouse), the OLS estimates are biased downward.

Control Function Results

Turning to our control function results in columns (2) and (4) of Table 2.4, we continue to find positive returns to spousal education. An increase in spousal education leads to a 4.79% increase in wages for husbands and a 2.51% for wives. Putting these results in perspective, note that the average annual wage is 28577.22 RMB for husbands and 21499.44 RMB for wives. The return to spousal education is 1368.85 RMB for husbands and 539.63 RMB for wives. Our control function results confirm that spousal education has a causal effect on wages and rule out theories that suggest that the positive relationship between spousal education and one's wages is merely a correlation.

The control function results are higher in magnitude than the OLS estimates, suggesting that the OLS estimates are biased downward. Positive assortative mating in China is well-documented in the literature (Han 2010, Nie and Xing 2019). Downward-biased OLS estimates suggest that there exists negative assortative mating on unobservable characteristics, meaning that individuals tend to marry someone with unobservable complementary traits such as personality.

Productivity vs. Labor Supply

Productivity and labor supply are two major determinants of wages. We continue to use our control function model to explore whether spousal education affects wages by increasing productivity or increasing labor supply or increasing both. Productivity is measured as the log of hourly wage. Labor supply is measured as annual working hours. Table 2.5 displays the results. We first notice that spousal education has a positive effect on productivity but a negative effect on labor supply. An additional year of spousal schooling is associated with a 5.06% increase in productivity for husbands, and a 3.09% increase in productivity for wives.

The magnitudes of effects are very similar to those of effects on annual wages. The negative effect of spousal education on wages is small in magnitude. An increase in spousal education leads to a 22-hour reduction in annual labor supply (5-minute reduction in daily labor supply) for husbands and a 16-hour reduction (4-minute reduction in daily labor supply) for wives. Table 2.5 indicates that spousal education affects wages by increasing productivity rather than increasing labor supply.

Heterogeneity in Effects of Spousal Education on Wages

Heterogeneity by Spousal Education

Based on Welch (1970), a better-educated spouse may have a higher allocative effect, meaning that they can more efficiently gather, process, and interpret information, therefore providing better advice on decision making. This will help individuals reach better decisions and lead to higher wages later (Loh 1996). In addition, a better-educated spouse may provide a higher-quality network that may help an individual transition to a better-paying job. Therefore, it is possible that there is heterogeneity in return to spousal schooling across different spousal education levels as spouses of different educational levels are equipped with different skill sets related to the allocative effect and different networks.

To explore whether spousal education level matters to the return to spousal schooling, we re-run our analysis by two groups. One group includes individuals whose spousal schoolings are at least 12 years.[5] The other group contains the rest of the sample. Results in Table 2.6 show that the return to spousal schooling is substantially higher among individuals with a better-educated spouse. One additional year of schooling received by a better-educated spouse increases wages by 11.4% for husbands, and 6.96% for wives. An increase in schooling of a lower-educated spouse increases wages only for husbands. The effect is 3.56%.

[5] 12 years is the median years of schooling in our sample. 12-year-schooling indicates having a middle school diploma.

Heterogeneity by Year

In China, wives typically specialize in housework, while husbands typically specializ in market work. The dramatic demographic changes during the past few decades have influenced the gender role in Chinese households. On one hand, reduced fertility greatly freed women from heavy housework. On the other hand, women's education has tremendously increased due to the education expansion (Wu and Zhang 2010). Along with the increase in economic returns to education (Zhang et al. 2005), women are more able and willing to exert themselves to build a career in the labor market. At the same time, the increase in women's education and wages leads to a growing incentive for husbands to support their wives' careers "given the shared benefits of labor market success within marriage" (Jolly 2019). If spousal education affects wages through gender-based specialization, then we should expect the return to spousal schooling for wives to increase over time as women have been more involved in the labor market.

In Table 2.7, we explore whether the effect of spousal education on wages varies over time. We re-do our analysis by survey year. Our results suggest that the return to spousal education for husbands is relatively consistent throughout the years. The return to spousal education is 6.23% for year 2002, 3.69% for year 2007, and 4.81% for year 2013. All the estimates are statistically significant at the 1 percent level. Return to spousal schooling for wives appears a different pattern. Spousal education does not have an effect on wages for wives in year 2002 and year 2007, while there exists a positive effect of spousal education on wages in year 2013. The return to spousal schooling for wives in year 2013 is 3.46%. Table 2.7 provides two interesting findings. First, there is a long-standing positive causal relationship between spousal education and wages for husbands. Second, the cross-productive effect from husbands to wives emerged recently. The findings are consistent with our prediction that spousal education affects wages through gender-based specialization, suggesting specialization is a potential mechanism underlying the positive relationship between spousal education and wages.

Heterogeneity by Dominant Earner Status

Besides gender, income is an important determinant for household specialization. We define a husband as a dominant earner if the husband earns more than his wife. Likewise, we define a wife as a dominant earner if the wife earns more than her husband. Spouses who make more earnings in their households are more likely to specialize in the labor market. We should expect the return to spousal education to be larger among dominant earners.

Table 2.8 looks at the heterogeneity in return to spousal education by dominant earner status. We highlight two findings. First, we do not detect substantial heterogeneity in return to spousal schooling by dominant earner status for wives. Second, return to spousal education is higher for non-dominant earner husbands. The effect is 10.9% for non-dominant earner husbands, while 7.65% for dominant earner husbands. The results are not consistent with our earlier prediction. One possible explanation is that gender plays a larger role in the specialization. Wives, regardless of their dominant earner status, are more likely to specialize in housework. While husbands, regardless of their dominant earner status, are more likely to specialize in the labor market. Non-dominant earner husbands are more likely to have a better-educated wife than dominant earner husbands, meaning that they are more likely to benefit from wives' knowledge and network to increase their wages.

Heterogeneity by Same Occupation Status

How spouses sharing the same occupation influences the return to spousal education is not clear. On one hand, couples who share the same occupation may be more likely to share work-related knowledge and learn from each other. In this case, we should expect the return to spousal education is higher among spouses who share the same occupation. On the other hand, couples who share the same occupation may have overlapped networks, meaning the benefit of expanding each other's network is more limited. If the network is the main channel through which spousal education affects wages, then we should expect the return to spousal education is more prominent among spouses who are in different occupations.

In Table 2.9, we conduct an analysis by whether spouses share the same occupation. We find that different patterns appear for husbands and wives. The return to spousal education when spouses share the same occupation is 6.34% for husbands and 6% for wives. The return to spousal education when spouses are in different occupations is 8.28% for husbands and 4.92% for wives. Our finding suggests that husbands benefit more from having a spouse in a different occupation, while wives benefit less from having a spouse in a different occupation. These findings suggest that spousal education affects wages through different channels for husbands and wives. Since husbands are more likely to specialize in the labor market, they can derive more benefits from a wider network and gain more wages. Wives tend to specialize in housework, and a wider network is less important than knowledge sharing and learning from the spouse.

2.4.3 Summary

Specialization

The positive effect of spousal education on wages is consistent with the specialization theory (Becker 1991). Spouses could specialize in different tasks and then exchange with minimal cost within marriage. Since more-educated spouses may have higher productivity in their specialized areas, having a better-educated spouse may increase gains from marriage. In China, specialization is typically gender-based. If specialization is the channel through which spousal education affects wages, the gains from specialization measured by wages should be more prominent for husbands. This is indeed what we find from our results. Return to spousal education for husbands is 2.28% higher than that for wives (Table 2.4). This pattern is persistent in all other analyses we conduct earlier. Our analyses on heterogeneity in return to spousal education provide additional evidence that is in favor of specialization being the potential channel.

Other channels

Specialization in home production might be less viable for wives who are better-educated and more productive in market work. If specialization is the only channel through which spousal education affects wages, we should expect the return to spousal education for husbands is smaller for husbands with a better-educated wife. However, our results show the opposite. This finding suggests that specialization is not the sole mechanism underlying the positive relationship between spousal education and wages. Besides specialization, spouses might experience gains from marriage through providing advice, expanding network, sharing knowledge, and learning from each other. These channels provide theoretical support for the finding that return to spousal education increases with spousal education level. A better-educated spouse is more likely to provide better advice, higher-quality network, sharing more productive knowledge, and hence is more likely to create more gains for husbands in the labor market. Our analyses on heterogeneity in return to spousal education by dominant earner status and by same occupation status lend support to this.

2.4.4 More discussion

The sample we have used so far is restricted to households for which both husband and wife are participating in the labor market. One concern is that our sample may not be a randomly selected sample from the underlying population, and our estimation may encounter the selection-bias problem. If abler individuals are more likely to participate in the labor market and specialize in market work, the return to spousal education should be more prominent among these individuals. If there exists a positive selection into the labor force, our estimates are biased upward. In this case, our results can be considered as the upper bound of the return to spousal schooling.

2.5 Conclusion

This paper investigates the effect of spousal education on wages using the method proposed by Klein and Vella (2010). Utilizing the nonlinearity of control terms induced by heteroskedasticity, we are able to estimate the causal effect of spousal education in the Chinese context. We find strong evidence of a cross-productivity effect from husband to wife as well as from wife to husband. An additional year of schooling received by the wife increases her husband's wage by 4.79%. The cross-productivity effect for wives is smaller. An additional year of schooling received by the husband increases his wife's wage by 2.51%. We also find that spousal education affects wages through increasing productivity rather than labor supply. Finally, Our results suggest there is substantial heterogeneity in return to spousal education by gender, spousal education level, year, dominant earner status, and whether spouses are in the same occupation.

The finding that spousal education has a cross-productivity effect has three important implications. First, it shows one's education not only contributes to one's own earnings but also contributes to the spouse's earnings. Only considering the effect of education on one's own earnings may actually underestimate the effect of education. Second, the presence of the cross-productivity effect indicates that education has positive externalities within marriage, which provide additional justifications for government intervention to improve education. It also has important implications for policy making regarding alleviating income inequality. Third, our results also suggest that spousal education is not a valid instrumental variable for one's own education in estimating return to education since spousal education has a direct impact on one's earnings.

Table 2.1: Summary statistics

	Husband		Wife	
	Mean	Std. Dev	Mean	Std. Dev
Panel A: Individual characteristics				
Age	42.56	7.21	40.57	7.12
Years of schooling	11.90	3.22	11.46	3.11
Work experience	23.59	8.60	21.62	8.64
Has an urban Hukou	96.89%	17.35%	95.69%	20.30%
Han	96.71%	17.84%	96.62%	18.06%
Birth cohort				
-1960	34.47%	47.53%	26.23%	43.99%
1961-1970	42.52%	49.44%	43.91%	49.63%
1971-	23.01%	42.09%	29.87%	45.77%
Panel B: Family background				
Father's years of schooling	6.50	4.18	6.84	4.11
Mother's years of schooling	4.66	4.06	5.13	4.03
Father has/had a white collar job	36.07%	48.02%	38.29%	48.61%
Mother has/had a white collar job	17.17%	37.72%	18.42%	38.76%
Panel C: Labor market outcomes				
Annual wage (RMB)	28577.22	26051.26	21499.44	30624.95
Hourly wage (RMB)	13.24	13.87	10.50	17.31
Hours worked in a year	2320.50	654.19	2249.90	666.66
N	7081		7081	

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013.

Table 2.2: Education equation

	Years of schooling		$ln(\hat{v}^2)$	
	(1)	(2)	(3)	(4)
	Husband	Wife	Husband	Wife
Han	0.287	0.114	-0.283*	-0.236
	(0.205)	(0.190)	(0.145)	(0.151)
Urban Hukou	2.115***	2.067***	0.340**	-0.0294
	(0.206)	(0.165)	(0.146)	(0.131)
Father's years of schooling	0.123***	0.100***	-0.00609	-0.00956
	(0.0118)	(0.0110)	(0.00833)	(0.00872)
Mother's years of schooling	0.0502***	0.119***	-0.0170*	-0.00476
	(0.0124)	(0.0115)	(0.00879)	(0.00912)
Mother has/had a white collar job	0.481***	0.356***	-0.0915	-0.107
	(0.116)	(0.105)	(0.0821)	(0.0835)
Father has/had a white collar job	0.270***	0.493***	-0.176***	-0.107
	(0.0912)	(0.0833)	(0.0644)	(0.0661)
Birth cohort (base group: born before 1960)				
1960-1970	0.898***	0.433***	-0.0472	0.171***
	(0.0827)	(0.0831)	(0.0584)	(0.0659)
1970-	1.906***	1.683***	0.141*	0.223***
	(0.102)	(0.0950)	(0.0720)	(0.0754)
Province fixed effect (base group: Beijing)				
Shanxi	-0.880***	-0.935***	0.251*	0.420***
	(0.198)	(0.185)	(0.140)	(0.147)
Liaoning	-1.177***	-1.204***	-0.0146	0.279*
	(0.196)	(0.183)	(0.138)	(0.145)
Shanghai	-0.938***	-1.414***	-0.354*	0.120

Continued on next page

Table 2.2 – continued from previous page

	Years of schooling		$ln(\hat{v}^2)$	
	(1)	(2)	(3)	(4)
	Husband	Wife	Husband	Wife
	(0.259)	(0.242)	(0.183)	(0.192)
Jiangsu	-0.906***	-1.378***	0.301**	0.369***
	(0.180)	(0.168)	(0.127)	(0.133)
Zhejiang	-1.286***	-1.552***	0.195	0.195
	(0.242)	(0.226)	(0.171)	(0.179)
Anhui	-0.985***	-1.562***	0.143	0.541***
	(0.186)	(0.174)	(0.131)	(0.138)
Shandong	-1.238***	-1.248***	0.207	0.182
	(0.242)	(0.226)	(0.171)	(0.179)
Henan	-1.048***	-1.431***	0.113	0.347***
	(0.179)	(0.168)	(0.127)	(0.133)
Hubei	-0.593***	-1.105***	0.0785	0.158
	(0.179)	(0.168)	(0.127)	(0.133)
Hunan	-0.741**	-1.477***	0.159	0.345
	(0.293)	(0.274)	(0.207)	(0.217)
Guangdong	-0.816***	-1.402***	0.156	0.163
	(0.174)	(0.163)	(0.123)	(0.129)
Chongqing	-0.932***	-1.287***	0.274*	0.423***
	(0.205)	(0.192)	(0.145)	(0.153)
Sichuan	-1.395***	-1.816***	0.226*	0.416***
	(0.184)	(0.172)	(0.130)	(0.137)
Yunan	-0.577***	-0.874***	0.442***	0.716***
	(0.202)	(0.189)	(0.143)	(0.150)
Gansu	-0.712***	-1.174***	0.0125	0.131

Continued on next page

Table 2.2 – continued from previous page

	Years of schooling		$ln(\hat{v}^2)$	
	(1)	(2)	(3)	(4)
	Husband	Wife	Husband	Wife
	(0.218)	(0.204)	(0.154)	(0.162)
Constant	8.418***	8.368***	0.940***	0.658***
	(0.323)	(0.280)	(0.228)	(0.222)
N	7081	7081	7081	7081

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use OLS regression. Standard errors in parentheses. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.3: Validity test

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity Ho: Constant variance		
	Years of schooling	
	(1)	(2)
	Husband	Wife
Statistic	12.68	13.44
P-value	0.0004	0.0002

Table 2.4: Wage equation

	Husband		Wife	
	(1) OLS	(2) CF	(3) OLS	(4) CF
Own years of schooling	0.0617***	0.0797***	0.0832***	0.119***
	(0.00346)	(0.00923)	(0.00447)	(0.0103)
Spousal years of schooling	0.0237***	0.0479***	0.0171***	0.0251**
	(0.00341)	(0.00729)	(0.00411)	(0.0104)
N	7081	7081	7081	7081

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. Dependent is log(annual wage) for all columns. Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The CF standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.5: Productivity vs. Labor supply

	Productivity		Labor supply	
	(1) Husband	(2) Wife	(3) Husband	(4) Wife
Own years of schooling	0.0841***	0.114***	-42.95***	-36.66***
	(0.00797)	(0.00722)	(8.460)	(6.512)
Spousal years of schooling	0.0506***	0.0309***	-22.12***	-16.48**
	(0.00617)	(0.00723)	(7.044)	(6.583)
N	7042	7039	7078	7081

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use control function approach. Dependent variable is log(hourly wage) in columns (1) and (2). Dependent variable is annual working hours in columns (3) and (4). Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.6: Spousal education effect by spousal education level

	Husband		Wife	
	(1) Spousal schooling >= 12	(2) Spousal schooling <12	(3) Spousal schooling >= 12	(4) Spousal schooling <12
Own years of schooling	0.0905***	0.0835***	0.135***	0.115***
	(0.0166)	(0.0149)	(0.0164)	(0.0174)
Spousal years of schooling	0.114***	0.0356**	0.0696**	0.0312
	(0.0225)	(0.0181)	(0.0289)	(0.0273)
N	3693	3388	4046	3035

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use control function approach. Dependent is log(annual wage) for all columns. Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.7: Spousal education effect by year

	Husband			Wife		
	(1) 2002	(2) 2007	(3) 2013	(4) 2002	(5) 2007	(6) 2013
Own years of schooling	0.0967***	0.0851***	0.0775***	0.103***	0.121***	0.118***
	(0.0148)	(0.0206)	(0.0145)	(0.0166)	(0.0227)	(0.0158)
Spousal years of schooling	0.0623***	0.0369**	0.0481***	0.0191	0.0218	0.0346**
	(0.0122)	(0.0161)	(0.0113)	(0.0171)	(0.0208)	(0.0151)
N	3231	1587	2263	3231	1587	2263

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use control function approach. Dependent is log(annual wage) for all columns. Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.8: Spousal education effect by dominant earner status

	Husband		Wife	
	(1) Dominant earner	(2) Non-dominant earner	(3) Dominant earner	(4) Non-dominant earner
Own years of schooling	0.0925***	0.0920***	0.0994***	0.112***
	(0.00950)	(0.0219)	(0.0170)	(0.0124)
Spousal years of schooling	0.0765***	0.109***	0.0653***	0.0625***
	(0.00796)	(0.0172)	(0.0147)	(0.0127)
N	4867	2214	1682	5399

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use control function approach. Dependent is log(annual wage) for all columns. Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

Table 2.9: Spousal education effect by occupation

	Husband		Wife	
	(1) Spouse in same occupation	(2) Spouse in different occupation	(3) Spouse in same occupation	(4) Spouse in different occupation
Own years of schooling	0.0902***	0.0958***	0.120***	0.112***
	(0.0138)	(0.0140)	(0.0166)	(0.0139)
Spousal years of schooling	0.0634***	0.0828***	0.0600***	0.0492***
	(0.0114)	(0.0114)	(0.0168)	(0.0138)
N	2781	4300	2781	4300

Data source: urban households from the Chinese Household Income Project (CHIP) wave 2002, 2007, and 2013. All columns use control function approach. Dependent is log(annual wage) for all columns. Control variables include Han, Hukou status, work experience, work experience squared, cohort dummies, province dummies, and wave dummies. Standard errors in parentheses. The standard errors and p-statistics are calculated from bootstrapping with 299 replications. * p <= 0.1, ** p <=0.05, *** p <= 0.01.

2.6 Appendix

Estimation of the empirical model using the Klein & Vella (2010) Approach

$$W_i = \alpha + \beta_1 S_i^{self} + \beta_2 S_i^{spouse} + X_{0i}\delta + \rho_1 \frac{H_{ui}(X_{0i})}{\hat{H}_{v1i}(X_{1i})}\hat{v}_{1i} + \rho_2 \frac{H_{ui}(X_{0i})}{\hat{H}_{v2i}(X_{2i})}\hat{v}_{2i} + e_i$$

proceeds as follows:

1. Regress S_i^{self} on X_{1i} and obtain $\hat{v}_{1i}$.
2. Regress S_i^{spouse} on X_{2i} and obtain $\hat{v}_{2i}$.
3. Regress $ln(\hat{v}_{ci})$ on X_{ci} to obtain θ_{cv} and compute $\hat{H}_{cvi}(X_{ci}) = \sqrt{exp(X_{ci}\hat{\theta}_{cv})}$, $c = 1, 2$.
4. Obtain consistent estimates via Nonlinear least-squares estimation:

$$\min_{\alpha,\beta_1,\beta_2,\delta,\rho_1,\rho_2,\theta_u} \sum_i \begin{bmatrix} W_i - \alpha - \beta_1 S_i^{self} - \beta_2 S_i^{spouse} - X_{0i}\delta \\ - \rho_1 \sqrt{exp(X_{0i}\theta_u)} \dfrac{\hat{v}_{1i}}{\hat{H}_{v1i}(X_{1i})} - \rho_2 \sqrt{exp(X_{0i}\theta_u)} \dfrac{\hat{v}_{2i}}{\hat{H}_{v2i}(X_{2i})} \end{bmatrix}^2$$

5. Estimate θ_u again by regressing $ln(\hat{u}_i^2)$ on X_{0i} and compute $\hat{H}_{ui}(X_{0i})$, where

$$\hat{u}_i = W_i - \alpha - \hat{\beta}_1 S_i^{self} - \hat{\beta}_2 S_i^{spouse} - X_{0i}\hat{\delta}.$$

6. Estimate parameters via OLS

$$W_i = \alpha + \beta_1 S_i^{self} + \beta_2 S_i^{spouse} + X_{0i}\delta + \rho_1 \frac{\hat{H}_{ui}(X_{0i})}{\hat{H}_{v1i}(X_{1i})}\hat{v}_{1i} + \rho_2 \frac{\hat{H}_{ui}(X_{0i})}{\hat{H}_{v2i}(X_{2i})}\hat{v}_{2i} + e_i$$

7. Compute standard errors through bootstrap.

Chapter 3

Heterogeneity in Child Quantity-Quality Trade-off: A Machine Learning Approach

3.1 Introduction

Economists have long been interested in understanding the relationship between family size and child quality. This is not only because family environment is a primary component to child's quality(Black et al. 2005), but also because understanding this relationship is important to policy makers. The theoretical quantity-quality model (Becker and Lewis 1973, Becker and Tomes 1976) predicts a negative effect of family size on child quality. This quantity-quality trade-off has become the main justification for family planning campaigns to increase population quality by curbing population growth. We use recently developed machine learning methods to investigate the heterogeneous effects of family size. We find the mother's age at first birth, her income and education levels play a large role in the negative effect of additional children on health outcomes.

Many empirical studies have tested the quantity-quality trade-off using data from various countries. They find mixed results. Some studies find that child quality is not significantly affected by family size (Kessler 1991; Guo and VanWey 1999; Black et al. 2005; Angrist et al. 2010; Zhong 2014). Other studies find evidence of a quantity-quality trade-off (Rosenzweig and Wolpin 1980; Cáceres-Delpiano 2006; Li et al. 2008; Liu 2014; Rosenzweig and Zhang 2009). A third set of studies suggests a positive effect of family size on child quality (Lee 2008, Qian 2009).

Testing for the existence of a quantity-quality trade-off is complicated by the endogeneity of family size. If parents who place a greater value on child quality more prefer fewer children, then the relationship between family size and child quality will be driven by parental preferences rather than family size. A commonly used method to establish the causal effect of family size on quality is to use instrumental variables (IV). Empirical studies have used the birth of twins (Rosenzweig and Wolpin 1980; Black et al. 2005; Cáceres-Delpiano 2006; Li et al. 2008; Angrist et al. 2010), twinning by birth order (Rosenzweig and Zhang 2009), child sex composition (Angrist et al. 1998; Angrist et al. 2010; Lee 2008), and the One Child Policy (OCP) (Qian 2009; Liu 2014) to extract exogenous variation in family size. One of the distinguishing features of the IV approach is that, without a homogeneity assumption, the IV estimate is the local average treatment effect (LATE) of a group of compliers (Angrist and Imbens 1995). It is possible that the compliers of different instrument variables are so disparate that the average treatment effects of compliers using different instrument variables differ from each other significantly. The mixed results from using the IV approach suggests that there might be heterogeneity in the effect of family size on child quality.

There are several theories purporting to explain the potential heterogeneity in the impact of family size on child quality. Rosenzweig and Wolpin (1980) point out that whether there is a quantity-quality trade-off depends on whether child quantity and child quality are substitutes or complements. If parents are facing a budget constraint on how much they can invest in total child quality, then having an additional sibling reduces the average amount of resources invested in each child, hence reducing individual child quality. However, it is possible that, for some families, additional siblings benefit from existing children by stabilizing the parental relationship (Becker 1998, Black et al. 2005), increasing the likelihood of the mother staying at home to provide child care (Gelbach 2002; Black et al. 2005; Ruhm 2008), or generating other positive spillovers (Bandura and Walters 1977). It is also possible that some parents adjust to an exogenous increase in family size by working longer hours, consuming less leisure, or investing less in themselves rather than decreasing quality inputs on each child (Angrist et al. 2010).

While the relationship between family size and child quality has been under intensive investigation in the literature, few studies have looked beyond LATE. This paper uses recently developed machine learning methods to examine the potentially heterogeneous effect of family size on child quantity in the Chinese context. We find substantial heterogeneity in treatment effects of additional children on child quality. The heterogeneity in treatment effects is most pronounced for girls. Giving birth later in a mother's reproductive life cycle, being more educated and having higher incomes all contribute to treatment effect heterogeneity.

This paper makes several contributions to the literature. First, we move beyond LATE and allow for the heterogeneous treatment effects of family size on child quality. Moreover, as measures of quality, we focus not only on children's education but also on children's health. Examining the heterogeneity of treatment effects of family size on child quality is crucial to understanding the process of human capital investment within households. It also has strong policy implications as family planing policies might have different impacts on households with disparate characteristics.[1]

Second, the machine learning method we propose advances the traditional approach in several ways. Traditional approaches to explore heterogeneous effects involve analyzing subgroups and including interaction terms in the model. These methods require the researcher

[1] For related studies please consult Hedrich (2011) and Brinch et al. (2017).

to have a thorough understanding of the research question to define subgroups. Additionally, the potential for cherry-picking problems may arise and unexpected subgroups may be missed (Lee et al. 2020). This approach involves interacting the treatment with various covariates, usually one covariate at time. This raises the probability of spurious findings (Davis and Heller 2017). Commonly used linear approaches may also fail to capture some nonlinear treatment effects since there might be nonlinearities in child quality from changes in family size (Løken et al. 2012). In contrast, machine learning methods use data-driven algorithms to detect heterogeneity which avoids leaving out important heterogeneities. Furthermore, flexible modeling alleviates the concern of the existence of nonlinear relationship between family size and child outcomes. Therefore this method makes the discovery of treatment heterogeneity more reliable for policy determination. We find that being a girl, mothers age at first birth, parental education and income levels play a larger role in the quantity-quality trade-off. Hence, policy-makers can fine-tune what segment of the population they need to concentrate on to alleviate the detrimental effects of having more children on child quality.

Third, this paper joins the literature on the effects of the OCP on child outcomes. It fills an important gap that exists in the literature by investigating the heterogeneity the OCP's treatment effects. As such we use data from the 1976 Statistical Yearbook of China where we collect information on provincial characteristics prior to the implementation of the One-Child-Policy (OCP). We use this to control for fertility preference prior to the OCP. We also use data from the China Health and Nutrition Survey (CHNS), a longitudinal dataset that collects detailed information on households on health, education, and a variety of other demographic information.

The paper is structured as follows: the next section will discuss the data used in our estimation. Following this, we lay out the empirical methodology in section 3.3. In section 3.4 we explain the results, including the IV and the estimation results of machine learning algorithm. Section 3.5 concludes.

3.2 Data

The first source of data we use is from the 1976 Statistical Yearbook of China where we collect information on provincial characteristics prior to the implementation of the One-Child-Policy.

The second data source we use in this paper is from the China Health and Nutrition Survey (CHNS), a longitudinal dataset that collects detailed information of households on health, education, and other basic demographic information. From this dataset we calculate height-for-age z-score (HAZ) and weight-for-age z-score (WAZ) as the two measures for child's health. HAZ and WAZ are constructed using the British 1990 Growth Reference. For child's education, we use relative educational attainment, and a dummy variable for being currently enrolled in school.[2]

In our analysis of the quantity-quality trade-off, we restrict our sample to first-born children. We impose a few restrictions to our sample. First, we exclude children from households with twins because the birth of twins results in shorter birth spacing. This could potentially interfere with our main results. Second, we restrict our sample to children aged between 6 and 17 at the time of survey. Before age 6, children are not in school and after age 17, the family influence is diminishing. Third, we restrict our sample to children who were born in 1976 or after. Before the adoption of the OCP, the family planning policy was known as the "later(marriage), longer(intervals), fewer(children)" (Chen and Huang 2018). The recommended birth spacing is 4 years (Liu 2014). Since the OCP was first carried out in some provinces in 1979, for couples who followed the previous family planning policy and had their firstborn in 1976 or after, their decision to have a second child would be heavily affected by the implementation of the strict OCP. Finally, we exclude firstborns whose province of residence is not available. In addition, we drop the observations from Chongqing because we were unable to obtain the pre-OCP provincial characteristics of Chongqing as Chongqing

[2]Following Rosenzweig and Wolpin (1980), relative educational attainment is constructed as $Educ_{igw}/\overline{Educ}_{gw}$, where $Educ_{igw}$ is the years of schooling of child i at age g from wave w and $\overline{Educ}_{gw}$ is the average years of schooling of children at age g in wave w.

was a part of Sichuan province at that time.

We present summary statistics in Table 3.2. The firstborns in our sample are relatively evenly distributed between boys and girls. 90% of the children are currently enrolled in school. The average age of the sample is 11 years and the average years of schooling is 5.48, implying that, on average, children in our sample are in their last year of primary school. We break down our sample into boys and girls. We highlight three findings from comparing boys' and girl's summary statistics. First, boys and girls are very similar in terms of family background, such as parental health and parental educational attainment. Second, we do not observe large disparities between boys' and girls' education measures. Third, HAZ and WAZ diverge between boys and girls. Girls' HAZ and WAZ are much lower.

3.3 Empirical Methodology

3.3.1 Heterogeneity in exposure to the One Child Policy

The One Child Policy was introduced to curb rapid population growth in China. Although it was a national policy, its implementation varied across provinces. First, the year of implementation of the OCP differed. The official document about the implementation of the OCP was released in 1979 but the actual implementation year among provinces ranged from 1979 to 1984. Second, provinces made modifications to the OCP over the years. The timing and the content of the modifications are often different. For example, the OCP faced strong resistance among rural households, particularly among those whose firstborn was a girl (Zhang 2017). Later, many provinces relaxed their OCP to allow rural households to have a second child. Among these provinces, the year of the rule relaxation ranged from 1985 to 1998. In some provinces, all rural households were eligible for this relaxation, while in other provinces, only rural households with a female firstborn were permitted to have a second child. Several provinces, such as Jiangsu, Zhejiang, Jilin, never carried out this relaxation. Table 3.1 presents the details of heterogeneity in the OCP implementation.

In addition to the divergence in the OCP implementation, the birth year of the mother also contributes to heterogeneity in the household's exposure to the OCP. The fertility of households with a mother whose prime fertility years (21 – 35) were not covered by the OCP is less likely to be influenced. If a mother entered her prime fertility years after the implementation of the OCP, then the longer she was exposed to the OCP, the greater impact the OCP would have on her fertility decision.

3.3.2 Instrumental Variable

Substantial heterogeneity in a household's exposure to the OCP generates a unique source of exogenous variation in family size. Borrowing the idea from Huang (2021), we use the share of a mother's prime fertility years covered by the strict OCP as the instrumental variable:

$$\text{coverage} = \frac{\text{number of prime fertility years covered by the strict OCP}}{\text{total number of prime fertility years}}$$

The strict OCP is defined as only one child being allowed per couple. Suppose province A started implementing the OCP in 1979. Starting from 1985, eligible households were allowed to have a second child. Now we consider a household in province A with a mother who was born in 1954. The mother entered her prime fertility years in 1975 and exited in 1989. If the household was not eligible for having a second child, then 11 (from 1979 to 1989) of the mother's most fertile years were exposed to the strict OCP. *coverage* for this household would be 11/15. If the household qualified for the OCP relaxation, then only 7 (from 1979 to 1985) years of the mother's most fertile years were subject to the strict OCP. In this case, *coverage* would be 7/15. The larger *coverage* is, the more influence the OCP has on the household's fertility choice. We should expect there to be a negative relationship between *coverage* and number of siblings.

3.3.3 Validity of IV

Our IV exploits the provincial variation in the OCP policies to account for exogenous variation in family size. One main concern is that variation in the OCP regulations is associated with provincial characteristics that may also influence fertility. For example, provinces that carried out the OCP relaxation policies sooner might have a stronger preference for a larger family. To alleviate this concern, we include pre-OCP provincial characteristic variables in the model to control for the potential correlation between the OCP polices and preexisting provincial fertility preferences.

Another concern is that couples might manipulate their eligibility for the OCP relaxation to have more children. The relaxed OCP mainly allowed two types of couples to have more than one child: ethnic minority couples and rural couples whose first child was a girl. An individual who desired more children could secure the eligibility by marrying a member of an ethnic minority. Past studies (Huang and Zhou 2015) show that the OCP induced more inter-ethnic marriages; however, it is less concerning in our analysis because nearly all of our sampled provinces required both spouses to be ethnic minority to be qualified for the OCP relaxation. The only exception is Guangxi province, which allowed couples with one ethnic minority spouse to have a second child. However, its policy only lasted for a short period of time before it was tightened to require both spouses to be an ethnic minority.

Rural couples could manipulate their eligibility status by utilizing sex-selective abortion to choose the gender of their first child. However, they had little incentive to do so. Rural couples who were pregnant with a boy were unlikely to terminate the pregnancy due to the strong son preference. Rural couples who were pregnant with a girl had little incentive to terminate their pregnancy because they were allowed to have a second child.

The validity of our IV strategy hinges on the assumption that family size is the only channel through which OCP policies affect child quality. Empirical studies (Ebenstein 2010, Chen et al. 2013) show that the OCP leads to sex ratio distortion among high-order births. This is concerning to our analysis because this means that the OCP might also affect sibling sex composition. Sibling sex composition has its own effect on child quality as parents might allocate resources differently between boys and girls (Behrman et al. 1986).

To examine the empirical relationship between sibling sex composition and our IV, we estimate the following regression:

$$SibSex_{ipt} = \beta_2 coverage_i + D_{it}\gamma_3 + C_p\delta_3 + w_t + \eta_{ipt} \tag{3.1}$$

where $SibSex_{ipt}$ is a measure of sibling sex composition. In this analysis, we use two measures of sibling sex composition: an indicator for a male second birth and the fraction of male siblings. D_{it} is a set of individual characteristics, including child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. C_p is a set of pre-OCP provincial characteristic variables.

Consistent with Ebenstein (2010), we only find a significant relationship between sex composition and our IV for girls. Results in Table 3.3 show that, among households with female firstborns, mothers who are always ineligible for having a second child are 12.3% less likely to have a male second birth compared with those who are always eligible. Firstborns from an always ineligible household also have a lower fraction of male siblings. One possible explanation for the significant negative signs in column (3) and (4) is that ineligible couples have more incentive to use sex-selective abortion on their first birth, while eligible couples have more incentive to use it on their second birth. In the female first-born subsample, it is possible that the ineligible couples have a weaker son preference and less incentive to use sex selective abortion on their second birth than those eligible couples.

Results in Table 3.3 suggest that our IV estimates of the effect of family size on child quality for firstborn boys are not confounded by sibling sex composition. However, the IV estimates for firstborn girls should be interpreted with caution. It is plausible to assume that firstborn girls with a male sibling are allocated with less resources than those with a female sibling. As ineligibility of having a second child leads to lower chances of having a male sibling, our IV estimates for firstborn girls are biased downward. We should consider those as the lower bound of the true estimates of the effect of family size for firstborn girls.

3.3.4 Local average treatment effect

We use a 2SLS approach to access the local average treatment effect of family size on quality of children. In the first stage, number of siblings is instrumented by *coverage* using equation (3.2).

$$nsib_{ipt} = \beta_1 coverage_i + D_{it}\gamma_1 + C_p\phi_1 + w_t + v_{ipt} \tag{3.2}$$

where $nsib_{ipt}$ is the number of siblings child i from province p in wave t. D_{it} contains a vector of individual control variables; C_p is set of provincial control variables [3]; w_t is wave fixed effects.

In the second stage, we use equation (3.3).

$$Y_{ipt} = \theta \widehat{nsib_{ipt}} + D_{it}\gamma_2 + C_p\phi_2 + w_t + \epsilon_{ipt} \tag{3.3}$$

where Y_{ipt} is the quality measure of child i in province p in wave t; $\widehat{nsib_{ipt}}$ is estimated from the first stage; the other variables are the same as defined in equation (3.2).

We consider child quality in terms of both health and education. For health, we use HAZ and WAZ. For education, we use relative educational attainment and a dummy variable for being currently enrolled in school.

The individual control variables include parental years of schooling, child's age, mother's age at first birth, and mother's age at first birth squared. In analysis focused on child's health, we also include parental height and parental weight. We replace the missing values of parental height/weight with the average sample parental height/weight. In addition,

[3] Note that we do not include province fixed effect because our IV is constructed based on provincial variations in the OCP implementation. Including province fixed effect would absorb some of the useful variation in the IV.

we include dummy variables for missing parental height/weight in case parents with some missing characteristics in the data set are characteristically different from those parents with no missing variables.

Provincial control variables include sex ratio, birth rate, log of GDP per capita, share of non-agricultural population, share of primary industry in GDP, and share of secondary industry in GDP to control preexisting provincial features such as fertility preferences.

3.3.5 Conditional local average treatment effect

The goal of this paper is to study heterogeneity in quantity-quality trade-offs. To do so, we estimate the conditional local average treatment of number of sibling on children's quality using the instrumental forest within the generalized random forest framework developed by Athey et al. (2019). The method we use is an adaptive nearest neighbor approach, of which the weighting is derived from a random forest technique.

We are interested in estimating the following model for individual i, $i = 1, \dots, n$:

$$Y_i = \tau(X_i)W_i + \mu(X_i) + \epsilon_i \tag{3.4}$$

where Y_i is a child quality measure of individual i; W_i is the treatment, in our analysis, it is the number of siblings individual i has; $\tau(X_i)$ captures the causal effect of the number of siblings on the quality of i; $\mu(X_i)$ is a nuisance parameter; and ϵ_i is a noise term that is correlated with W_i. Because of the correlation between ϵ_i and W_i, we need to use an instrumental variable Z_i to generate a consistent estimate for $\tau(X_i)$. In this paper, we use $coverage_i$, the share of a mother's prime fertility years covered by the strict OCP, as our instrumental variable. X_i contains the same covariates as in the 2SLS approach along with wave dummies.

$\tau(X_i)$ is obtained by minimizing equation (3.5), an empirical version of two moment functions based on the exclusion assumption of the instrumental variable: $\mathbb{E}[Z_i(Y_i - \tau(x)W_i -$

$\mu(x))|X_i = x] = 0$ and $\mathbb{E}[Y_i - \tau(x)W_i - \mu(x)|X_i = x] = 0$.

$$(\hat{\tau}(x), \hat{\mu}(x)) \in \underset{\tau(x),\mu(x)}{\operatorname{argmin}} \left\{ \left\| \sum_{i=1}^{n} \alpha_i(x) \begin{pmatrix} Y_i - \tau(x)W_i - \mu(x) \\ Z_i \left(Y_i - \tau(x)W_i - \mu(x)\right) \end{pmatrix} \right\|_2 \right\} \tag{3.5}$$

The resulting instrumental forest estimator can be written as :

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha_i(x)[Z_i - \overline{Z}(x)][Y_i - \overline{Y}(x)]}{\sum_{i=1}^{n} \alpha_i(x)[Z_i - \overline{Z}(x)][(W_i - \overline{W}(x)]} \tag{3.6}$$

where $\alpha_i(x)$ is some kind of similarity weights that measure the relevance of individual i to the estimation of $\tau(x)$; $\overline{Z}(x) = \sum_{i=1}^{n} \alpha_i(x)Z_i$; $\overline{Y}(x) = \sum_{i=1}^{n} \alpha_i(x)Y_i$; $\overline{W}(x) = \sum_{i=1}^{n} \alpha_i(x)W_i$.[4]

Similarity weights, $\alpha_i(x)$, are derived from the random forest technique. In the traditional random forest method (Breiman (2001)), many trees are grown in the forest and each terminal leaf of the trees is associated with a specific prediction value. To obtain a prediction for a point of interest of x, x is pushed down through each tree till it hits a terminal leaf and a prediction for x from each tree is observed. The final prediction for x is done by averaging over predictions from all the trees in the forest. Each tree is grown by recursively splitting a random subset of covariate space and each split is chosen to maximize the prediction accuracy of the tree.

Instrumental forest, instead of looking for predictions, counts how often individual i ends up in the same terminal leaf with x among all trees. Similarity weights, $\alpha_i(x)$, are calculated as the following:

$$\alpha_i(x) = \frac{1}{B} \sum_{b=1}^{B} \frac{I\{X_i \in \mathcal{N}_b(x)\}}{|\mathcal{N}_b(x)|}$$

[4]To improve the performance of the instrumental forest in practice, Athey et al. (2019) suggest using the centered outcome $\widetilde{Y}_i$, centered treatment $\widetilde{W}_i$, centered instrumental variable $\widetilde{Z}_i$ to replace Y_i, W_i, Z_i, respectively. $\widetilde{Y}_i = Y_i - \hat{y}^{(-i)}(X_i)$, where $\hat{y}^{(-i)}(X_i)$ is leave-one-out estimate of the marginal expectation of Y_i, computed without using the ith observation. $\widetilde{T}_i$ and $\widetilde{Z}_i$ are computed likewise. $\overline{Y}(x) = \sum_{i=1}^{n} \alpha_i(x)\widetilde{Y}_i$, likewise for $\overline{W}(x)$ and $\overline{Z}(x)$.

where B is the number of trees in the forest; $\mathcal{N}_b(x)$ is the number of individuals that fall into the same terminal leaf as x in tree b; $I\{X_i \in \mathcal{N}_b(x)\}$ is an indicator function that takes on value 1 if individual i ends up in the same terminal leaf as x in tree b. The more frequently individual i falls into the same terminal leaf as x, the higher value $a_i(x)$ receives for individual i in the estimation of $\tau(x)$.

Compared to the traditional random forest, two other features of the instrumental forest used in this paper are worth highlighting. First, the splitting criterion is different. The traditional random forest focuses on delivering predictions; as a result, each split is chosen to maximize prediction accuracy. In the instrumental forest, each split seeks to maximize the heterogeneity in treatment effects across partitions. This splitting criterion helps improve the expected accuracy in predicting treatment effects (Athey et al. 2019). Second, trees grown in the instrumental forest are "honest" trees, meaning that a tree is constructed by using one subsample, while similarity weights derived from this tree are estimated using a different subsample. In other words, $\alpha_i(x)$ is obtained by using trees constructed without individual i. This method is similar in structure to cross-fitting.

We conduct our analysis in R, using the package `grf` developed by Athey et al. (2019). The instrumental forest estimator is obtained by the function `instrumental_forest`. Without formal criteria to guide our choices of the parameters in the function, we set all the parameters at their default values except two parameters: the number of trees and the minimum number of observations in each leaf. Increasing the number of trees reduces Monte Carlo error at the price of increasing computational cost. We choose to grow 15000 trees because the improvement of the estimate stability dramatically slows down when more trees are grown.[5] The choice of the minimum number of observations in each leaf shows a trade-off between bias and variance. Large minimum size of each leaf produces overly simplified tree models that generate less heterogeneity. Previous studies typically set the minimum size between 1 and 10 (Davis and Heller 2017; O'Neill and Weeks 2018; Baiardi and Naghi 2020).

[5] We use the median prediction variance as the measure for estimate stability. Figures 3.1 and 3.2 display the relationship between the number of trees and estimate stability.

In this paper, we choose 10 as our minimum leaf size.

3.4 Results

3.4.1 Local average treatment effect

Results from 2SLS estimations are reported in Table 3.4 and 3.5. First, we find that an additional sibling has a significantly negative impact on children's WAZ. An increase in the number of siblings reduces first-born boy's WAZ by 1.066 standard deviations.[6] It reduces first-born girl's WAZ by 0.816 standard deviations. The magnitude of the effect is smaller for girls. Second, we only find a significant effect of sibling size on HAZ among boys. One additional sibling decreases boy's HAZ by 1.251. Third, all the coefficients on *coverage* are highly significant, meaning *coverage* is significantly related to the number of siblings. The F statistics are above the Stock-Yogo critical value of 10 % maximial IV size, which confirms that our IV is not weak.

Now we turn to the results of estimations using educational measures as dependent variables. We only find a significant impact of sibling size on relative educational attainment among female firstborns.[7] An increase in sibling size leads to a 0.149 decrease in relative educational attainment for girls. We do not find any significant effect of additional siblings on the likelihood of school enrollment for girls or boys. The fact that sibling size has a limited effect on children's education can by explained by the adoption of the Compulsory Education Law in 1986. The law requires that all children attend school for a minimum of nine years. Our results suggest that the Compulsory Education Law does protect children from the negative additional sibling effect on their education during their early school years.

[6] WAZ and HAZ are z-scores therefore the units of measurement are in standard deviations.

[7] Relative educational attainment is measured as a fraction of the mean group educational attainment. The group for calculating the mean is defined by age and wave.

3.4.2 Heterogeneity in local average treatment effects

The generalized random forest procedure adopted from Athey et al. (2019) allows for the estimation of heterogeneity in treatment effects. For girls, the distribution of treatment effects has a relatively tight dispersion. There is relatively wider treatment heterogeneity for boys, as visible in Figure 3.3. The range of treatment effects of additional siblings is much more spread out for boys. This pattern is much more apparent for the distribution of the effects of an additional sibling on health variables than on educational variables.

Another interesting finding is that the distributions of treatment effects on HAZ and WAZ exhibit different patterns among boys and girls. For girls, the distribution of treatment effects on HAZ centers around zero, while almost the entire distribution of treatment effects on WAZ falls below zero. For boys, the distribution of treatment effects on HAZ is relatively similar to the distribution of treatment effects on WAZ. HAZ captures long-term health consequences, while WAZ presents short-term health consequences. This finding suggests that most of the girls experience a negative effect of an additional sibling in the short run, but some of them are able to catch up and even benefit from having more siblings in the long run.

In addition to the distributions of treatment effects, we also report quartile treatment effect means using all four dependent variables in Table 3.6. For boys, the quartile means range from -3.63 to 0.31 for HAZ, and from -2.40 to 0.95 for WAZ. For girls, the quartile means range from -1.26 to 0.71 for HAZ, and from -1.94 to -0.24 for WAZ. The ranges of quartile means for educational variables are much smaller. For boys, the quartile means range from -0.47 to 0.15 for relative education, and from -0.29 to 0.31 for school enrollment. For girls, the quartile means range from -0.83 to 0.29 for relative education, and from -0.21 to 0.03 for school enrollment.

The next step before examining treatment effect heterogeneity is to see whether the generalized random forest algorithm actually captures treatment heterogeneity in general. To perform this simple test we examine whether the subgroup that is hypothesized to have the largest treatment effect does in fact have a larger treatment effect than the other subgroup. In the spirit of Davis and Heller (2017), we first divide our sample into two subsamples based on the treatment effects estimated by the generalized random forest approach. One subsample contains individuals whose treatment effects are above the median treatment effect. The other subsample contains the rest of the individuals. We then repeat the 2SLS regression discussed in Section 3.3.4 within each subgroup and compare the coefficients on sibling size across two subgroups. We present results in Table 3.7.

Consistent with the result in the graphs depicting the distribution of treatment effects we find the largest heterogeneity in treatment effects for the two health-related dependent variables — HAZ and WAZ. In the 2SLS regressions estimating HAZ and WAZ for the subgroup with the largest treatment effects additional siblings reduce health outcomes the most. Furthermore, the difference between the subgroup with the largest treatment effects and the smallest treatment effects are significant for both sexes and both dependent variables. In contrast, for the education outcome variables the treatment effect for the most negative treatment effect subgroup is insignificant as is the difference between this group's treatment effect and the rest of the sample. The results of this simple check of the heterogeneity of treatment effects is in accordance with the graph in Figure 3.3. The results suggest that the generalized random forest algorithm is only able to capture the heterogeneity in treatment effects when the dependent variables are health related variables. Since we do not have enough evidence of whether the algorithm can also capture heterogeneity in treatment effect for education related variables, we focus our heterogeneity analysis on health related variables only.

In Table 3.8 we look at covariate means for the below-median and above-median treatment groups in estimating HAZ, separately for the two sexes. For both boys and girls, children who experience a larger negative effect of sibling size are from households with lower parental health measures, lower parental education, and mother giving birth at an older age. When the dependent variable is WAZ, the differences in coefficients display remarkable similarity, especially for the boys.

The generalized random forest algorithm provides statistics on variable importance. The number of times a variable is used to split the covariate space can be used to determine its importance in the estimation of treatment heterogeneity. Athey et al. (2019) prove that splitting to minimize mean squared errors in the leaf nodes is equivalent to maximizing treatment heterogeneity.[8]

In Figure 3.4 the bars indicate the relative importance of variables for HAZ and WAZ estimation for girls and boys separately. As expected, father's and mother's height variables play an important role in growing the regression trees in the random forest algorithm. Furthermore, since height and weight are correlated both of these variables play an important role in estimations of both of the health-related child quality estimations — HAZ and WAZ. Parental education and mother's age at first birth also play an important role in all estimations. This has important implications from a policy perspective. Governments can implement incentives to improve the education level of the population and help women to give birth during the most optimal times of their life-cycle.

Parental education

In Figure 3.5 and 3.6 we examine the heterogeneity of treatment effects looking at their distribution, color-coded by quartiles, for different levels of parent's education. Some clear patterns emerge in these graphs. When HAZ is the dependent variable, increasing either parent's educational level moves the distribution to the right — the quantity-quality trade-off is attenuated for higher-educated mothers and fathers of girls alike. There is no clear pattern for distributional shifts for boys. It is possible that as parental educational levels go up, there are more resources to invest in child quality.[9] However, if parents set aside more resources to invest in boys regardless of income level, this channel is muted.

[8]Athey et al. (2019) also use a penalty structure for variance for the treatment and control outcomes in the leaves.

[9]Chen and Li (2009) find positive effects of mother's education on height-for-age z-scores using Chinese data

It is also of import that when the dependent variable is WAZ the pattern described above is not observed. The reason for this might be the difference between the two biometrics. Children's weights respond faster to changes in nutrition, while height is slower to respond and it is thought to be a repository of long term nutritional investments. In other words, a child's weight can rebound quickly after bouts of malnutrition but height will reflect accumulated nutritional investments.

Mother's age at first birth

Mother's age at birth of first child also plays an important role in treatment heterogeneity. Figure 3.7 displays the distributional changes as mother's age at birth is increased. As evident in the graphs, the quantity-quality trade-off is amplified for mothers giving birth at a later age.

Two potential channels can explain the positive relationship between mother's age at first birth and the quantity-quality trade-off. It is possible that mothers who gave birth at an older age are better educated, and as a result, face a higher opportunity cost of providing child care themselves. Another possibility is that older mothers experience more income loss when they have an additional child compared to younger mothers. We test these two hypothesis using a simple regression:

$$Y_{ipt} = \beta_3 \widehat{nsib_{ipt}} + \alpha_1 nsib_{ipt} \widehat{\times old\_mother_i} + D_{it}\gamma_4 + C_p\delta_4 + w_t + \epsilon_{ipt} \tag{3.7}$$

Equation (3.7) is the same as equation (3.3) except we include an interaction term between number of siblings and a dummy variable old_mother_i. old_mother_i takes on value 1 if mother's age at first birth is above the sample median. The dependent variables are whether the mother provides child care herself and the log of mother's income.

$\widehat{nsib_{ipt}}$ and $nsib_{ipt} \widehat{\times old\_mother_i}$ are estimated using equation (8) and equation (9), respectively.

$$nsib_{ipt} = \beta_4 coverage_i + \alpha_2 (coverage_i \times old\_mother_i) \qquad (3.8)$$
$$+ D_{it}\gamma_5 + C_p\delta_5 + w_t + v_{ipt}$$

$$nsib_{ipt} \times old\_mother_i = \beta_5 coverage_i + \alpha_3 (coverage_i \times old\_mother_i) \qquad (3.9)$$
$$+ D_{it}\gamma_6 + C_p\delta_6 + w_t + \eta_{ipt}$$

Table 3.10 displays the results. Having more children increases the probability of a mother providing child care by nearly 40%. The effect is not significantly different between younger mothers and older mothers. An increase in family size decreases mother's income by 28%. The income penalty for having more children is even larger for older mothers. Our findings suggest that a larger income penalty might be the channel behind the positive relationship between mother's age at first birth and the quantity-quality trade-off.

The results in table 3.10 are consistent with the literature on this subject. Putz and Engelhardt (2014) find more dramatic wage losses from giving birth at a later age. They offer several explanations for this "late birth wage gap". One is that women may go through a transitional phase in the years during which later births occur.[10] Another possible explanation is statistical discrimination. Putz and Engelhardt (2014) infer this since they only find a late-birth wage penalty when they use a mother's biological age, but not when they use the number of elapsed years from joining the labor force to giving birth, as a dependent variable.

It is worth noting in the above graph that the late birth penalty is a feature of the mothers of boys as well. This effect is in stark contrast to the treatment effect heterogeneity

[10] In analyzing Ukrainian data Nizalova et al. (2016) find that delaying birth and low education incur the largest birth-related wage penalties.

graphs on previous pages. While there was no change in quantity-quality trade-offs for boys for varying levels of parents education there is a dramatic late-birth penalty in terms of child quality as apparent in the graph. There may need to be more research conducted in this area to examine the reasons for this.

3.5 Conclusion

In this paper we use a recently developed machine learning approach to study the child quantity-quality trade-off. First we use the severity of the OCP implementation across Chinese provinces combined with the overlap of these policies with a mother's fertile years to discern exogenous variation in family size. The resulting estimation of LATE is a starting point in the examination of the heterogeneous treatment effects. We utilize a generalized random forest algorithm because it is an apt technique to detect treatment heterogeneity since it is minimizing mean square errors by contemporaneously maximizing treatment heterogeneity in the covariate space.

We detect large heterogeneous effects of family size on health-related quality measures but not on educational outcomes. Our findings suggest that parental education, income, the sex of the child and the timing of mother's first birth play the largest role in the quantity-quality trade-off. The quality penalty, especially on health outcomes, is the largest for girls. Children with more educated parents experience a smaller quality penalty. Also, an increase in mother's age at first birth induces a child quality penalty. Our findings suggest that the previous results in the literature estimating the effect of the quantity-quality trade-off may be contradictory to each other because of the instruments they were using to deal with the endogeneity of family size. As such these studies detected different complier groups and therefore different treatment effects of quantity on quality. Our approach uses a data-driven algorithm to pinpoint all sources of treatment heterogeneity. Because of this, our study can inform policy makers on precisely where to concentrate their work in their efforts to maximize child quality.

Table 3.1: Heterogeneity in the OCP implementation

Province	OCP from	Exemptions			
		Both spouses minority	One spouse minority	Both spouses arigultural	Both spouses agricultural with a girl
Beijing	1979				
Liaoning	1980	1982-1984		1985b-	1985-
Heilongjiang	1979	1981-1983 1994a-			1990-
Shanghai	1979				
Jiangsu	1979				
Shandong	1980	1984-			1986-
Henan	1981	1990c-			1990-
Hubei	1981				1988-
Hunan	1982	1990d-		1990e-	1987-
Guangxi	1982	1989a-	1985-1988a		1989-
Guizhou	1982	1982-1998		1982-1987 1988e-	1998-

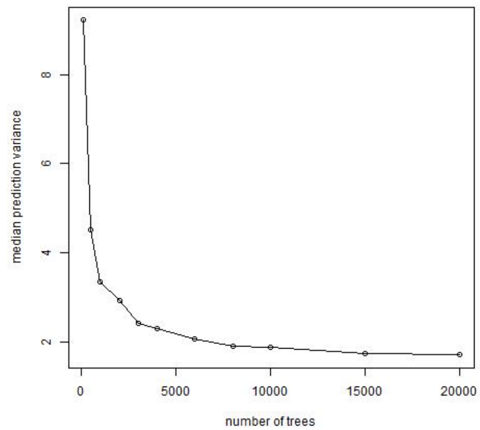
(a) only minorities with a total population less than 10 million (Manchu and Zhuang have populations exceeding 10 millions); (b) one spouse must belong to a minority, whose total population is less than 10 million; (c) both spouses must be agricultural; (d) one spouse must be agricultural; (e) one spouse must belong to a minority.

Source: Huang (2021)

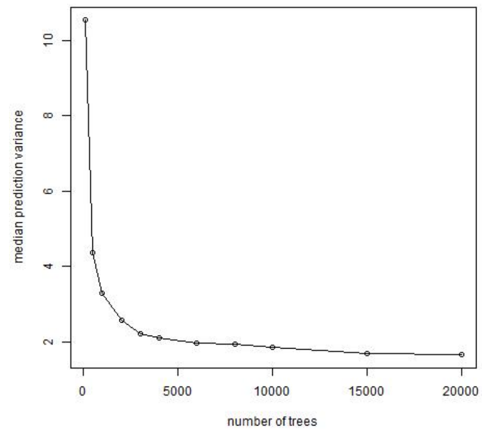
Table 3.2: Summary statistics

	All		Boys		Girls	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Boy	51%	49%				
Age	11.20	3.30	11.19	3.31	11.20	3.28
HAZ	-0.42	1.35	-0.37	1.37	-0.46	1.33
WAZ	-0.41	1.33	-0.26	1.36	-0.56	1.27
Years of schooling	5.48	3.13	5.43	3.13	5.54	3.13
Currently enrolled in school	0.90	0.29	0.90	0.28	0.90	0.29
Father's height (cm)	167.73	5.79	167.71	5.76	167.74	5.82
Mother's height (cm)	156.88	5.85	157.13	5.62	156.61	6.06
Father's weight (kg)	64.91	10.06	64.93	9.84	64.89	10.29
Mother's weight (kg)	55.73	8.56	55.88	8.63	55.58	8.49
Mother's age at first birth	24.62	3.48	24.62549	3.39	24.63	3.57
Father's years of schooling	9.32	3.44	9.331241	3.43	9.30	3.45
Mother's years of schooling	8.24	4.06	8.264473	4.04	8.21	4.08

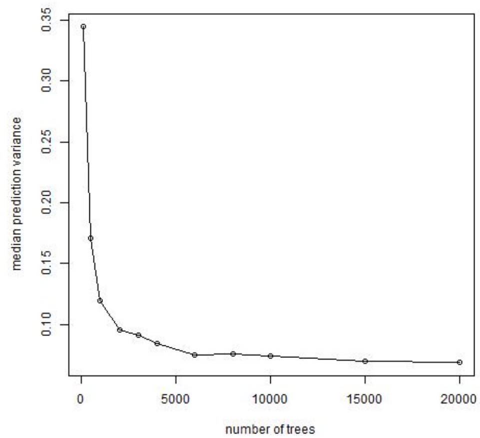
Notes: Our sample comes from the China Health and Nutrition Survey (CHNS). 8461 observations are included in the sample. HAZ is height-for-age z-score. WAZ is weight-for-age z-score.



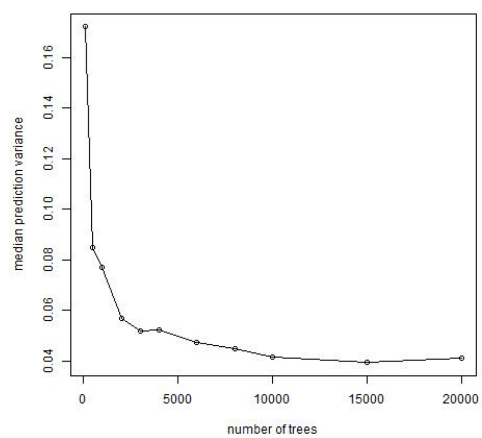
(a) HAZ median prediction variance



(b) WAZ median prediction variance



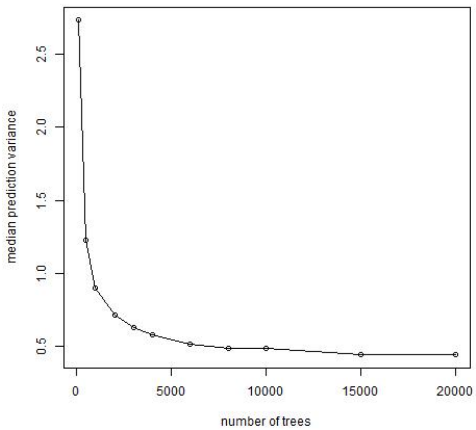
(c) relative education median prediction variance



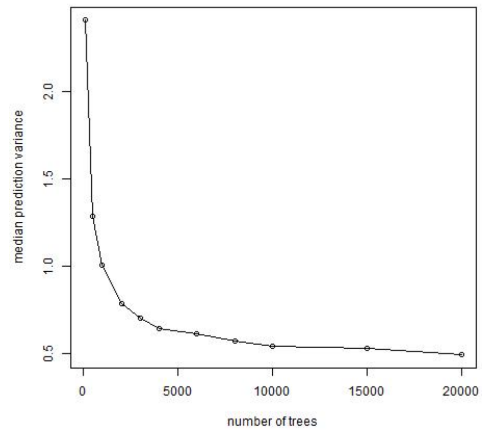
(d) school enrollment median prediction variance

Figure 3.1: Number of trees and median variance for boys

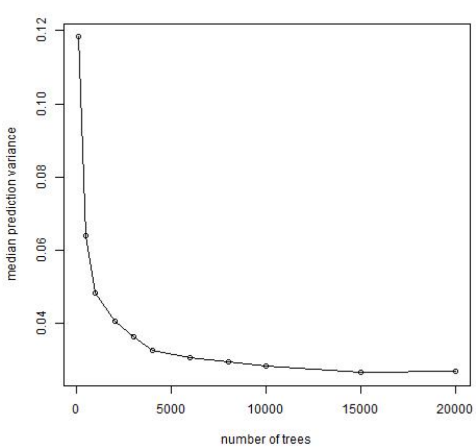
Notes: Variance of predicted treatment effect is obtained using `instrumental_forest` from R package `grf`. All the parameters are set at their default values except "num.trees". In (a)-(d), independent variable is number of siblings; instrumental variable is *coverage*, the share of a mother's prime fertility years covered by the strict OCP; covariate variables include a set of individual characteristic variables, a set of pre-OCP provincial characteristic variables and wave dummies. In (a) and (b), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. In (c) and (d), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Provincial characteristic variables are the same for all panels, including sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976.



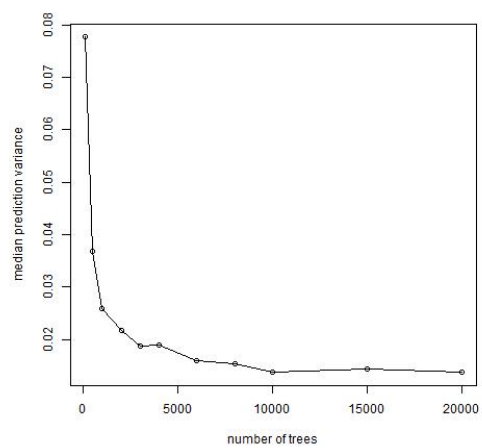
(a) HAZ median prediction variance



(b) WAZ median prediction variance



(c) relative education median prediction variance



(d) school enrollment median prediction variance

Figure 3.2: Number of trees and median variance for girls

Notes: Variance of predicted treatment effect is obtained using `instrumental_forest` from R package `grf`. All the parameters are set at their default values except "num.trees". In (a)-(d), independent variable is number of siblings; instrumental variable is *coverage*, the share of a mother's prime fertility years covered by the strict OCP; covariate variables include a set of individual characteristic variables, a set of pre-OCP provincial characteristic variables and wave dummies. In (a) and (b), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. In (c) and (d), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Provincial characteristic variables are the same for all panels, including sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976.

Table 3.3: Exposure to the strict OCP and sibling sex composition

	Boys		Girls	
	(1) Male second birth	(2) Fraction of male siblings	(3) Male second birth	(4) Fraction of male siblings
coverage	0.0135	1.799	-0.123***	-10.85***
	(0.0498)	(4.7090)	(0.0362)	(3.1830)
N	1609	1609	2095	2095

Notes: OLS regression is used in all columns. *coverage* is the share of a mother's prime fertility years covered by the strict OCP. A set of individual characteristic variables, a set of pre-OCP provincial characteristic variables, and wave dummies are included in all the regressions. Individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Provincial characteristic variables include sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 3.4: 2SLS: Height for age z-score and weight for age z-score

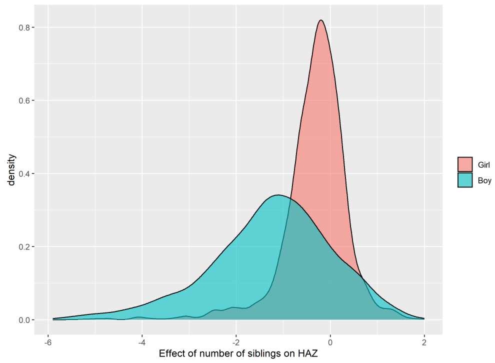
	Boys		Girls	
	(1) HAZ	(3) WAZ	(5) HAZ	(7) WAZ
Number of siblings	-1.251***	-1.066***	-0.304	-0.816***
	(0.363)	(0.337)	(0.198)	(0.209)
First-stage: Number of siblings				
coverage	-0.186***	-0.189***	-0.296***	-0.296***
	(0.0282)	(0.0282)	(0.0337)	(0.0336)
Cragg-Donald Wald F statistics	43.763	44.771	77.097	77.670
Stock-Yogo critical value: 10% maximal IV size	16.38	16.38	16.38	16.38
Individual controls	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes	Yes
N	3648	3681	3456	3479

Notes: This table reports 2SLS regression results for HAZ and WAZ among boys and girls separately. The independent variable of interest, number of siblings, is instrumented by *coverage*, the share of a mother's prime fertility years covered by the strict OCP. Individual controls include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. Province controls include nclude sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

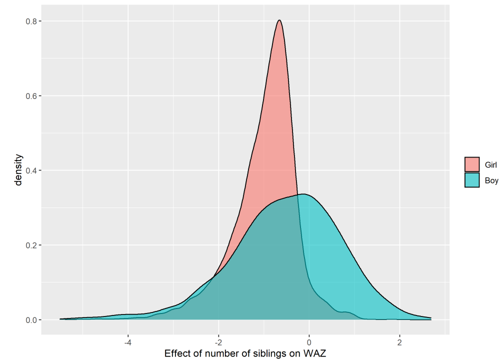
Table 3.5: 2SLS: relative education and school enrollment

	Boys		Girls	
	(5) Relative education	(7) School enrollment	(6) Relative education	(8) School enrollment
Number of siblings	-0.158	-0.0480	-0.149**	-0.0185
	(0.103)	(0.0739)	(0.0670)	(0.0510)
First-stage: Number of siblings				
coverage	-0.197***	-0.200***	-0.289***	-0.299***
	(0.0284)	(0.0277)	(0.0336)	(0.0331)
Cragg-Donald Wald F statistics	48.217	51.828	73.991	81.331
Stock-Yogo critical value: 10% maximal IV size	16.38	16.38	16.38	16.38
Individual controls	Yes	Yes	Yes	Yes
Province controls	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes	Yes
N	3648	3646	3462	3452

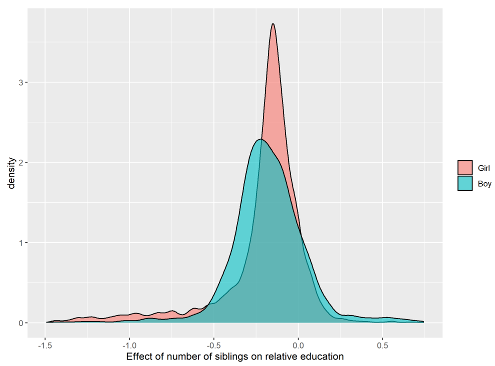
Notes: This table reports 2SLS regression results for relative education and school enrollment status among boys and girls separately. The independent variable of interest, number of siblings, is instrumented by *coverage*, the share of a mother's prime fertility years covered by the strict OCP. Individual controls include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Province controls include nclude sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.



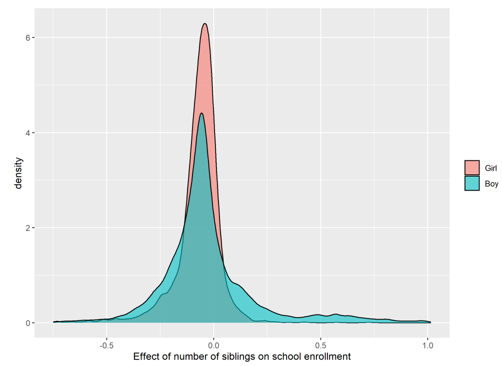
(a) Distribution of effects of number of siblings on HAZ by gender



(b) Distribution of effects of number of siblings on WAZ by gender



(c) Distribution of effects of number of siblings on relative education by gender



(d) Distribution of effects of number of siblings on relative education by gender

Figure 3.3: Distribution of effects of number of siblings by gender

Notes: Predicted treatment effect is obtained using `instrumental_forest` from R package `grf`. "num.trees" is set to be 15000. "min.node.size" is set to be 10. The rest parameters are set at their default values. In (a)-(d), independent variable is number of siblings; instrumental variable is *coverage*, the share of a mother's prime fertility years covered by the strict OCP; covariate variables include a set of individual characteristic variables, a set of pre-OCP provincial characteristic variables and wave dummies. In (a) and (b), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. In (c) and (d), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Provincial characteristic variables are the same for all panels, including sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976.

Table 3.6: Mean treatment effects by quartile

	HAZ		WAZ		Relative Educ		Enrolled in school	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Mean	-1.41967	-0.26605	-0.58843	-0.98916	-0.16829	-0.21519	-0.02607	-0.07343
Mean in quartile 1 (0 - 25%)	-3.62738	-1.2622	-2.40255	-1.94055	-0.46563	-0.83078	-0.29172	-0.2085
Mean in quartile 2 (25% - 50%)	-1.55903	-0.41205	-0.83298	-1.07248	-0.23932	-0.19432	-0.0902	-0.07771
Mean in quartile 3 (50% - 75%)	-0.79682	-0.09441	-0.06766	-0.7019	-0.12462	-0.12065	-0.02931	-0.03606
Mean in quartile 4 (75% - 100%)	0.305095	0.705417	0.950052	-0.24063	0.156737	0.285109	0.30726	0.028657

Notes: Predicted treatment effect is obtained using `instrumental_forest` from R package `grf`. "num.trees" is set to be 15000. "min.node.size" is set to be 10. The rest parameters are set at their default values. In (a)-(d), independent variable is number of siblings; instrumental variable is *coverage*, the share of a mother's prime fertility years covered by the strict OCP; covariate variables include a set of individual characteristic variables, a set of pre-OCP provincial characteristic variables and wave dummies. In (a) and (b), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. In (c) and (d), individual characteristic variables include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies. Provincial characteristic variables are the same for all panels, including sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976.

Table 3.7: Heterogeneity in local average treatment effect by predicted treatment effect

	HAZ		WAZ		Relative Educ		Enrolled in school	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Subgroup with more	-3.6584***	-0.9252***	-2.9751***	-2.1818***	-0.2392	0.0600	-0.0756	-0.0343
negative estimated effects	(0.9695)	(0.3356)	(0.9223)	(0.6028)	(0.1653)	(0.1459)	(0.0846)	(0.0867)
Rest of sample	-0.1278	0.2022	0.1260	-0.1906	-0.0706	-0.2798***	-0.0009	-0.0262
	(0.5285)	(0.2649)	(0.5680)	(0.2560)	(0.1491)	(0.0782)	(0.1255)	(0.0679)
P-value, test of subgroup difference	0.0014	0.0084	0.0042	0.0024	0.4488	0.0400	0.6216	0.9415

Notes: We divide our sample into two subsamples based on the treatment effects estimated by using `instrumental_forest` from R package `grf`. Subgroup with more negative estimated effects contains individuals whose estimated treatment effects are below the median treatment effect. Rest of sample contains individuals whose estimated treatment effects are above the median treatment effect. We repeat 2SLS regressions in Table 3.4 and 3.5 using these two subgroups separately. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 3.8: Summary statistics by predicted treatment effect on HAZ

	HAZ					
	Boys			Girls		
Variable	Subgroup with more negative effects	Rest of sample	Difference	Subgroup with more negative effects	Rest of sample	Difference
child's age	11.2438	11.1571	0.0867	11.5695	10.9490	0.6205***
mother's age at first birth	26.0754	23.5776	2.4978***	24.8749	24.4583	0.4166***
father's height	166.5394	168.5676	-2.0282***	166.0556	168.9491	-2.8935***
mother's height	156.4396	157.6337	-1.1941***	156.6803	156.5745	0.1057
father's weight	63.4930	65.9803	-2.4873***	62.5956	66.5294	-3.9338***
mother's weight	54.6330	56.7928	-2.1598***	57.7451	54.0413	3.7038***
father's years of schooling	9.1965	9.4443	-0.2479**	8.4369	10.0430	-1.6061***
mother's years of schooling	8.2108	8.3061	-0.0953	7.0844	9.0788	-1.9944***

Notes: Subgroup with more negative effects contains individuals whose estimated treatment effects of family size on HAZ are below the mean estimated treatment effect. Rest of sample contains individuals whose estimated treatment effects of family size on HAZ are above the mean estimated treatment effect. We report covariate means for each subgroup. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Table 3.9: Summary statistics by predicted treatment effect on WAZ

	WAZ					
	Boys			Girls		
Variable	Subgroup with more negative effects	Rest of sample	Difference	Subgroup with more negative effects	Rest of sample	Difference
child's age	11.2470	11.1542	0.0928	10.5014	11.7165	-1.2151***
mother's age at first birth	25.7475	23.8014	1.9462***	25.2796	24.1642	1.1154***
father's height	167.1398	168.1405	-1.0006***	167.2004	168.1376	-0.9373***
mother's height	155.8743	158.0570	-2.1826***	156.1625	156.9476	-0.7852***
father's weight	63.9790	65.6404	-1.6613***	63.8573	65.6387	-1.7814***
mother's weight	53.8134	57.4095	-3.5961***	55.7574	55.4570	0.3004
father's years of schooling	9.0311	9.5875	-0.5564***	9.4136	9.2212	0.1924*
mother's years of schooling	8.1587	8.3478	-0.1891	8.5627	7.9445	0.6182***

Notes: Subgroup with more negative effects contains individuals whose estimated treatment effects of family size on WAZ are below the mean estimated treatment effect. Rest of sample contains individuals whose estimated treatment effects of family size on HAZ are above the mean estimated treatment effect. We report covariate means for each subgroup. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

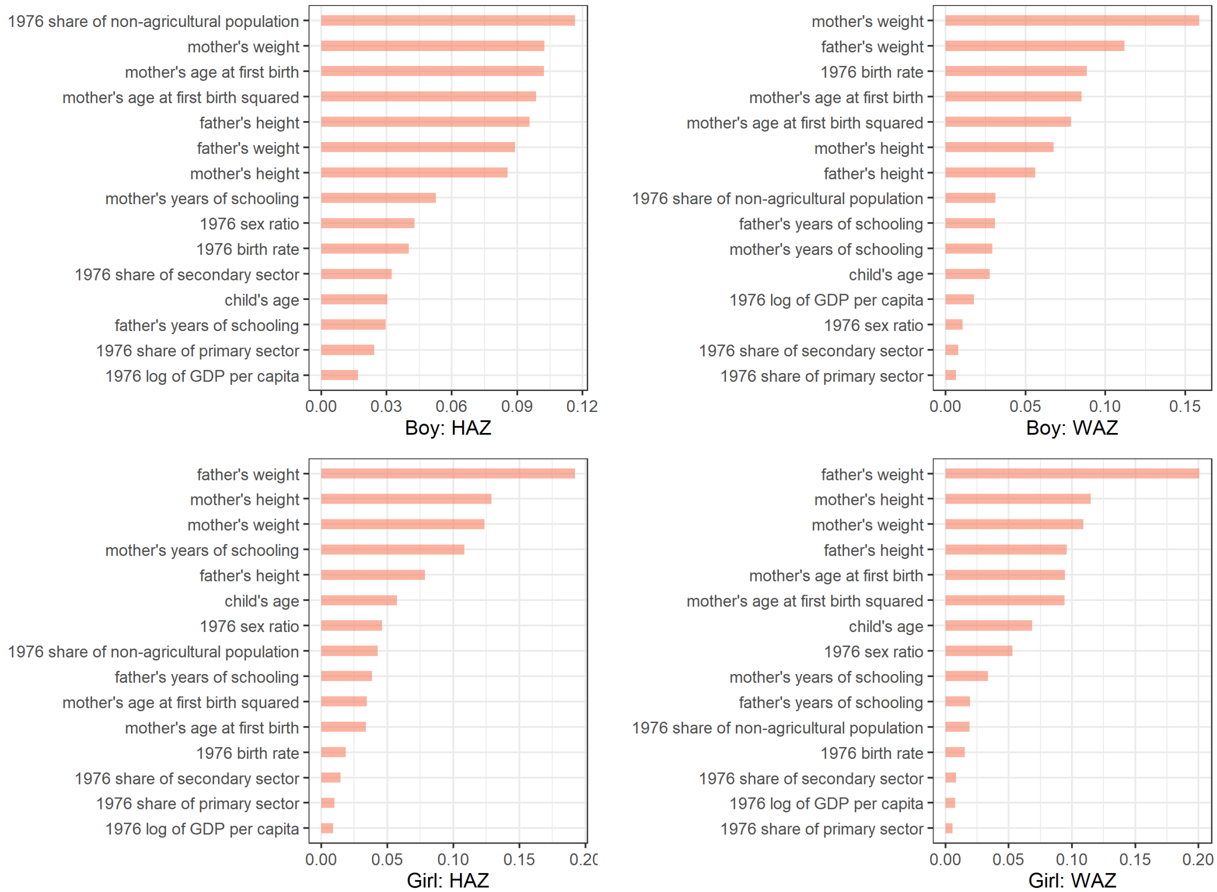
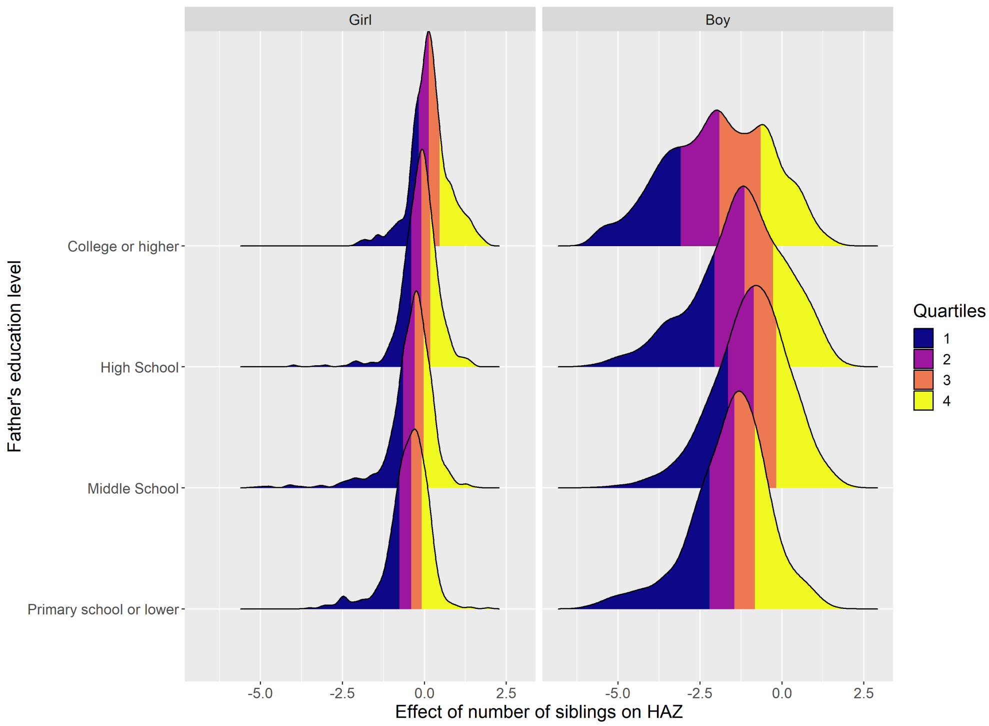
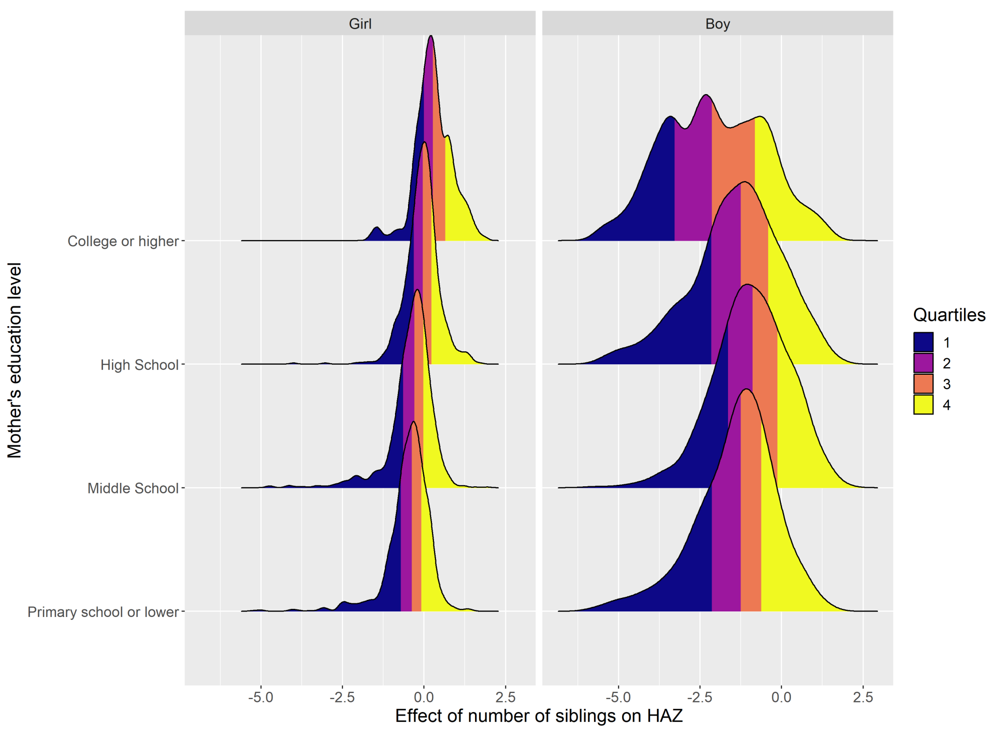


Figure 3.4: Health variable importance

Notes: Variable importance is obtained using `instrumental_forest` from R package `grf`. The variable importance measures how frequently a covariate is used in the tree splitting process.

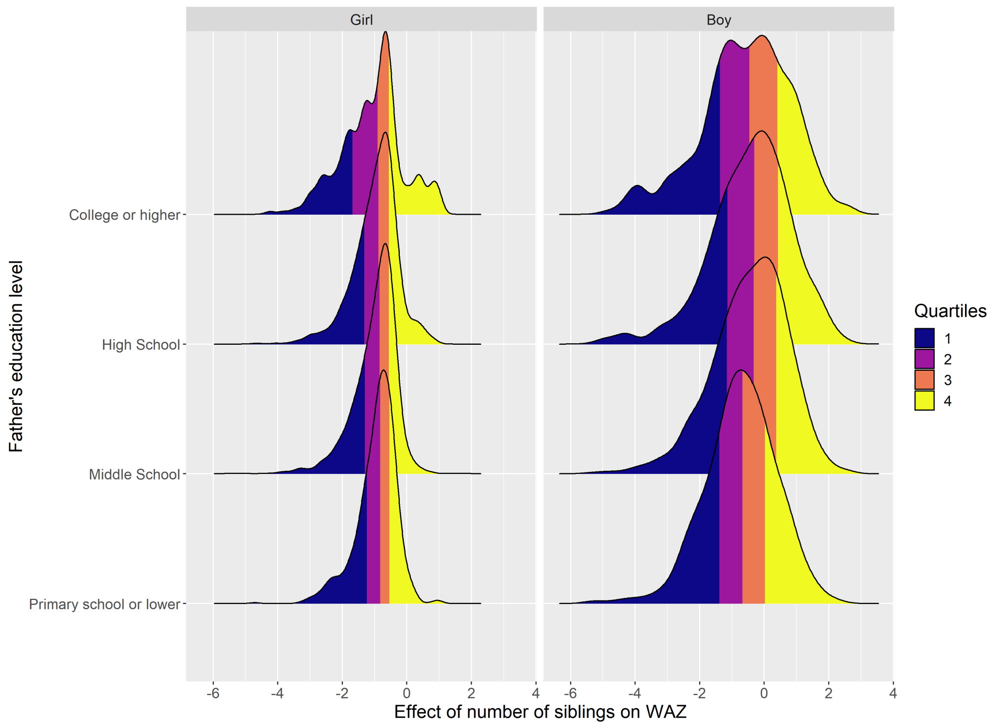


(a) by father's education

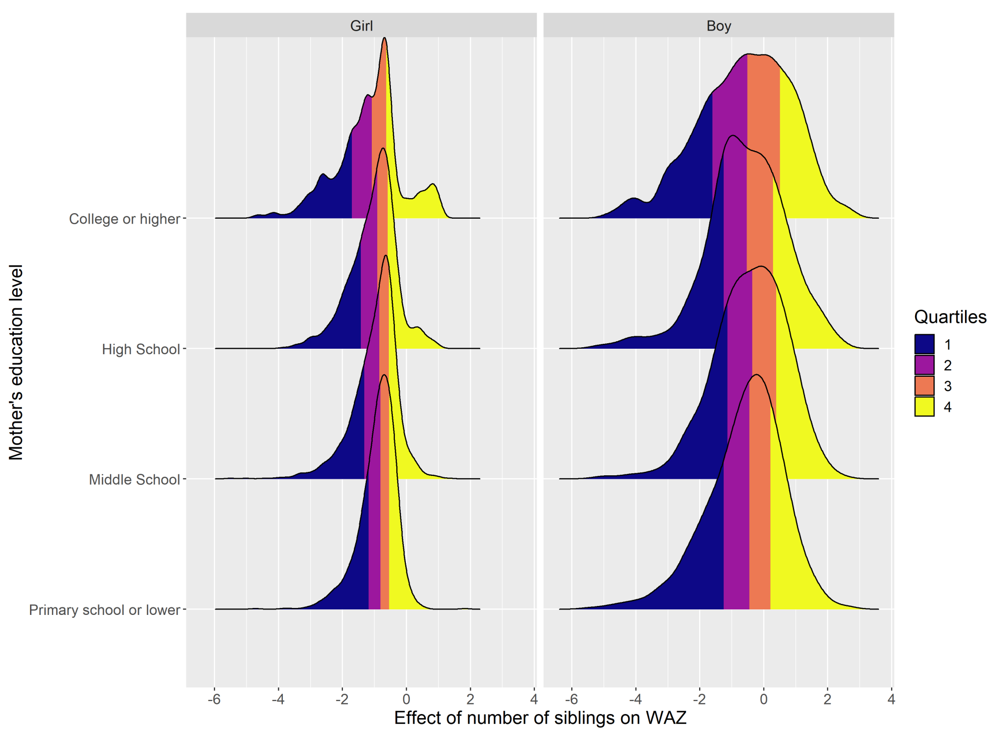


(b) by mother's education

Figure 3.5: distribution of effects of number of children on HAZ by parent's education level

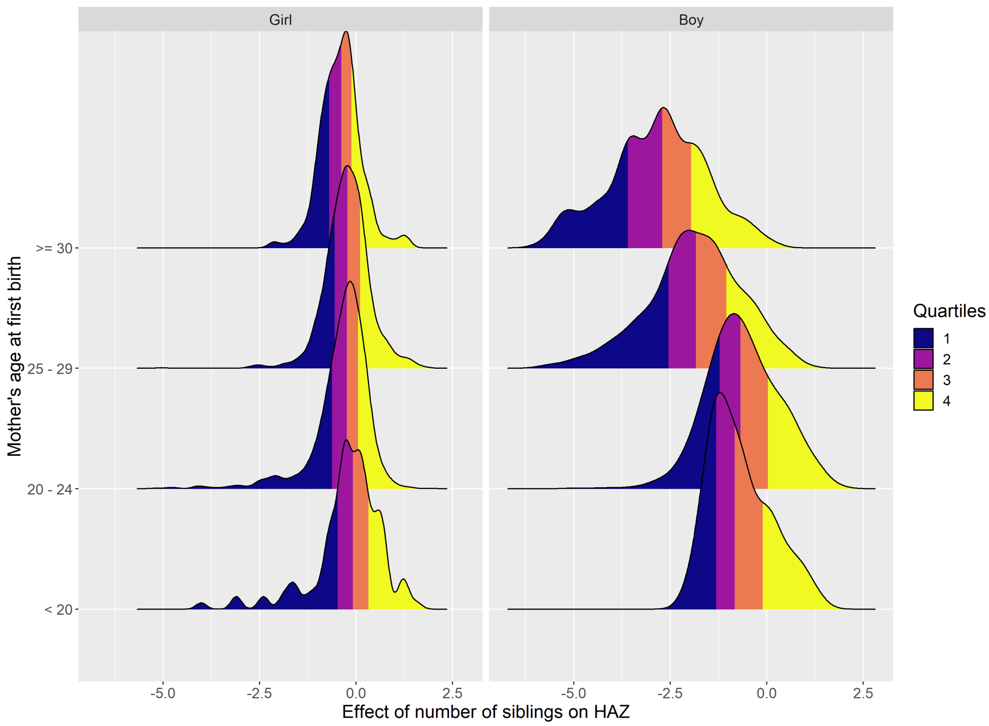


(a) by father's education

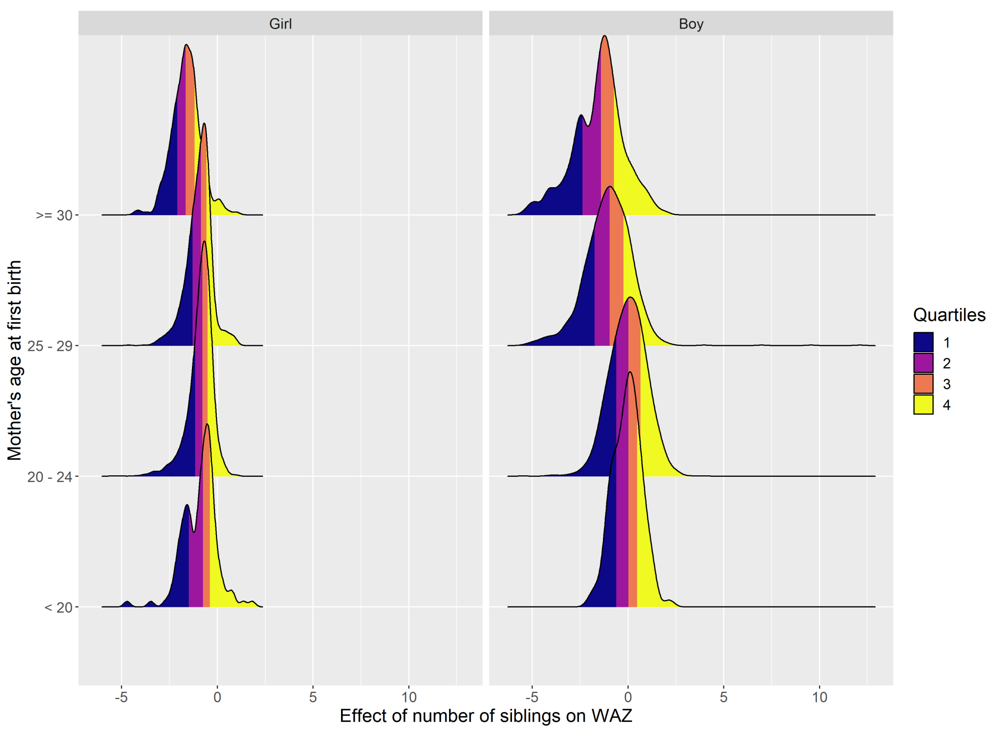


(b) by mother's education

Figure 3.6: distribution of effects of number of children on WAZ by parent's education level



(a) HAZ



(b) WAZ

Figure 3.7: distribution of effects of number of children on health by child's age group

Table 3.10: Effect of number of siblings by mother's age at first birth

	(1) child care	(2) log of mother's income
number of siblings	0.389***	-0.280*
	(0.106)	(0.161)
number of siblings $\times$old mother	-0.0270	-0.292*
	(0.0578)	(0.172)
N	3796	2836

Notes: Both columns report 2SLS estimates. old mother is a dummy variable that takes on value 1 if mother's age at first birth is above the sample median mother's age at first birth. Both number of siblings and number of siblings $\times$ old mother are instrumented by *coverage* and *coverage*$\times$ old mother. *coverage* is the share of a mother's prime fertility years covered by the strict OCP. Individual controls, province controls, and wave dummies are included in regressions of column (1) and (2). Individual controls include child's age, mom's age at first birth, mom's age at first birth squared, parental education-level dummies, parental weight, parental height, dummy variable for missing parental weight, dummy variable for missing parental height. Province controls include nclude sex ratio in 1976, birth rate in 1976, log of GDP per capita in 1976, share of noon-agricultural population in 1976, share of primary industry in GDP in 1976, and share of secondary industry in GDP in 1976. Standard errors are reported in parentheses. * p <0.1, ** p <0.05, *** p <0.01.

Bibliography

Ahn, N. (1994). "Effects of the one-child family policy on second and third births in hebei, shaanxi and shanghai." *Journal of Population Economics*, *7*(1), 63–78.

Angrist, J., and Imbens, G. (1995). "Identification and estimation of local average treatment effects."

Angrist, J., Lavy, V., and Schlosser, A. (2010). "Multiple experiments for the causal link between the quantity and quality of children." *Journal of Labor Economics*, *28*(4), 773–824.

Angrist, J. D., Evans, W. N., et al. (1998). "Children and their parents' labor supply: Evidence from exogenous variation in family size." *American Economic Review*, *88*(3), 450–477.

Antonucci, T. C., and Jackson, J. S. (1990). "The role of reciprocity in social support."

Athey, S., Tibshirani, J., and Wager, S. (2019). "Generalized random forests." *The Annals of Statistics*, *47*(2), 1148–1178.

Baiardi, A., and Naghi, A. (2020). "The value added of machine learning to causal inference: Evidence from revisited studies."

Bandura, A., and Walters, R. H. (1977). *Social learning theory*, vol. 1. Englewood cliffs Prentice Hall.

Banerjee, A., Meng, X., and Qian, N. (2010). "The life cycle model and household savings: Micro evidence from urban china." *National Bureau of Demographic Dividends Revisited*, *21*.

Becker, G. S. (1974). "A theory of social interactions." *Journal of political economy*, *82*(6), 1063–1093.

Becker, G. S. (1991). *A treatise on the family: Enlarged edition*. Harvard university press.

Becker, G. S. (1998). "A treatise on the family."

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

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

Behrman, J. R., Pollak, R. A., and Taubman, P. (1986). "Do parents favor boys?" *International Economic Review*, 33–54.

Benham, L. (1974). "Benefits of women's education within marriage." *Journal of political Economy*, *82*(2, Part 2), S57–S71.

Bernheim, B. D., Shleifer, A., and Summers, L. H. (1986). "The strategic bequest motive." *Journal of labor Economics*, *4*(3, Part 2), S151–S182.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). "The more the merrier? the effect of family size and birth order on children's education." *The Quarterly Journal of Economics*, *120*(2), 669–700.

Bound, J., and Solon, G. (1999). "Double trouble: on the value of twins-based estimation of the return to schooling." *Economics of Education Review*, *18*(2), 169–182.

Breiman, L. (2001). "Random forests." *Machine learning*, *45*(1), 5–32.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). "Beyond late with a discrete instrument." *Journal of Political Economy*, *125*(4), 985–1039.

Bures, R. M., Koropeckyj-Cox, T., and Loree, M. (2009). "Childlessness, parenthood, and depressive symptoms among middle-aged and older adults." *Journal of Family Issues*, *30*(5), 670–687.

Cáceres-Delpiano, J. (2006). "The impacts of family size on investment in child quality." *Journal of Human Resources*, *41*(4), 738–754.

Cáceres-Delpiano, J., and Simonsen, M. (2012). "The toll of fertility on mothers' wellbeing." *Journal of health economics*, *31*(5), 752–766.

Cai, F., Giles, J., and Meng, X. (2006). "How well do children insure parents against low retirement income? an analysis using survey data from urban china." *Journal of public Economics*, *90*(12), 2229–2255.

Canning, D., and Schultz, T. P. (2012). "The economic consequences of reproductive health and family planning." *The Lancet*, *380*(9837), 165–171.

Cao, Y. (2019). "Fertility and labor supply: evidence from the one-child policy in china." *Applied Economics*, *51*(9), 889–910.

Chen, Y., and Fang, H. (2018). "The long-term consequences of having fewer children in old age: evidence from china's "later, longer, fewer" campaign."

Chen, Y., and Huang, Y. (2018). "The power of the government: China's family planning leading. group and the fertility decline since 1970." Tech. rep., GLO Discussion Paper.

Chen, Y., and Li, H. (2009). "Mother's education and child health: Is there a nurturing effect?" *Journal of health economics*, *28*(2), 413–426.

Chen, Y., Li, H., and Meng, L. (2013). "Prenatal sex selection and missing girls in china: Evidence from the diffusion of diagnostic ultrasound." *Journal of Human Resources*, *48*(1), 36–70.

Chen, Y., Wang, L., and Zhang, M. (2018). "Informal search, bad search?: the effects of job search method on wages among rural migrants in urban china." *Journal of Population Economics*, *31*(3), 837–876.

Chunling, L. (2006). "Sociopolitical change and inequality in educational opportunity: Impact of family background and institutional factors on educational attainment (1940-2001)." *Chinese Sociology & Anthropology*, *38*(4), 6–36.

Cong, Z., and Silverstein, M. (2011). "Intergenerational exchange between parents and migrant and nonmigrant sons in rural china." *Journal of Marriage and Family*, *73*(1), 93–104.

Cox, D. (1987). "Motives for private income transfers." *Journal of political economy*, *95*(3), 508–546.

Cox, D., Hansen, B. E., and Jimenez, E. (2004). "How responsive are private transfers to income? evidence from a laissez-faire economy." *Journal of Public Economics*, *88*(9-10), 2193–2219.

Croll, E., Kane, P., and Davin, D. (1985). *China's one-child family policy*. Springer.

Davis, J., and Heller, S. B. (2017). "Using causal forests to predict treatment heterogeneity: An application to summer jobs." *American Economic Review*, *107*(5), 546–50.

Duflo, E. (2001). "Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment." *American economic review*, *91*(4), 795–813.

Ebenstein, A. (2010). "The "missing girls" of china and the unintended consequences of the one child policy." *Journal of Human resources*, *45*(1), 87–115.

Edlund, L., Li, H., Yi, J., and Zhang, J. (2007). "Sex ratios and crime: evidence from china's one-child policy." Tech. rep., IZA Discussion Papers.

Fang, H., and Feng, J. (2018). "The chinese pension system."

Farré, L., Klein, R., and Vella, F. (2013). "A parametric control function approach to estimating the returns to schooling in the absence of exclusion restrictions: an application to the nlsy." *Empirical Economics*, *44*(1), 111–133.

Ge, S., Yang, D. T., and Zhang, J. (2018). "Population policies, demographic structural changes, and the chinese household saving puzzle." *European Economic Review*, *101*, 181–209.

Gelbach, J. B. (2002). "Public schooling for young children and maternal labor supply." *American Economic Review*, *92*(1), 307–322.

Gibson, J., Olivia, S., and Rozelle, S. (2011). "How widespread are nonlinear crowding out effects? the response of private transfers to income in four developing countries." *Applied Economics*, *43*(27), 4053–4068.

Golley, J., and Kong, S. T. (2018). "Inequality of opportunity in china's educational outcomes." *China Economic Review*, *51*, 116–128.

Gove, W. R., and Geerken, M. R. (1977). "The effect of children and employment on the mental health of married men and women." *Social forces*, *56*(1), 66–76.

Griliches, Z. (1979). "Sibling models and data in economics: Beginnings of a survey." *Journal of political Economy*, *87*(5, Part 2), S37–S64.

Guifu, C., and Hamori, S. (2009). "Economic returns to schooling in urban china: Ols and the instrumental variables approach." *China Economic Review*, *20*(2), 143–152.

Guo, G., and VanWey, L. K. (1999). "Sibship size and intellectual development: Is the relationship causal?" *American Sociological Review*, 169–187.

Han, H. (2010). "Trends in educational assortative marriage in china from 1970 to 2000." *Demographic Research*, *22*, 733–770.

Hanewald, K., Jia, R., and Liu, Z. (2021). "Why is inequality higher among the old? evidence from china." *China Economic Review*, *66*, 101592.

Harvey, A. C. (1976). "Estimating regression models with multiplicative heteroscedasticity." *Econometrica: Journal of the Econometric Society*, 461–465.

Hedrich, J. P. (2011). "The heterogeneous impact of the quantity-quality tradeoff: Looking beyond the mean."

Huang, C., Li, H., Liu, P. W., and Zhang, J. (2009). "Why does spousal education matter for earnings? assortative mating and cross-productivity." *Journal of Labor Economics*, *27*(4), 633–652.

Huang, W., and Zhou, Y. (2015). "One-child policy, marriage distortion, and welfare loss."

Huang, Y. (2021). "Family size and children's education: Evidence from the one-child policy in china." *Population Research and Policy Review*, 1–26.

Isacsson, G. (2007). "Twin data vs. longitudinal data to control for unobserved variables in earnings functions–which are the differences?" *Oxford Bulletin of Economics and Statistics*, *69*(3), 339–362.

Islam, A., Smyth, R., et al. (2010). "Children and parental health: Evidence from china."

Tech. rep., Citeseer.

Jain-Chandra, M. S., Khor, N., Mano, R., Schauer, J., Wingender, M. P., and Zhuang, J. (2018). *Inequality in China–trends, drivers and policy remedies*. International Monetary Fund.

Jensen, R. T. (2004). "Do private transfers 'displace'the benefits of public transfers? evidence from south africa." *Journal of Public Economics*, *88*(1-2), 89–112.

Jolly, N. A. (2019). "Female earnings and the returns to spousal education over time." *Journal of Family and Economic Issues*, *40*(4), 691–709.

Kessler, D. (1991). "Birth order, family size, and achievement: Family structure and wage determination." *Journal of Labor Economics*, *9*(4), 413–426.

Kim, J., and Choi, K. (2008). "Closing the gap: Modeling within-school variance heterogeneity in school effect studies." *Asia Pacific Education Review*, *9*(2), 206–220.

Klein, R., and Vella, F. (2010). "Estimating a class of triangular simultaneous equations models without exclusion restrictions." *Journal of Econometrics*, *154*(2), 154–164.

Lee, J. (2008). "Sibling size and investment in children's education: An asian instrument." *Journal of Population Economics*, *21*(4), 855–875.

Lee, K., Bargagli-Stoffi, F. J., and Dominici, F. (2020). "Causal rule ensemble: Interpretable inference of heterogeneous treatment effects." *arXiv preprint arXiv:2009.09036*.

Lei, L. (2013). "Sons, daughters, and intergenerational support in china." *Chinese Sociological Review*, *45*(3), 26–52.

Lei, X., Giles, J., Hu, Y., Park, A., Strauss, J., and Zhao, Y. (2012). "Patterns and correlates of intergenerational non-time transfers: evidence from charls." *World Bank Policy Research Working Paper*, (6076).

Lei, X., Strauss, J., Tian, M., and Zhao, Y. (2015). "Living arrangements of the elderly in china: evidence from the charls national baseline." *China economic journal*, *8*(3), 191–214.

Li, B., and Zhang, H. (2017). "Does population control lead to better child quality? evidence from china's one-child policy enforcement." *Journal of Comparative Economics*, *45*(2), 246–260.

Li, H., Yi, J., and Zhang, J. (2011). "Estimating the effect of the one-child policy on the sex ratio imbalance in china: Identification based on the difference-in-differences." *Demography*, *48*(4), 1535–1557.

Li, H., Zhang, J., and Zhu, Y. (2008). "The quantity-quality trade-off of children in a developing country: Identification using chinese twins." *Demography*, *45*(1), 223–243.

Li, H., Zhang, J., Zhu, Y., et al. (2005). "The effect of the one-child policy on fertility in china: Identification based on the differences-in-differences." *Department of Economics, The Chinese University of Hong Kong, Discussion Papers*, *19*.

Liu, H. (2014). "The quality–quantity trade-off: evidence from the relaxation of china's one-child policy." *Journal of Population Economics*, *27*(2), 565–602.

Loh, C., and Remick, E. J. (2015). "China's skewed sex ratio and the one-child policy." *The China Quarterly*, *222*, 295–319.

Loh, E. S. (1996). "Productivity differences and the marriage wage premium for white males." *Journal of Human Resources*, 566–589.

Løken, K. V., Mogstad, M., and Wiswall, M. (2012). "What linear estimators miss: The effects of family income on child outcomes." *American Economic Journal: Applied Economics*, *4*(2), 1–35.

Marbach, M., and Hangartner, D. (2020). "Profiling compliers and noncompliers for instrumental-variable analysis." *Political Analysis*, *28*(3), 435–444.

Millimet, D. L., and Roy, J. (2016). "Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous." *Journal of Applied Econometrics*, *31*(4), 652–677.

Millimet, D. L., and Wang, L. (2006). "A distributional analysis of the gender earnings gap in urban china." *The BE Journal of Economic Analysis & Policy*, *5*(1).

Neuman, S., and Ziderman, A. (1992). "Benefits of women's education within marriage: results for israel in a dual labor market context." *Economic Development and Cultural*

Change, *40*(2), 413–424.

Neumark, D. (1999). "Biases in twin estimates of the return to schooling." *Economics of Education Review*, *18*(2), 143–148.

Nie, H., and Xing, C. (2019). "Education expansion, assortative marriage, and income inequality in china." *China Economic Review*, *55*, 37–51.

Ning, M., Gong, J., Zheng, X., and Zhuang, J. (2016). "Does new rural pension scheme decrease elderly labor supply? evidence from charls." *China Economic Review*, *41*, 315–330.

Nizalova, O. Y., Sliusarenko, T., and Shpak, S. (2016). "The motherhood wage penalty in times of transition." *Journal of Comparative Economics*, *44*(1), 56–75.

Oliveira, J. (2016). "The value of children: Inter-generational support, fertility, and human capital." *Journal of Development Economics*, *120*, 1–16.

O'Neill, E., and Weeks, M. (2018). "Causal tree estimation of heterogeneous household response to time-of-use electricity pricing schemes." *arXiv preprint arXiv:1810.09179*.

Putz, T., and Engelhardt, H. (2014). "The effects of the first birth timing on women's wages: A longitudinal analysis based on the german socio-economic panel." *Zeitschrift für Familienforschung*, *26*(3), 302–330.

Qian, N. (2009). "Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural china. nber working paper no. 14973." *National Bureau of Economic Research*.

Qian, X., and Smyth, R. (2008). "Measuring regional inequality of education in china: widening coast–inland gap or widening rural–urban gap?" *Journal of International Development: The Journal of the Development Studies Association*, *20*(2), 132–144.

Rosenzweig, M. R., and Wolpin, K. I. (1980). "Testing the quantity-quality fertility model: The use of twins as a natural experiment." *Econometrica: journal of the Econometric Society*, 227–240.

Rosenzweig, M. R., and Zhang, J. (2009). "Do population control policies induce more human capital investment? twins, birth weight and china's "one-child" policy." *The Review of Economic Studies*, *76*(3), 1149–1174.

Ruhm, C. J. (2008). "Maternal employment and adolescent development." *Labour economics*, *15*(5), 958–983.

Shen, Z., Zheng, X., and Yang, H. (2020). "The fertility effects of public pension: Evidence from the new rural pension scheme in china." *Plos one*, *15*(6), e0234657.

Short, S. E., and Fengying, Z. (1998). "Looking locally at china's one-child policy." *Studies in family planning*, 373–387.

Song, Y. (2014). "Losing an only child: the one-child policy and elderly care in china." *Reproductive Health Matters*, *22*(43), 113–124.

Umberson, D., and Gove, W. R. (1989). "Parenthood and psychological well-being: Theory, measurement, and stage in the family life course." *Journal of family issues*, *10*(4), 440–462.

Wang, L. (2013). "How does education affect the earnings distribution in urban china?" *Oxford Bulletin of Economics and Statistics*, *75*(3), 435–454.

Wang, L., Béland, D., and Zhang, S. (2014). "Pension fairness in china." *China Economic Review*, *28*, 25–36.

Welch, F. (1970). "Education in production." *Journal of political economy*, *78*(1), 35–59.

Welch, F. (1974). "Benefits of women's education within marriage: Comment." *Journal of Political Economy*, *82*(2, Part 2), S72–S75.

Wong, Y.-C. (1986). "Entrepreneurship, marriage, and earnings." *The Review of Economics and Statistics*, 693–699.

Wu, X., and Li, L. (2012). "Family size and maternal health: evidence from the one-child policy in china." *Journal of Population Economics*, *25*(4), 1341–1364.

Wu, X., and Li, L. (2014). "The motives of intergenerational transfer to the elderly parents in china: consequences of high medical expenditure." *Health economics*, *23*(6), 631–652.

Wu, X., and Zhang, Z. (2010). "Changes in educational inequality in china, 1990–2005: Evidence from the population census data." In *Globalization, changing demographics, and* *educational challenges in East Asia*, Emerald Group Publishing Limited.

Xie, Y., and Zhu, H. (2009). "Do sons or daughters give more money to parents in urban china?" *Journal of Marriage and Family*, *71*(1), 174–186.

Yang, H. (1996). "The distributive norm of monetary support to older parents: A look at a township in china." *Journal of Marriage and the Family*, 404–415.

Zeng, S., Li, F., and Ding, P. (2020). "Is being an only child harmful to psychological health?: evidence from an instrumental variable analysis of china's one-child policy." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *183*(4), 1615–1635.

Zeng, Y., and Hesketh, T. (2016). "The effects of china's universal two-child policy." *The Lancet*, *388*(10054), 1930–1938.

Zhang, J. (2017). "The evolution of china's one-child policy and its effects on family outcomes." *Journal of Economic Perspectives*, *31*(1), 141–60.

Zhang, J., Zhao, Y., Park, A., and Song, X. (2005). "Economic returns to schooling in urban china, 1988 to 2001." *Journal of comparative economics*, *33*(4), 730–752.

Zhao, Y. (1997). "Labor migration and returns to rural education in china." *American Journal of Agricultural Economics*, *79*(4), 1278–1287.

Zhao, Y., Strauss, J., Yang, G., Giles, J., Hu, P., Hu, Y., Lei, X., Park, A., Smith, J. P., and Wang, Y. (2013). "China health and retirement longitudinal study–2011–2012 national baseline users' guide." *Beijing: National School of Development, Peking University*, 1–56.

Zhong, H. (2014). "The effect of sibling size on children's health: a regression discontinuity design approach based on china's one-child policy." *China Economic Review*, *31*, 156–165.

Zhu, H., and Walker, A. (2018). "Pension system reform in china: Who gets what pensions?" *Social Policy & Administration*, *52*(7), 1410–1424.

Zimmer, Z., and Kwong, J. (2003). "Family size and support of older adults in urban and rural china: Current effects and future implications." *Demography*, *40*(1), 23–44.