Educational Dividend or Educational Disparity? The Consequences of Fertility Decline During the One-Child Policy era in China

Hanzhi Hu

Department of Sociology

University of Oxford

(Hanzhi.hu@sociology.ox.ac.uk)

Abstract

Rapid fertility decline is witnessed in developing countries during the second half of the twentieth century. However, little is known about the consequences of fertility decline on average education and educational inequality at the macro-level over time. Using data from China General Social Survey(CGSS) and China Family Panel Survey(CFPS) (N=44,918), this paper contributes to the literature by answering two questions regarding the educational consequences of fertility decline after the implementation of the One-Child Policy (OCP) in China with simulations. First, has fertility decline improved human capital by limiting average sibship size? Results show that the fertility decline during the OCP era brought modest education dividends in terms of improvement in the average year of schooling. Second, how does the differential fertility between groups contribute to educational inequality? Counterfactual simulations show that its impact on the educational disparity between males and females is limited. However, it has a marked impact on the rural-urban disparity in education.

(155 words)

Keywords: Fertility Decline; Sibship Size Effect; Educational Inequality; One-Child Policy; China

Introduction

Rapid fertility decline is witnessed in developing countries during the second half of the twentieth century (Bongaarts 2008). This demographic process is argued to affect economic outcomes, socioeconomic inequality, and the population's well-being. The consequences of the fertility decline have long been the interest to scholars. Fertility decline is believed to have an impact on economic growth (Cuaresma, Lutz, and Sanderson 2014), poverty reduction (Wietzke 2020), and inequality (Yount et al. 2014). Scholars argue that fertility decline can boost productivity and benefit economic growth by creating the demographic dividend (Bloom and Williamson 1998), which derives from the beneficial changes in age-dependency ratios. Following this argument, the "East Asian Growth Miracle" is attributed to countries' fertility decline and the subsequent changes in age structure (Bloom and Williamson 1998; Bloom, Canning and Sevilla 2003). Studies also found that fertility decline raises income per capita (Ashraf et al., 2013) and positively affects the growth of GDP (Li and Zhang 2007). Studies on the relationship between fertility rates and poverty emphasize the impact of the heterogeneous trends of fertility decline. Scholars argued that compared with the higher-status group, fertility decline happens later and slower for the lower-status group, which can increase inequality and influence poverty rates (Eastwood and Lipton 1999). Moreover, fertility decline can influence gender inequality and intensify gender gaps in societies with son preference (Yount, Zureick-Brown, Halim and LaVilla. 2014). For instance, as fertility rates fell in India, girls' excess mortality grew (Das Gupta and Bhat 1997). In China, the fertility decline is accompanied by an increasing male-to-female sex ratio at birth (Junhong 2001; Li, Yi and Zhang 2011).

When discussing the impact of fertility decline on economic growth, poverty reduction, and inequality, improvement in human capital is an often mentioned mechanism (Lee, Mason, and

Miller 2000; Wietzke 2020). For instance, Cuaresma, Lutz and Sanderson (2014) argued that the demographic dividend of fertility decline is an education dividend. The improvements in educational attainment explain a substantial portion of productivity and income growth brought by changes in age structure after the fertility decline. Fertility decline can increase educational attainment since, according to the "Quantity-Quality" trade-off thesis coined by Becker and Lewis (1973), at the household level, parents can offer each child more resources that benefit educational attainment as the number of children in the household decreases. Empirical studies have discussed the variations of the sibship size effect on education over cohorts (Parish and Willis 1993; Pong 1997; Maralani 2008;) and across countries (Vogl 2016; Li, Dow and Rosero-Bixby 2017). Scholars also investigated the mechanism of the effect (Chen, 2020) and its robustness against unmeasured confounders (Black et al., 2005; Angrist, Lavy and Schlosser, 2010). Though several lines of studies in recent years question the robustness of the negative relationship between sibship size and educational attainment (Maralani, 2008; Angrist et al. 2010; Åslund and Grönqvist, 2010; Bagger et al., 2013), numerous empirical studies have reported evidence supporting the "Quantity-Quality" trade-off thesis (Rosenzweig and Zhang, 2009; Ponczek and Souza, 2012; Liu, 2014; Choi et al., 2020).

If the impact of sibship size on education holds, we should expect fertility decline to bear consequences on macro-level trends in educational attainment. As fertility declines, the benefits of small sibship size at the family level might lead to improvements in the human capital of the society. Li and colleagues (2017) provided an exception. They showed with simulations that fertility decline in Latin America and Asia can explain only a small portion of the increase in educational attainment. Their study does not include China in the analysis. China provides a unique case to understand the consequences of fertility decline. It experienced a rapid fertility decline since the 1970s; however, its fertility decline did not happen entirely voluntarily like in other countries. The fertility trends in China are affected by multiple governments' birth planning suggestions, campaigns, and regulation policies, among which the One-child Policy launched in 1979 forcefully set limits to people's fertility behaviours. Whether the rapid fertility decline observed in China has translated into an improvement in human capital has not been empirically answered. The first goal of this paper is to show whether the fertility decline observed during the OCP era in China improved human capital?

Furthermore, the heterogeneity in fertility between groups should influence between-group inequality in educational attainment. Differential fertility between groups might translate into educational disparities as groups with higher fertility rates suffer from larger sibship size penalties. Moreover, during demographic transitions, fertility rates can decline differently across groups. For instance, the "leader-follower model" suggests that higher socioeconomic groups start to have fewer births earlier and faster than the lower-status groups (Bongaarts 2003; Wietzke, 2020). The divergence or convergence in fertility rates between groups over time can lead to long-term changes in between-group disparities in education. As in China, the implementation of the One-child Policy set different birth limits across groups, which might forcefully create differences in fertility between groups and lead to education inequality. First, previous studies suggested the decrease in sibship size for females should promote gender equality, as more women enjoyed the only-child premium and obtained similar attainment as men (Ye & Wu 2010; Wu, Ye & He 2012). Moreover, during the One-child Policy era, the relaxation of the regulations for the rural population in 1984 created differences in policy-targeted fertility between urban and rural areas. Only the urban population was strictly forced to limit their fertility, which might enlarge the fertility differences between rural and urban China. However, whether the fertility differences between groups have further translated into between-group inequality in education has not been empirically tested.

Using recent data sources, I observe the completed education for cohorts born from 1950 to 1993 and examine the consequence of fertility decline after the implementation of the One-Child Policy (OCP). Results show that the fertility decline during the OCP era brought modest education dividends in terms of improvement in the average year of schooling. Counterfactual simulations show that its impact on the educational disparity between genders is limited. However, during the OCP era, the divergent trends of fertility decline for urban and rural populations have a marked impact on the rural-urban disparity in education.

In the literature review section, I will review previous studies on fertility decline and its implications and focus on the relationship between fertility and educational attainment. Then, I will review the trends in fertility and educational attainment in China from 1950 to 1990 and discuss why and how we should expect differential fertility to contribute to gender inequality and rural-urban inequality. Then I introduce the data and empirical strategy and report the results. I end by discussing the implications of the findings for understanding human capital growth and inequality in Chinese society.

Literature Review

As a feature of demographic transitions, fertility decline and its implications have long been the interest to scholars. Though the causal relationships are still debatable, fertility decline is believed to be at least associated with economic growth (Cuaresma, Lutz, and Sanderson 2014), poverty reduction (Wietzke 2020), and inequality (Yount et al. 2014).

Scholars argue that fertility decline can boost productivity and benefit economic growth by creating the demographic dividend (Bloom and Williamson 1998), which derives from the beneficial changes in age-dependency ratios. Following this argument, the "East Asian Growth Miracle" is attributed to countries' fertility decline and the subsequent changes in age structure

(Bloom and Williamson 1998; Bloom, Canning and Sevilla 2003). Studies also found that fertility decline raises income per capita (Ashraf et al., 2013) and positively affects the growth of GDP (Li and Zhang 2007). Improvement in human capital is an often mentioned mechanism that links fertility decline and economic growth (Lee, Mason, and Miller 2000; Wietzke 2020). For instance, Cuaresma, Lutz and Sanderson (2014) argued that the demographic dividend of fertility decline is an education dividend. The improvements in educational attainment explain a substantial portion of productivity and income growth brought by age structure changes after the fertility decline.

Fertility decline can increase educational attainment since, according to the "Quantity-Quality" trade-off thesis coined by Becker and Lewis (1973), at the household level, parents can offer each child more resources that benefit educational attainment as the number of children in the household decreases. Empirical studies have discussed the variations of the sibship size effect on education over cohorts (Parish and Willis 1993; Pong 1997; Maralani 2008;) and across countries (Vogl 2016; Li, Dow and Rosero-Bixby 2017; Choi et al., 2020). Several studies record a shift in the association from positive to negative over cohorts. Maralani (2008) found that the association between sibling size and children's schooling in urban Indonesia was positive for older cohorts (the 1950s) but turned negative for more recent cohorts (1970s). Vogl (2016) also found that, in 48 developing countries, before 1960, children from larger families were more likely to attain higher education, but the pattern had reversed in later cohorts. A similar pattern is also found in Taiwan (Parish and Willis 1993) and Malaysia (Pong 1997). This change in the association has been attributed to the evolution of societal preferences and structural changes in other social regimes. Vogl (2016) emphasized that the reversal of sibling size and education association is due to the reversal of differential fertility by parents' education. For earlier cohorts, wealthier and more educated families tend to have more children; however, the association flips with economic development.

Scholars also investigated the mechanism of the effect (Chen, 2020) and its robustness against unmeasured confounders (Black et al., 2005; Angrist et al., 2010). Though numerous empirical studies have reported evidence supporting the "Quantity-Quality" trade-off thesis (Rosenzweig and Zhang, 2009; Ponczek and Souza, 2012; Liu, 2014; Choi et al., 2020), several lines of studies in recent years question the robustness of the negative relationship between sibship size and educational attainment (Maralani, 2008; Angrist et al., 2010; Åslund and Grönqvist, 2010; Bagger et al., 2013). Scholars have argued that confounders might bias the negative association between sibship size and education. Previous studies have usually used two approaches to overcome the endogeneity. First is to employ instrumental variables, like twins' birth, sex composition of children, or policy intervention, that triggers exogenous changes in sibling size (Black, Devereux and Salvanes, 2005; Rosenzweig & Zhang, 2009; Angrist et al. 2010). The other is to use the individual fixed-effect models to account for time-invariant confounders (Guo and VanWey 1999; Chen, 2020). Extant studies have reported mixed findings depending on the methods and educational outcome. For instance, Black et al. (2005) show that while applying the birth of twins as an instrumental variable finds no effect of increased family size on educational outcomes, using the same-sex composition as the instrument shows a positive effect in Norway. Cáceres-Delpiano (2006) reports no significant impact of the sibling size increase from twin births on grade retention, but it does reduce the probability of attending a private school in the U.S. Using the fixed-effect method, Guo and VanWey (1999) find that sibship size has little effect on verbal skills but positively affects math skills.

Compared with the abundant studies on the effect of sibship size on education, little is known about whether fertility decline can improve educational attainment at the macro-level by decreasing average sibship size. Li and colleagues (2017) showed with simulations that fertility decline in Latin America and Asia can explain only a small portion of the increase in educational attainment. Their study does not include China in the analysis. However, like other developing countries, China experienced a rapid fertility decline since the 1970s due to socioeconomic changes and policy interventions. Whether the fertility decline promoted average educational attainment at the macro-level is the first research question this paper aims to answer. The first hypothesis is proposed as,

H1: Fertility decline during the OCP era increases average educational attainment.

Besides the overall decline in fertility rates, the heterogeneity in fertility decline is argued to bear important consequences. Studies argue that the heterogeneous trends of fertility decline between groups are the driving forces of the relationship between fertility rates and inequality. Scholars argued that compared with the higher-status group, fertility decline happens later and slower for the lower-status group, which can increase inequality and influence poverty rates (Eastwood and Lipton 1999). The divergence in fertility rates between the higher- and lower-status group tends to worsen inequality and poverty rates, while the convergence in fertility rates reduces inequality. Not only by socioeconomic status, in the literature, differential fertility by various parents' characteristics, such as income (Lam, 1986; De La Croix, D., & Doepke, M., 2003), education (Westoff, 1954; Vogl, 2016), and race (Preston, 1974; Kuo and Hauser, 1995; Mare, 1997), is well documented. Demographers and stratification scholars are interested in examining the consequence of differential fertility between groups. Take the gradient of fertility across education levels as an example. If parents with less education tend to have more children, and large sibship size penalizes educational attainment, scholars argue that the differential fertility by parental education can lead to 1) lower average education over generations; 2) increase between-group inequality in education. Empirical studies provided mixed evidence. Preston and Campbell (1993) claimed that micro-level differential fertility by IQ does not translate into changes in the distribution of IQ in the population over generations with a two-sex model. Mare (1997) examined the differential fertility between white and

African-American women in the United States and found that the effect of differential fertility on the trend of average educational attainment is not large enough to create meaningful changes over generations, and its impact is offset mainly by intergenerational mobility. However, Volg (2016) pointed out that, among the 48 developing countries examined, results varied by countries for whether the differential fertility over the skill and income distribution has increased or depressed the mean educational attainment. As for between-group inequality, Mare (1997) showed that the racial differences in fertility between white and African-American women have little impact on educational inequality between races. As in China, as will be reviewed later, fertility rates differ across various characteristics, like regions, education levels, and socioeconomic status. Moreover, the regulations of the One-child Policy set different birth limits across groups. For instance, while the urban population is under the strict one-child limit, the rural population is subjected to more relaxed regulations, which might enlarge the fertility differences between rural and urban China and lead to more inequality in education. The second hypothesis is proposed as,

H2: The fertility differences during the OCP era lead to more inequality in education between rural and urban China.

Moreover, fertility decline can influence gender inequality as scholars asserted that fertility decline is associated with changes in gender attitudes and norms in society (Yount et al., 2014). However, empirical studies show that the relationship between fertility decline and gender inequality depends on the social context. Girls can benefit from fertility declines in societies with lower son preference, while in societies with strong son preference, studies show that decline in fertility intensifies the gender gaps in mortality and nutrition. For instance, as fertility rates fell in India, girls' excess mortality grew (Das Gupta and Bhat 1997). In China, the fertility decline is accompanied by an increasing male-to-female sex ratio at birth (Junhong 2001; Li, Yi and Zhang 2011). In terms of educational attainment, previous studies have speculated that the fertility decline in China should promote gender equality as more women enjoyed the only-child premium and obtained similar attainment as men (Ye & Wu 2010; Wu, Ye & He 2012). The third hypothesis is proposed as,

H3: The fertility differences during the OCP era reduce inequality in education between genders.

The Chinese Context

Trends in Fertility in China From 1950 to 1990

The fertility level in China has experienced drastic changes since 1950. The total fertility rates (TFR) in China throughout the 1950s and 1960s were around 5 to 6 births per mother (Banister 1987). In the early 1970s, the Chinese government promoted a "Two is Enough" policy. The policy advocated that couples should delay marriage and childbearing and have no more than two children with wider birth spacing. It is also known as the "*wanxishao*" (later, longer and fewer) campaign. During the 1970s, China witnessed its most repaid fertility decline. The TFR plummeted from 5.8 in 1970 to 2.7 in 1979, more than half in 10 years (Cai 2008). The "two-child policy" and a subsequent series of fertility control policies culminated in the One-Child Policy in 1979. The initial statement of the policy is to encourage one child per couple. However, the implementation of the policy became stricter over time. In 1983, it developed into forced sterilization and abortion campaign. The campaign backfired on itself as the rebellion of the society led to severe unsettlement threatening security. In response, the central government relaxed the fertility control in 1984 and transferred the power to provincial governments to set fertility limits. So under the guidance of the provincial governments, lower-level governments execute the fertility control according to the local conditions. (See Gu et al. (2007) for a detailed summary of the local fertility policy variations at the provincial level.)

The fertility policies of the 31 provinces of mainland China fall into the following categories. For the urban population, one child limit was enforced. For the rural population, there were three scenarios:1) In six provinces, including Beijing and Shanghai, all residents follow the one-child rule; 2) in 19 provinces, rural residents are allowed to have a second child if the first birth is a girl. It is also called a "1.5-child policy"; 3) All rural couples can have two children in the remaining five provinces. Aside from the above, 26 provinces permitted couples who are both the only child in their families to have two children. Moreover, ethnic minorities are exempted from fertility control. (See Gu et al. 2007: p134-p135, Table 1, for a summary of provincial fertility policies in China revised or changed through the 1980s to late 1990s.)

During the 1980s, the observed TFR fluctuated around 2.5 (Feeney and Wang 1993). As entering the 1990s, the TFR dropped further. The 1990 census reported a TFR at 2.3 (Cai 2008). Analysis based on surveys and census suggests that the TFR in China has declined to the under replacement level in the 1990s, from 1.8 in 1991 to 1.2 in 2000 (Retherford et al. 2005). Though issues like underreporting have led to a debate on the exact level of TFR in China (Hermalin and Liu 1990; Goodkind 2004), the consensus is that during the 1990s, fertility rates continued to decline and fell below the replacement level.

Although population policies are argued to play an essential role in shaping the fertility trends in China (Lavely and Freedman 1990; Goodkind 2017), scholars emphasized that socioeconomic factors became key driving forces, especially for more recent trends (Zhao and Zhang 2018). Earlier work pointed out that well before the implementation of government programs in the 1960s and 1970s, evidence of voluntary family birth control and fertility declines were observed among the better-educated population and residents in urban areas (Greenhalgh 1989; Lavel and Freedman 1990). More recent work stressed that social, economic, and cultural transformations became the main factors leading to fertility decline since the mid-1990s (Zhao 2015). After several decades of high-speed development, people's fertility intentions and behaviours experienced profound changes. For example, studies recorded an increase in average age at first marriage and first birth (Feng and Quanhe 1996), contributing to the fertility decline. Moreover, the rising costs of child-rearing and education, along with other factors, have led to people's voluntary choices of having fewer children (Zhang 1990). It is challenging to distinguish policy intervention's impact from socioeconomic factors' impact on fertility decline, as these forces are interwoven and often work hand in hand.

While much attention has been paid to why and how the overall TRF changes over time in China, it is less discussed whether social groups experienced the decline differently. However, there are good reasons to expect heterogeneity in fertility decline. Policy interventions and individual characteristics, such as education, social status, and preference, all can create divergent fertility trends within the population. First, regarding individual characteristics, similar to other countries, fertility decline in China started from the educated groups. Lavely and Freeman (1990) mentioned that better-educated elites intentionally practised birth control before government intervention. Norm or preference is another important factor to consider due to the tradition of son preference in China (Murphy, Tao and Lu 2011). Parents with son preference might have multiple births to reach their desired result if no policy or resource constraints; or practice sex-selective abortion (Goodkind 2011; Den Boer and Hudson 2017) if having multiple births is not an option. Second, as for policy regulations, as mentioned, the strict 1979 One-Child Policy and its ancillaries set varied birth limits for social groups. While most urban residents were subjected to a strict one-child limit, rural residents and ethnic minorities can have multiple births with or without conditions. Moreover, studies pointed out the heterogeneity in people's degree of acceptance and compliance to the policies, as urban and industrialized areas show higher acceptance than others (Merli and Smith 2002). Thus, we should expect the fertility decline started earlier and more rapidly for groups with more education, socioeconomic status, weaker son preference, and living in urban areas.

Trends in Educational Attainment and Inequality in China Since 1950

It is well documented in the literature that educational attainment in China has increased over time since 1950. Cohort studies show that the average year of schooling increased for both men and women in rural and urban China (Wu and Zhang 2010: Fig. 2; J.Treiman 2013: Fig. 1).

The gender gap in education closed over cohorts (Wu and Zhang 2010), while the rural-urban gap persisted. The average year of schooling for males increased from around seven years for the 1950 cohort to around ten years for the 1980 cohort, while the average year of schooling for females increased from less than six years for the 1950 cohort to around ten years for 1980 cohorts. The level of educational attainment for men and women converged over time (Treiman 2013). Meanwhile, the urban-rural gap persisted. The average schooling for urban males born in 1950 is around 10, but the number for rural males is slightly above six. The gap was maintained over cohorts. When it comes to the 1980 cohort, the average year of schooling for urban males is 12 years, while it is around eight years for rural males. A similar pattern is observed for females that the urban-rural gap in education was maintained at around four years over cohorts (Wu and Zhang 2010).

When explaining the improvement in educational attainment, fertility decline is an often mentioned driving force alongside educational expansion and economic development. However, empirical tests on the effect of fertility decline on human capital improvement are sparse. Moreover, scholars suggested that the fertility decline contributed to the convergence of educational attainment between genders. Scholars argue that fertility decline should have promoted gender equality as more girls enjoyed the only-child premium and obtained similar attainment as boys (Ye & Wu 2010; Wu, Ye & He 2012). Some empirical analysis showed that in urban one-child families, only daughters and only sons were being treated equally (Tsui and Rich 2002), and no difference was found in their educational achievement (Lee 2012). However, these findings cannot be generalized to the rural case where only-child is not as common as in urban areas. More importantly, whether the increase in the proportion of only-child girls is large enough to affect the overall trends in gender inequality is under doubt. Hannum and Xie (1994) found little evidence to link the fertility decline with the declines in gender stratification.

As for the urban-rural disparity, the household registration system (*hukou*)[1], which institutionalize the divide between rural and urban population, is regarded as the most critical determinant of rural-urban inequality (Wu and Treiman 2004). Fertility differential is a less mentioned factor but should have played a role in the process. The fertility rate is higher in rural than in urban China, and the local implementations of the One-Child Policy might enlarge the difference. Rural parents tend to have more children than urban parents due to the practical consideration of having more workforce in the household for agricultural production and cultural norms that large family size represents prosperity in traditional Chinese culture. Moreover, during the One-child Policy era, the relaxation of the regulations for the rural population in 1984 created differences in policy-targeted fertility between urban and rural areas. Only the urban population was strictly forced to limit their fertility, which might enlarge the fertility differences between rural and urban China. Whether fertility differences affect

[1] *Hukou* refers to the household registration system in China. Under the household registration system, everyone is assigned a *hukou* (household registration certificate) at the time of birth. The *hukou* contains two folds of information. First, the category that one's *hukou* status falls in. The *hukou* status is divided into rural vs urban or agricultural vs non-agricultural. Second, the local administrative unit one's *hukou* is registered. The *hukou* must register to one and only one administrative unit at the lowest available level as one's official/permanent residential place and under the supervision of all the administrative authorities level above.

between-group inequality in China has not been empirically tested. To fill in the void, I compile several recent data sources to observe the completed education for cohorts born from 1950 to 1993 and examine the contribution of changes in sibling size during the One-Child Policy era on 1) average education and 2) between-group inequality over cohorts.

Data and Methods

Data and Measurements

I use China General Social Survey(CGSS) 2006, 2008, 2017, and China Family Panel Survey(CFPS) 2010-2018 to construct a cross-sectional dataset. CGSS is a nationally representative cross-sectional survey on mainland China launched in 2003 (Bian and Li 2012). Respondents' number of siblings is known only in the 2006, 2008, and 2017 waves. The CFPS is a nationally representative longitudinal dataset. The survey sample is drawn from 25 provinces or administrative equivalents, representing 94.5% of the total population in mainland China. The sampling is designed to achieve regional representation using multistage probability sampling with implicit stratification (Xie &Lu 2015). They are the most recent data sources available to observe the completed education for cohorts born after the implementation of the One-child Policy.

I focus on respondents born after the People's Republic of China (PRC) was founded (1949) and older than 25-years old for each survey year. The final sample contains 44,918 respondents born between 1950 and 1993 with complete information on their sibship size and educational attainment. Education attainment is measured as the year of education the respondent completed. Sibship size is measured as the total number of siblings, including those who had passed away at the survey time. I distinguish urban from rural population according to the *hukou* registration status of the mother when the respondent is 14 years old. Table A1 in the Appendix presents the sample size, average sibship size, and schooling for each cohort.

Methods

To portrait the trend over cohorts, I calculate the 5-year moving average of sibship size and educational attainment. The average for the focal cohort, t , is computed as,

$$\widehat{y_t} = \frac{y_{t-2}+y_{t-1}+y_t+y_{t+1}+y_{t+2}}{5} \quad (1.1)$$

For each cohort, I predict education attainment with sibship size, gender, *hukou* status, ethnicity, parental birth years and parents' education, then explore how fertility decline has contributed to educational attainment and inequality with counterfactual simulations. I first fixed sibship size at its pre-OCP level to see the difference between the hypothetical trend in educational attainment and the observed. Then, I look into the trends by gender and by regions (urban v.s. rural). I explore the effect of fertility on gender inequality and urban/rural inequality by making alternative assumptions about how the sibship size changes over time. One is to fix the sibship size at each group's pre-OCP level; The other is to let males and females, urban and rural residents have the same level of sibship size. The associations between sibship size and educational attainment is obtained with OLS regression. The potential endogeneity concern is discussed in the Robustness Check section.

Results

Sibship Size and Educational Attainment: Trends, Group Variations, and Inequality

Fig.1 depicts the changes in average sibship size and year of education over cohorts. Overall, the average sibship size decreased while the education attainment increased. As shown in panel a of Fig.1, the average sibship size decreased from around 3.6 for the oldest cohorts (the 1950s) to around 1.05 for the youngest cohorts (the 1990s). The average sibship size declined more rapidly for cohorts born before the OCP. For the 1970-1978 cohorts, the average sibship size decreased by 1.18, from 2.85 to 1.67, while for cohorts born after the implementation of the OCP (1979-1993 cohorts), their average sibship size decreased by 0.65, from 1.67 to 1.02. This echoes with previous findings that family size started to decline in China before the implementation of OCP, and it has been declined most rapidly during the 1970s (Hesketh, Lu & Xing, 2005). In terms of education, as shown in panel b of Fig.1, the average year of schooling increased over cohorts, from around 7.3 years for the 1950s to 13.1 years for the 1990s, except for fluctuations experienced by the 1960s cohorts.

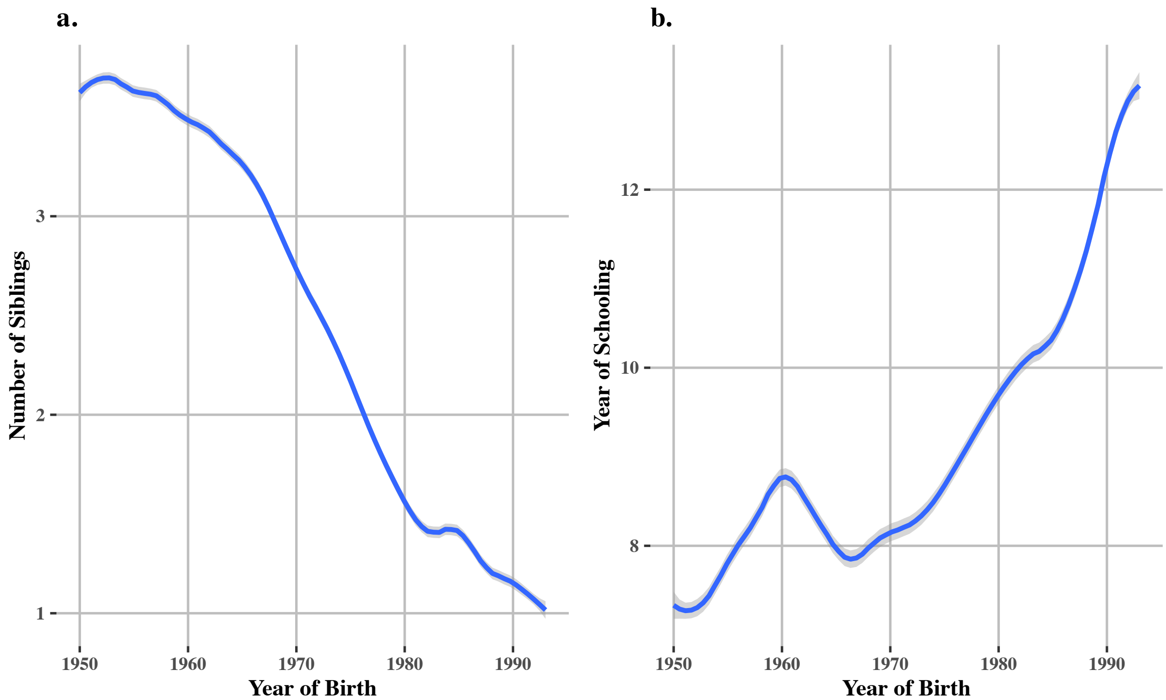

Figure.1 Cohort Trends of the Average Number of Sibling (Panel a) and Average Year of Schooling (Panel b)

As mentioned, due to the prevalence of son preference, we would expect females to have more siblings than males. We would also expect urban residents to have fewer siblings than rural residents since, on average, urban parents are more educated and with higher social status. Moreover, in rural areas, families are allowed to have multiple births conditioned or unconditioned on having a son. Thus, we would expect the group discrepancies by gender (male vs female) and regions (urban vs rural) in sibship size to be extravagated in the OCP era.

First, for gender differences, according to Panel.a in Fig.2, females had more siblings than males over cohorts. Gender difference in sibship size was maintained at around 0.3 during the observed period. Meanwhile, the gender difference in average education decreased over cohorts (Panel.b in Fig.2). The largest gender gap is observed for the 1960 cohort, where males had 1.6 years more education than females. The gender gap diminished afterwards and closed for the 1983 cohort and onwards. There is no statistically significant difference in educational attainment between genders for the OCP generations.

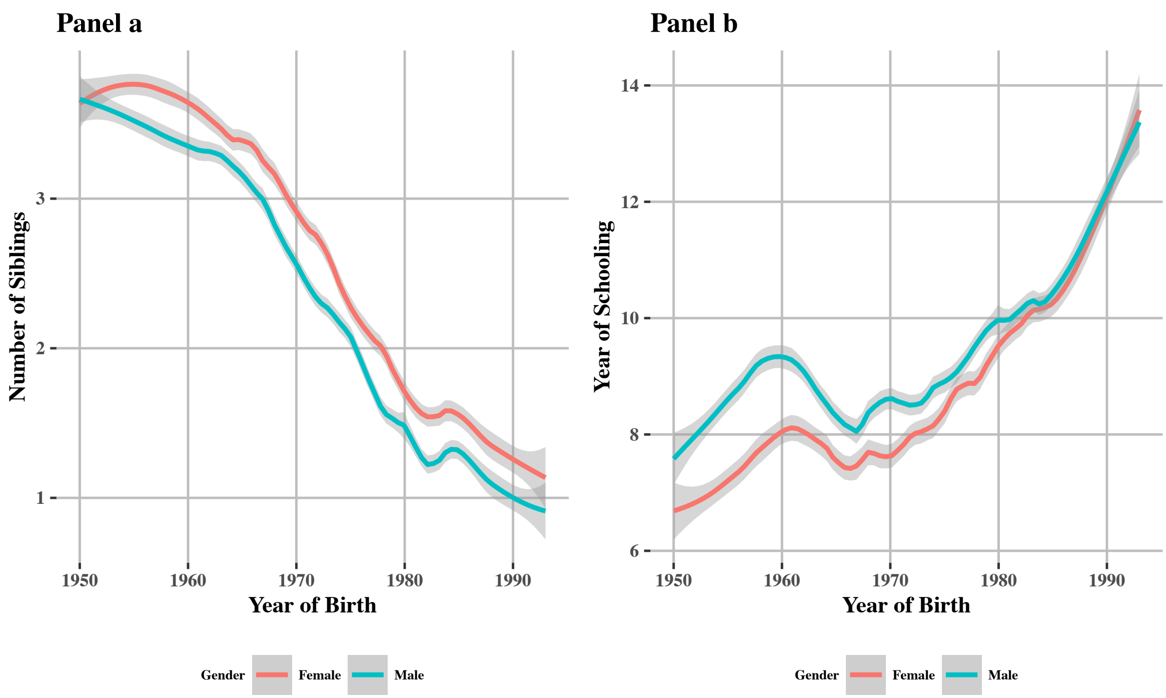

Figure.2 Cohort Trends of Sibling Size (Panel a) and Average Year of Schooling (Panel b) by Gender

As for the rural/urban discrepancy, rural residents had more siblings than urban residents, and the difference first decreased from around one for the 1960 cohort to 0.67 for the 1970 cohort (Panel.a in Fig.3). It then fluctuated at around 0.7 for the 1970s cohorts and increased after the implementation of the OCP. It oscillated at around one for the younger cohorts. What is worth noticing is that the average sibship size for urban residents fell below one after 1979, which suggests a high proportion of the only-child among urban residents. Fig.4 depicts the share of only-child within urban and rural residents. We can see that for cohorts born before 1970, the percentage of the only-child is less than 5% in both urban and rural populations. Since 1970, the share of the only-child among urban residents has increased dramatically. The number increased from less than 5% for the 1970 cohort to 40% for the 1979 cohort. After the implementation of OCP, the share of the only-child in urban increased by around 38%, from 40% to 78% for the 1993 cohort. In comparison, the increase in the share of only-child among rural residents was relatively modest. It slowly increased from less than 5% for the 1970 cohort to less than 10% for the 1980 cohort and around 20% for the 1993 cohort. As for educational attainment (Panel.b in Fig.3), over cohorts, though urban and rural residents all experienced improvement in average educational attainment, urban residents constantly had around four more years of education than rural residents, and this gap only started to diminish for recent cohorts (the 1990s).

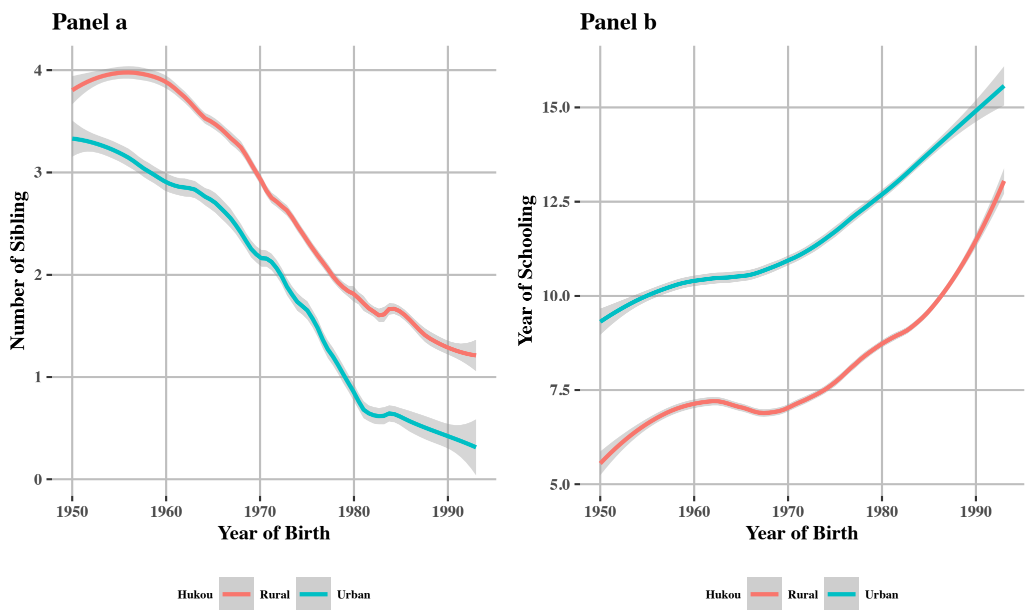

Figure.3 Cohort Trends of Sibling Size (Panel a) and Average Year of Schooling (Panel b) by *Hukou*

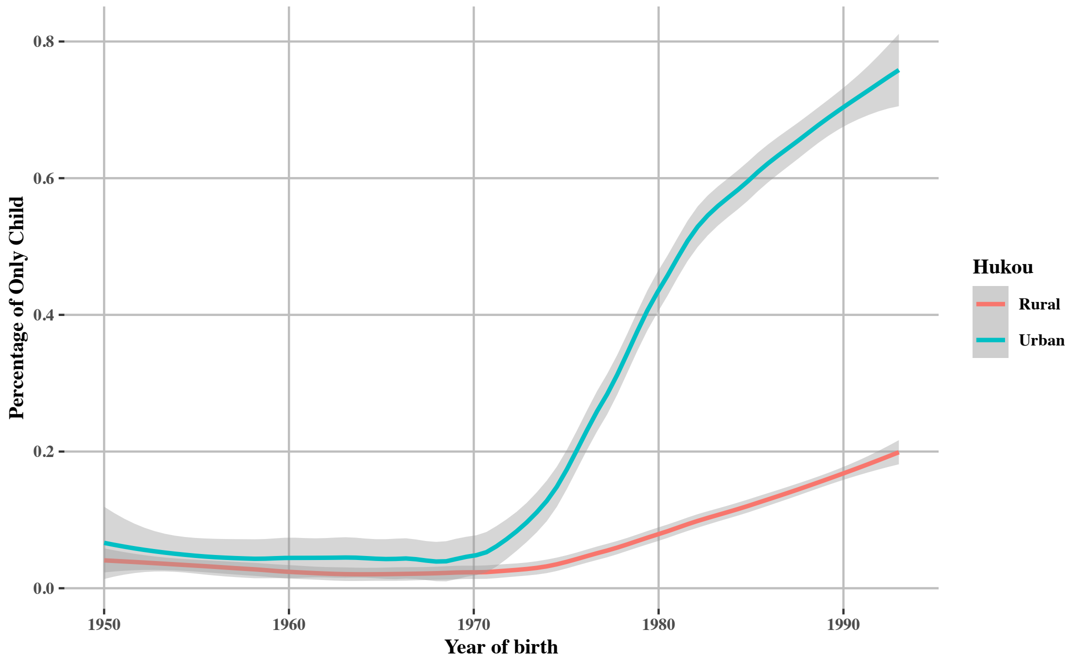

Figure 4. Cohort Trends of Share of Only Child in the Respondents by *Hukou*

To link the changes in sibship size with changes in educational attainment, we need to answer two questions: first, whether sibship size has a causal influence on one's educational attainment; second, how much of the gender inequality and urban/rural inequality in the year of schooling is explained by the differential sibship sizes. For the former, I first estimate the association between one's sibship size and educational attainment over cohorts with OLS regressions and then employ the exogenous changes in sibship size due to the One-Child Policy as instruments to address the endogeneity problem. For the second question, I report two sets of counterfactual analysis to see a) what would have happened if the sibship size had remained at pre-OCP level; b) what would have happened if males and females, rural and urban residents had had the same sibship size over cohorts.

Sibship Size and Educational Attainment

Fig.5. shows how the association between sibship size and educational attainment evolves over cohorts. The estimates for each cohort is obtained with OLS regression controlling for child's gender, *hukou* status, ethnicity, parental birth years and parents' education. In line with

previous work, sibship size is negatively associated with one's educational attainment (Li et al. 2008; Rosenzweig and Zhang 2009; Liu 2014; Chen, 2020), and the negative association became stronger over time (Lu and Treiman 2008; Choi et al., 2020). Having one more sibling is associated with around 0.1-unit less education (statistically insignificant, p>0.01) for the oldest cohorts (the 1950s). The association increased to around 0.6-unit less of education (p<0.000) for the youngest cohorts (the 1990s). For the early cohorts (the 1950s and 1960s) and some of the recent cohorts (the 1980s), the sibship size disadvantage is larger for females than for males (Panel.b in Fig.5). There is no statistically significant difference between urban and rural residents in the sibship size penalty (Panel.c in Fig.5).

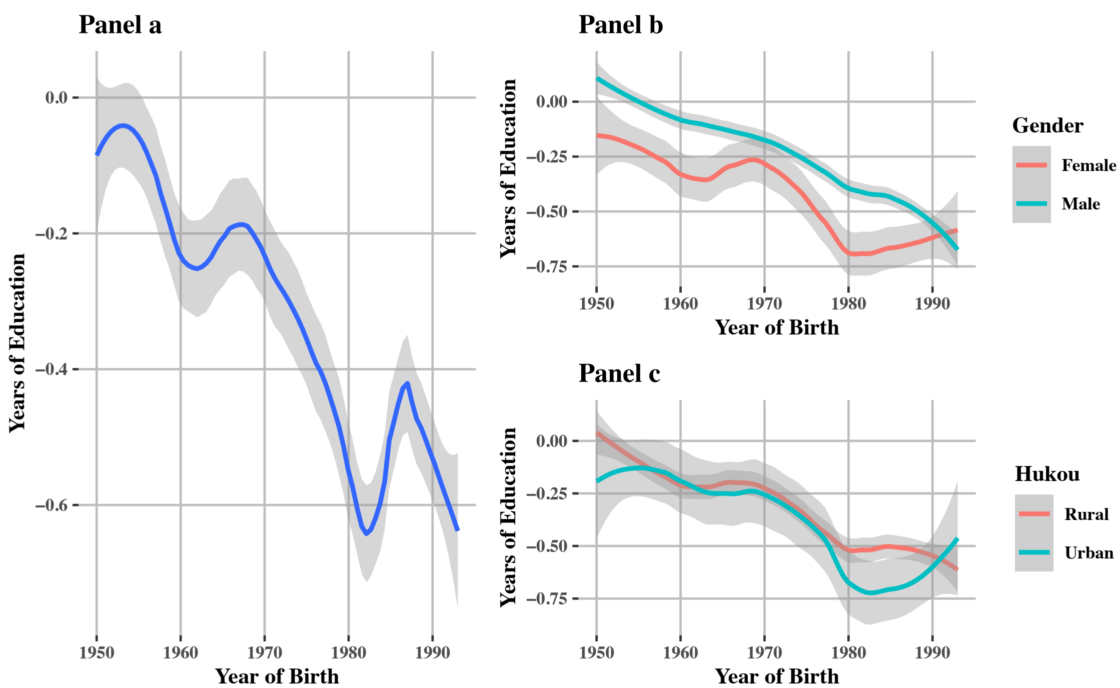

Figure.5 Cohort Trends in Sibship Size Disadvantage in the Years of Education (Panel a) and Trends by Gender (Panel b) and by *Hukou* Status (Panel c).
Note: Shadows show 95% confidence intervals of the coefficient. Parental education, ethnicity, gender, hukou, age and age square are controlled.

Differential Fertility and Group Inequality

First, I keep the average sibship size constant at the 1978 level, one year before the official implementation of the One-Child Policy, for cohorts born after the OCP. As shown in Fig.6, for the OCP generations (1979-1993 cohorts), their observed educational attainment increased by 3.84 years from 9.28 to 13.12. If the average sibship size had remained at its 1978 level, the increase in educational attainment for the OCP generations would have been 3.29 years, from 9.28 to 12.56, which is a 0.6-year or 9%-less than the observed.

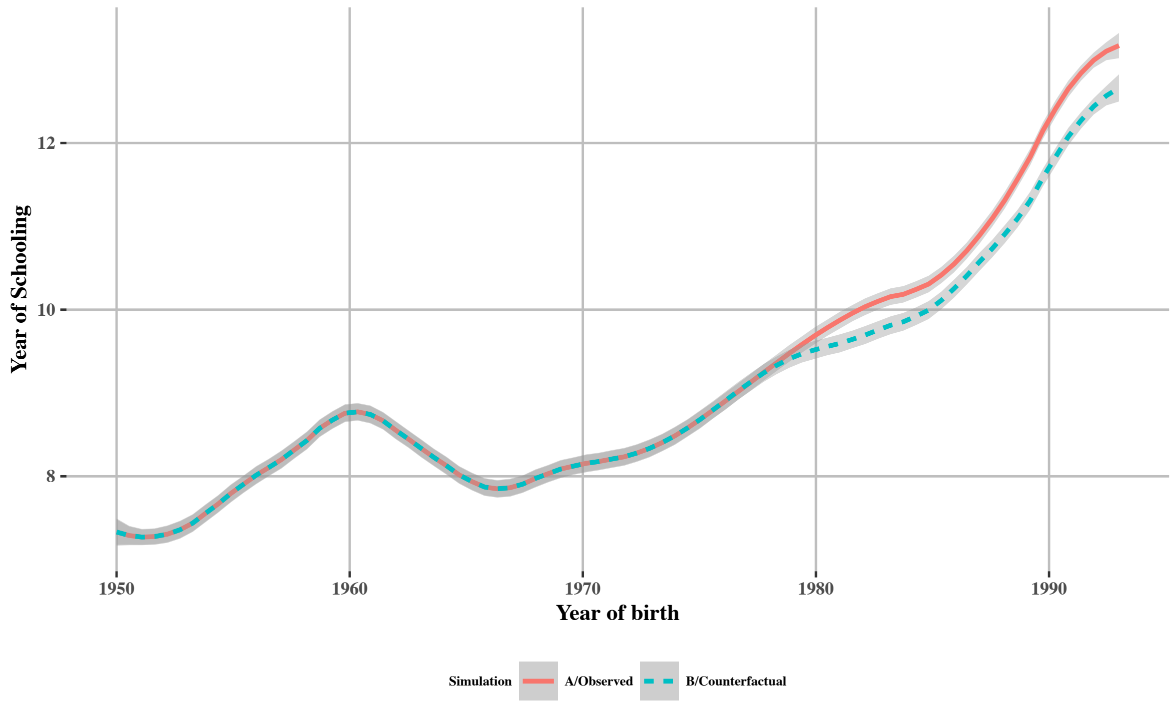

Figure.6 Observed and Counterfactual Trends (Sibling Size Fixed at 1978 Level) of Average Year of Schooling. Dash line represents the counterfactual scenario. Shadows show 95% confidence intervals.

As for the gender gap in education, it started to diminish before the OCP was implemented and closed for cohorts from 1983 onwards. If the sibship size had remained constant for males and females at the 1978 level, the hypothetical trends (Panel.a in Fig.7) suggest that it would have slightly suppressed the improvement in education for both genders and postponed the achievement of gender equality in education. The gender gap would not have closed until the 1990 cohort under the counterfactual scenario. Panel b in Fig.7 shows what if females had had the same level of sibship size as males. The hypothetical trends are almost the same as the observed. The gender gap in education would have closed as early as for the 1980 cohort, and females would have surpassed males in educational attainment for the 1990 cohorts. In all, the fertility decline during the OCP era has moderately accelerated the closing of the gender gap in education. However, the gender gap in education would have disappeared over cohorts even there had been no difference between genders in their sibship sizes.

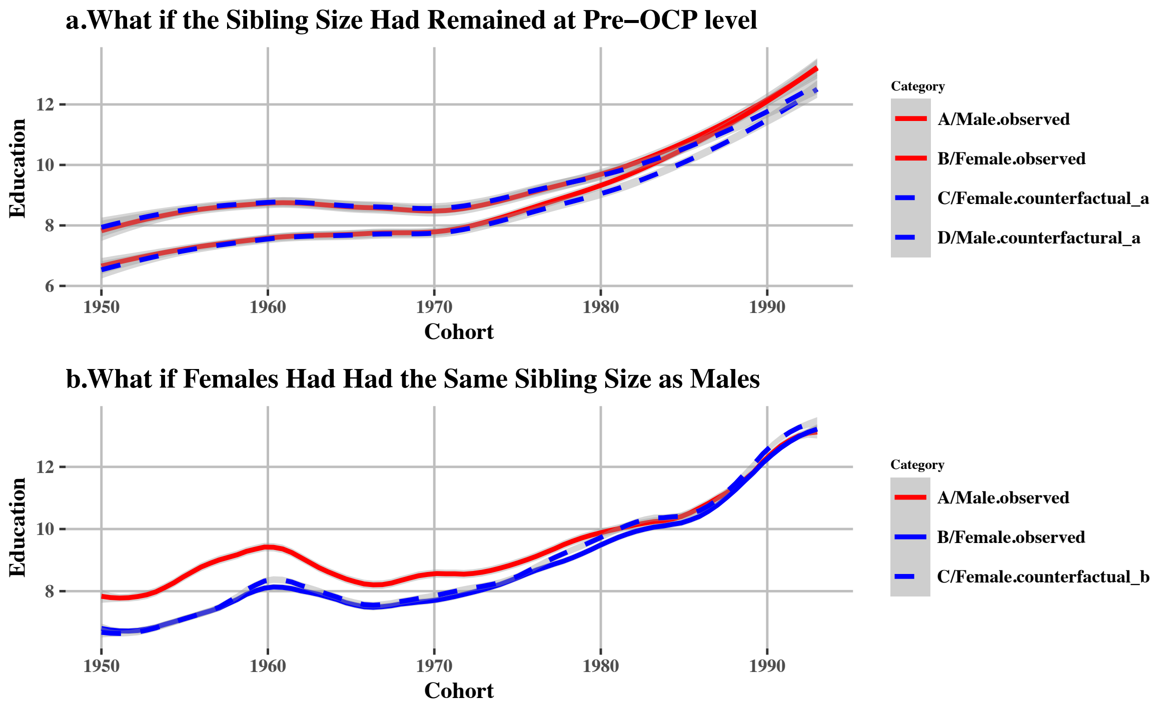

Figure.7 Observed and Two Counterfactual Trends of Gender Inequality in Schooling. Dash lines represent the counterfactual scenarios. Shadows show 95% confidence intervals.

As for rural/urban differences, panel.a in Fig 8 shows what would have happened if the sibship size had remained at the pre-OCP level. As shown, the rural/urban differences in education would have been significantly smaller if the sibship size for the two groups had remained at the 1978 level. First, the counterfactual sibship size has little impact on the rural residents' educational attainment. The development under the counterfactual is essentially the same as the observed. For rural residents, their observed average attainment for the 1990 cohort is 11.53 years and would have been 11.32 years under the counterfactual. However, urban residents' educational attainment would have increased less rapidly under the counterfactual. The observed attainment for the 1990 cohort urban residents is 15.10 years. If they had had the sibship size at the pre-OCP level, the average attainment would have been 13.13 years, a near two-year less of improvement. The observed urban/rural gap for the 1990 cohort is 3.57 years, and it would have been 1.81 years under the counterfactual, which is 49.3% less than the observed. Panel b in Fig.8 shows that if rural residents had had the same sibship size over cohorts as their urban counterparts, it would have increased the rural residents' educational attainment by 0.7 years for the 1970 cohort (7.13 vs 7.83), 0.98 years for the 1980 cohort (8.74 vs 9.72), and 1.08 years for the 1990 cohort (11.53 vs 12.61). The rural/urban gap would have decreased by 18.4% for the 1970 cohort, 24.9% for the 1980 cohort, and 30.3% for the 1990 cohort, respectively.

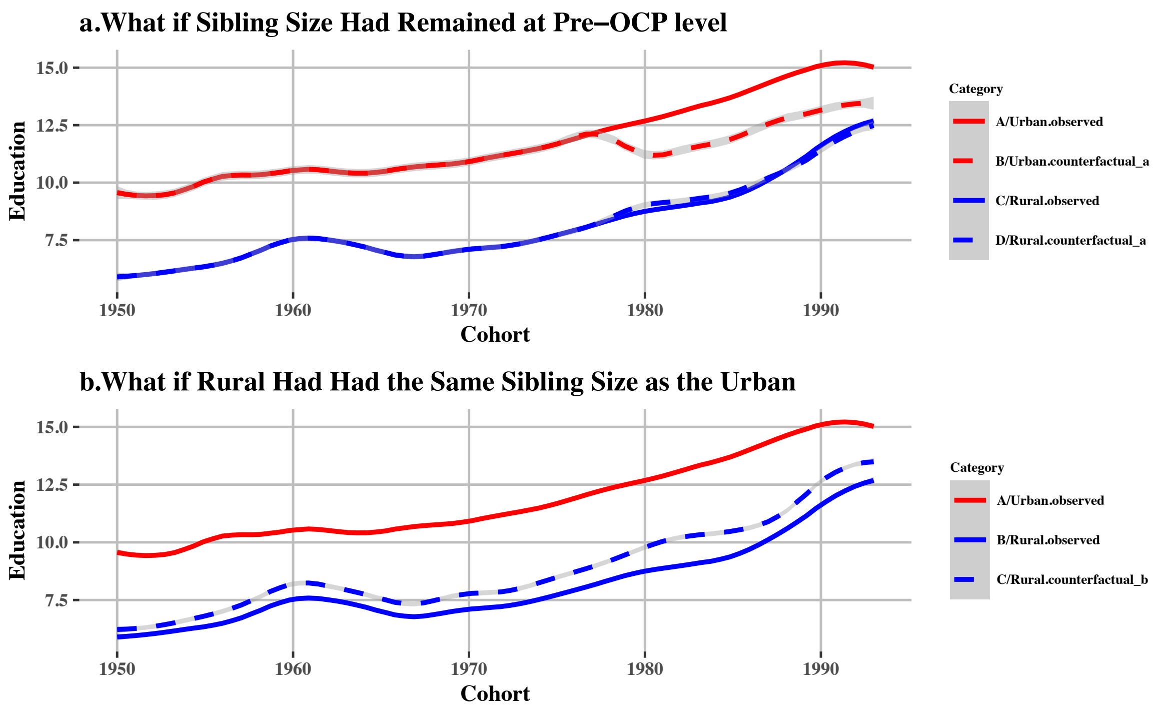

Figure.8 Observed and Two Counterfactual Trends of Urban/Rural Inequality in Schooling. Dash lines represent the counterfactual scenarios. Shadows show 95% confidence intervals.

Robustness Check

Internal Migration

When examining the cohort trends, I define rural and urban populations according to the mother's *hukou* status when the respondent is young. However, China has experienced a massive scale of internal rural-to-urban migration since the early 1980s. In 2007, about 10% of China's population were rural-to-urban migrants (Kwong 2011), while in 2011, 29.7% of the population with a rural origin lived in the urban areas. If only the policy intervention is considered, not isolating the internal migrants from the rural sample should not be a concern because one's fertility behaviour is subjected to local regulation where the person's *hukou* is registered. For a migrant who lives in the urban areas, if his/her mother has the rural *hukou* when the respondent is young, his/her sibship size is affected by the rural rather than the urban regulations. However, the assimilation theory of migration (Warner and Srole 1945) expects migrants and their children to adapt to the hosting environment and to have similar attitudes and behaviours as residents at their destinations. Thus, rural-to-urban migrants may voluntarily adopt the fertility strategies of urban residents to have fewer children. Moreover, migration experience also influences one's educational attainment. Empirical studies report mixed findings on the impact of migration on one's educational attainment (Liang & Chen 2007; Xu & Xie 2015; Wu & Zhang 2015). The concern is that if migrants are positively or negatively selected and intrinsically different in childbearing and rearing from their rural counterparts, including them in the rural sample might bias the results. According to the literature, the most advantaged migrants are those who successfully switch to urban *hukou* (Wu and Treiman 2004), as the government set quota on *hukou* conversion to 1.5 to 2.0 per thousand persons each year (Yilong 2003). The conversion can be achieved mainly by obtaining high education, party membership, and military services. Among the 32,815 rural cases in the sample, 3,033(9.24%) cases switched to urban *hukou* at the time of the survey. To test the robustness of the results against selection, I redo the analysis by excluding them from the sample. Results do not show a substantial difference as reported.

The Endogeneity Problem of Sibship Size

The simulation model is constructed based on the assumption that changes in sibship size are exogenous. However, as reviewed, previous studies raised concerns about the endogeneity problem of sibship size, and estimates from OLS regression might be biased. In response, I use the instrumental variable approach and compare the 2SLS estimates, and the OLS estimates to see whether and how the endogeneity problem can affect the conclusions.

In the literature, studies often employ twin births or the sex composition of the first two births as instruments for sibship size (Black, Devereux and Salvanes, 2005; Conley and Glauber 2006; Angrist et al., 2010). However, these instrumentals have drawbacks when applied to the Chinese context. Neither of them can capture the increase in sibship size from zero to one, which is the most common scenario in China during the observational periods. In practice, the variation in family size due to these instruments is confined to families that have at least two children.[2] Moreover, the sex composition itself is endogenous in a society with a strong son preference. Instead, I use the variations in the One-Child Policy in rural China as instruments. Several previous studies employed this policy intervention as an instrument but suffered two shortcomings. First, the policy regulation is often measured at the aggregated level, for instance, the urban and rural differences (Wu & Li, 2012) or provincial or county-level differences (Liu, 2014; Qian, 2009). Second, previous studies measured outcomes for children who reside in the same household as their parents often at relatively young ages (Liu, 2014). In this paper, I used the detailed policy regulations available at the community/village level: a) whether more than

[2] When using twin births as instrument, the effects of sibship size are uncovered by estimating the effect of a twin birth at birth N, on the outcomes of children born prior to this birth, conditional on families having at least N births. Due to the fact that twins and singleton child are different in their birth-endowment, like birth weight, and various other aspects that related to future development, it is not appropriate to use births of twin as instrument at N = 1, i.e., to compare twins with singleton child at first birth, to estimate the effect of increase in sibship size from 0 to 1. Similarly, the sex composition can only be used for families with at least two children to estimate the effect of family size increase from 2 to more.

one birth is allowed; 2) the amount of fine on the excess births, from 2010 China Family Panel Studies (CFPS) as instruments to assess sibship size's effect on the highest completed education. Details on the sample and variables used for the IV analysis are in the Appendix.

The instrument exogeneity assumption requires that the local fertility regulations should not be correlated with unmeasured confounders that affect a child's education. The most mentioned candidate of the unmeasured confounders in the literature is the parental preference for the high-quality child, which is also associated with a smaller sibship size. The local governments set the fertility policy. It can only be correlated with parental preference if a culture of "Quality-Quantity" trade-off preference exists at the community level. It is tested in previous studies whether the family size preferences varied at the community level. There is no evidence suggesting that communities with more relaxed fertility policies also have a higher proportion of parents preferring a larger family (Liu 2014). I further control for educational resources measured by educational expenditure and number of schools in the communities in 2010, in case of local governments with stricter fertility policy also invest more in education. Another potential violation of the exogeneity assumption is that local variations in policy may motivate migration for fertility purposes. Communities with fewer restrictions on fertility may have more migrants who wish to have more children, and if the migrants are intrinsically different in educating their children, the exogeneity assumption is violated. In response, I excluded the migrants from the analysis. I also checked the exogeneity assumption with a placebo test and the robustness of the results against potential measurement errors in the instrumental variables. Details are provided in the Appendix.

The OLS and IV results are presented in Table 1. All models controlled for the following variables: 1) the demographic features of the child: gender and age gap between the oldest and youngest children in the family; 2) the characteristics of the mother: her age at first birth and year of schooling; 3) father's year of schooling and the logarithm of annual net family income; 4) local educational resources measured by the educational expenditure of the government and number of schools in the communities in 2010.

Table 1. OLS and IV Estimates of the Effect of Sibling Size on Education Outcomes

	OLS	**IV**	
		1st Stage	**2SLS**
Outcome Variables		Sibling size	Education of the 1st born
Number of Sibling	**-.583***** (.107)		**-1.847*** (.851)
Instruments			
One Child Limit (No=0; Yes=1)		-.182** (.076)	
Log of Fine		-.291*** (.074)	
N	1,461		
adj. R^2	.157	.182	.072
1st Stage Statistics			
Partial R^2		.035	
Robust F-statistics		14.680	

Note: Four sets of variables: 1) the demographic features of the child: gender and age gap between the oldest and youngest children in the family; 2) the characteristics of the mother: her age at first birth and year of schooling; 3) father's year of schooling and the logarithm of annual net family income;4) annual educational expenditure and number of schools of the communities. are controlled. Parameters for control variables are omitted here. Standard errors in parentheses. Standard errors are adjusted to account for the clustering of residents within communities.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

According to the OLS estimates, a one-unit increase in sibship size is associated with a 0.583-year (p<.000), or 7-month, decrease in the educational attainment of the first child. As for the IV estimates, the first stage regression results show that the one-child fertility limit is negatively associated with sibship size. Living in communities where only one birth is allowed, *ceteris paribus*, is associated with a .185-unit (p<0.01) decrease in sibship size. As for the level of fine, it is negatively associated with sibship size. A 1% increase in the fine charged on excess birth is associated with a 0.3% decrease ($p<.000$) in sibship size. The two instruments together explained 3.5% of the variance in sibship size, and the F-statistics passed the threshold of 10 (Bound et al., 1995).

The IV models show the same patterns that the sibship size has a negative effect on the educational attainment of the firstborn children. Compared with the OLS estimate, the IV estimate has a sizeable standard error, and its 95% confidence interval contains the point estimate from the OLS regression. To decide which estimator is preferred, I tested the endogeneity of sibship size with the Hausman (1978) test. It compares the OLS estimator and the 2SLS estimator to see whether the difference is statistically significant. The test results give large p-values ($p=.091$), suggesting that there lacks sufficient evidence to reject the null hypothesis that the OLS estimate and the 2SLS estimate are statistically different. Thus, the OLS estimate is preferred for its efficiency.[3]

Conclusion and Discussion

This study answered two questions regarding the educational consequences of fertility decline using the case of China.

First, has fertility decline improved human capital via shrinking sibship size? The educational consequences of fertility decline or changes in sibship size depend on: a) the magnitude of fertility changes; b) the magnitude of the effect of sibship size on educational attainment. As For the former, the amount of sibship size changes during the OCP era is limited. After the official implementation of the One-Child Policy, China's average sibship size decreased by

[3] In the appendix, I discussed the validity of the instruments and provide evidence for the exclusion restriction assumption with placebo test and robustness check.

0.65 for the observed cohorts (1979-1993). Counterfactual analysis suggests that if the sibship size had remained at the pre-OCP level, the average educational attainment would have been 0.6-year less than the observed. During this period, the total increase in educational attainment is 3.84 years. The contribution of sibship size decline is relatively modest at 9%. Although during this period, the draconian policy intervention and rapidly developed social factors, like raising mothers' education and labour force participation rate, led to a further decline in fertility, the amount of decline observed in this period compared with the decline happened in the previous decade (1970-1979) is limited (1.18 vs 0.65). Thus, the contribution of fertility decline to human capital improvement during the OCP era is modest.

As for the second component, the magnitude of the effect of fertility on educational attainment, a methodological concern is whether the OLS estimates of the effect of fertility on educational attainment are biased. I employed two policy instruments from the local variations of the One-Child Policy in rural areas to address the endogeneity concern. IV results confirm a negative link between sibship size and educational attainment. The OLS estimate is preferred for its efficiency in the analysis since the difference between the OLS and IV estimates is not statistically significant. The decision is also made because IV analysis requires its own set of strong assumptions, and its estimate is a local effect for the compliers (LATE).

The second question answered is how does the differential fertility between groups contribute to educational inequality? First, results show that, over cohorts, females have more siblings than males, and the gap in sibship size between gender was maintained at 0.3. The observed gender gap in education closed for the 1983 cohort and onwards. Counterfactual analysis suggests that if females had had the same number of siblings as males, the gender gap in education would have disappeared slightly earlier for the 1980 cohort and onwards. The impact of sibship size on gender inequality in education is limited. Previous studies argue that fertility decline should have promoted gender equality as more girls enjoyed the only-child premium and obtained similar attainment as boys (Tsui and Rich 2002; Lee 2012). Results suggest that the magnitude of the impact is rather modest. Factors other than fertility drive the convergence of females' and males' educational attainment in China. Fig A1 in the appendix depicts the cohort trends of educational attainment by gender for those with and without siblings, respectively. As shown, there is no statistically significant difference in educational attainment between only-child females and only-child males over cohorts. For those with siblings, the gender gap closed over cohorts as well. It further illustrates that despite females having more siblings and the disadvantage of sibship size on education appears to be more vital for females than males, the sibship size does not play an important role in explaining the closed gender gap in education in education China.

Second, as for rural-urban discrepancy, results show that differential fertility has a marked impact on rural-urban inequality in education. Findings first highlight that fertility declined differently in rural and urban China. Over cohorts, rural residents had more siblings than urban residents, and the difference first decreased to 0.7-unit for the 1970s cohorts and increased to one after the implementation of the OCP. During the One-Child Policy era, urban residents' sibship size decreased more rapidly. It might be because only urban populations are subjected to the strict one-child limit. Fertility regulations are more relaxed for rural populations. The counterfactual analysis shows that if the average sibling size had remained at the pre-policy level, it would have suppressed the educational improvement for the urban residents but had little impact on rural residents. Thus, it would have diminished the urban-rural inequality in education. These findings highlight a less mentioned perspective: the differential fertility between groups in understanding educational inequality in contemporary China. By fixing the sibship size for urban and rural populations at the same level, findings from this analysis show that differential fertility can explain up to 30% of their education gap. It should inspire future research on the inequality of other dimensions where differences in fertility rates between groups are observed. For instance, a promising candidate is the inequality between ethnic groups (Han Chinese vs ethnic minorities). Ethnicity groups might cherish different childbearing and rearing values, which leads to discrepancies in fertility outcomes between groups. Moreover, ethnic minorities are exempted from the one-child limit during the OCP era, further enlarging the fertility gap between groups. As ethnic inequality in educational attainment (Hannum 2002) and occupation status (Hannum and Xie 1998) is documented, it is a fruitful avenue of research to look into the long-term impact of fertility differences on these ethnic inequalities.

One thing that needs to emphasize is that the findings of this study should not be interpreted as the impact of the one-child limit set by the One-Child Policy because, for one, the policy regulations varied across groups. Only urban populations and residents in several provinces are subjected to the strict one-child limit. For another, multiple factors have led to fertility decline during this period. The effect of policy intervention and socioeconomic factors on fertility decline are interwoven. Scholars of population policy might be interested in the contribution solely made by the One-Child Policy intervention. To answer this inquiry, one has to isolate the impact of policy on fertility decline from other socioeconomic factors. MacElroy and Yang (2000) provided an estimate of a third of a child in family size reduction due to policy intervention. R.Rosenzweig and Zhang (2009) used this result and computed the contribution of the One-Child Policy on human capital improvement, concluding that the policy at most increased schooling attainment by 4%. However, this result should be interpreted with caution since the estimate of the family size reduction due to policy intervention is based on local policy variations in rural areas. Thus, it should not be generalized nationally and directly compared with findings from this analysis. Evaluating the consequence of the One-Child Policy bears its own significance. Findings from this analysis suggest that future work should first focus on urban populations as they are subjected to the strict one-child limit; second, find approaches to isolate the impact of policy intervention from other socioeconomic factors.

Finally, based on the findings, the answer to the question raised in the title is, the fertility decline during the One-child Policy era in China brought modest education dividends in terms of improvement in the average year of schooling. Its impact on the educational disparity between genders is also limited. However, during this period, the divergent trends of fertility decline for urban and rural populations have a marked impact on the rural-urban disparity in education.

Reference

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

Ashraf, Q. H., Weil, D. N., & Wilde, J. (2013). The effect of fertility reduction on economic growth. *Population and development review*, *39*(1), 97-130.

Åslund, O. and Grönqvist, H., 2010. Family size and child outcomes: Is there really no trade-off?. *Labour Economics*, *17*(1), pp.130-139.

Bagger, J., Birchenall, J. A., Mansour, H., & Urzúa, S. (2013). Education, birth order, and family size (No. w19111). *National Bureau of Economic Research.*

Banister, J. (1987). *China's changing population*. Stanford University Press.

Becker, G. S., & Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of political Economy*, 81(2, Part 2), S279-S288.

Bian, Y., & Li, L. (2012). The Chinese general social survey (2003-8) sample designs and data evaluation. *Chinese Sociological Review*, *45*(1), 70-97.

Black, S. E., Devereux, P. J., & 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.

Bloom, D. E., & Williamson, J. G. (1998). Demographic transitions and economic miracles in emerging Asia. *The World Bank Economic Review*, *12*(3), 419-455.

Bloom, D., Canning, D., & Sevilla, J. (2003). *The demographic dividend: A new perspective on the economic consequences of population change*. Rand Corporation.

Bongaarts, J. (2003). Completing the fertility transition in the developing world: The role of educational differences and fertility preferences. *Population Studies*, *57*(3), 321-335.

Bongaarts, J. (2008). Fertility transitions in developing countries: Progress or stagnation?. *Studies in family planning*, *39*(2), 105-110.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430), 443-450.

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

Cai, Y. (2008). An assessment of China's fertility level using the variable-r method. *Demography*, *45*(2), 271-281.

Chen, S., (2020). Parental Investment After the Birth of a Sibling: The Effect of Family Size in Low-Fertility China. *Demography*, pp.1-27.

Choi, S., Taiji, R., Chen, M., & Monden, C. (2020). Cohort trends in the association between sibship size and educational attainment in 26 low-fertility countries. *Demography*, 57(3), 1035-1062.

Conley, D., & Glauber, R. (2006). Parental educational investment and children's academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility. *Journal of Human Resources*, 41(4), 722-737.

Cuaresma, J. C., Lutz, W., & Sanderson, W. (2014). Is the demographic dividend an education dividend?. *Demography*, *51*(1), 299-315.

Das Gupta, M., & Mari Bhat, P. N. (1997). Fertility decline and increased manifestation of sex bias in India. *Population Studies*, *51*(3), 307-315.

De La Croix, D., & Doepke, M. (2003). Inequality and growth: why differential fertility matters. *American Economic Review*, 93(4), 1091-1113.

Den Boer, A., & Hudson, V. (2017). Patrilineality, son preference, and sex selection in South Korea and Vietnam. *Population and Development Review*, 119-147.

Eastwood, R., & Lipton, M. (1999). The impact of changes in human fertility on poverty. *The Journal of Development Studies*, *36*(1), 1-30.

Feeney, G., & Feng, W. (1993). Parity progression and birth intervals in China: The influence of policy in hastening fertility decline. *Population and Development Review*, 61-101.

Feng, W., & Quanhe, Y. (1996). Age at marriage and the first birth interval: The emerging change in sexual behavior among young couples in China. *Population and Development Review*, 299-320.

Goodkind, D. M. (2004). China's missing children: the 2000 census underreporting surprise. *Population Studies*, *58*(3), 281-295.

Goodkind, D. (2011). Child underreporting, fertility, and sex ratio imbalance in China. *Demography*, *48*(1), 291-316.

Goodkind, D. (2017). The astonishing population averted by China's birth restrictions: Estimates, nightmares, and reprogrammed ambitions. *Demography*, *54*(4), 1375-1400.

Greenhalgh, S. (1989). *Fertility trends in China: approaching the 1990s* (Vol. 8). Population Council, Research Division.

Guo, G. and VanWey, L.K., (1999). Sibship size and intellectual development: Is the relationshipcausal?. *American Sociological Review*, pp.169-187.

Gu.B., Feng, W., Zhigang, G., & Erli, Z. (2007). China's local and national fertility policies at the end of the twentieth century. *Population and Development Review*, 129-147.

Hannum, E., & Xie, Y. (1994). Ineouality in China: 1949–1985. *Research in social stratification and mobility*, *13*, 73-98.

Hannum, E., & Xie, Y. (1998). Ethnic stratification in Northwest China: Occupational differences between Han Chinese and national minorities in Xinjiang, 1982–1990. *Demography*, *35*(3), 323-333.

Hannum, E. (2002). Educational stratification by ethnicity in China: Enrollment and attainment in the early reform years. *Demography*, *39*(1), 95-117.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.

Hesketh, T., Lu, L., & Xing, Z. W. (2005). The effect of China's one-child family policy after 25 years. *New Engl J Med*, *353*(11), 1171-1176.

Hermalin, A. I., & Liu, X. (1990). Gauging the validity of responses to questions on family size preferences in China. *Population and Development Review*, 337-354.

Kuo, H. H. D., & Hauser, R. M. (1995). Trends in family effects on the education of black and white brothers. *Sociology of Education*, 136-160

Kwong, J. (2011). Education and identity: The marginalisation of migrant youths in Beijing. *Journal of Youth Studies*, *14*(8), 871-883.

Lam, D. (1986). The dynamics of population growth, differential fertility, and inequality. *The American Economic Review*, 1103-1116.

Lavely, W., & Freedman, R. (1990). The origins of the Chinese fertility decline. *Demography*, *27*(3), 357-367.

Lee, R., Mason, A., & Miller, T. (2000). Life cycle saving and the demographic transition: The case of Taiwan. *Population and Development Review*, *26*, 194-219.

Lee, M. H. (2012). The one-child policy and gender equality in education in China: Evidence from household data. *Journal of family and economic issues*, *33*(1), 41-52.

Li, H., & Zhang, J. (2007). Do high birth rates hamper economic growth?. *The Review of Economics and Statistics*, *89*(1), 110-117.

Li, H., Yi, J., & 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, J., Dow, W. H., & Rosero-Bixby, L. (2017). Education gains attributable to fertility decline: Patterns by gender, period, and country in Latin America and Asia. *Demography*, *54*(4), 1353-1373.

Liang, Z., & Chen, Y. P. (2007). The educational consequences of migration for children in China. *Social Science Research*, 36(1), 28-47.

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.

Lu, Y., & Treiman, D. J. (2008). The effect of sibship size on educational attainment in China: Period variations. *American Sociological Review*, *73*(5), 813-834.

Junhong, C. (2001). Prenatal sex determination and sex-selective abortion in rural central China. *Population and Development Review*, *27*(2), 259-281.

Maralani, V. (2008). The changing relationship between family size and educational attainment over the course of socioeconomic development: Evidence from Indonesia. *Demography*, 45(3), 693-717.

Mare, R. D. (1997). Differential fertility, intergenerational educational mobility, and racial inequality. *Social Science Research*, 26(3), 263-291.

McElroy, M., & Yang, D. T. (2000). Carrots and sticks: fertility effects of China's population policies. *American Economic Review*, *90*(2), 389-392.

Merli, M. G., & Smith, H. I. (2002). Has the Chinese family planning policy been successful in changing fertility preferences?. *Demography*, *39*(3), 557-572.

Murphy, R., Tao, R., & Lu, X. (2011). Son preference in rural China: Patrilineal families and socioeconomic change. *Population and Development Review*, *37*(4), 665-690.

Parish, W. L., & Willis, R. J. (1993). Daughters, education, and family budgets Taiwan experiences. *Journal of Human Resources*, 863-898.

Ponczek, V. and Souza, A.P.(2012). New evidence of the causal effect of family size on child quality in a developing country. *Journal of Human Resources*, 47(1), pp.64-106.

Pong, S. L. (1997). Sibship size and educational attainment in Peninsular Malaysia: Do policies matter?. *Sociological Perspectives*, 40(2), 227-242.

Preston, S. H. (1974). Differential fertility, unwanted fertility, and racial trends in occupational achievement. *American Sociological Review*, 492-506.

Preston, S. H., & Campbell, C. (1993). Differential fertility and the distribution of traits: The case of IQ. *American Journal of Sociology*, 98(5), 997-1019.

Qian, N. (2009). Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural China (No. w14973). *National Bureau of Economic Research*.

Retherford, R. D., Choe, M. K., Chen, J., Xiru, L., & Hongyan, C. (2005). How far has fertility in China really declined?. *Population and Development Review*, *31*(1), 57-84.

Rosenzweig, M. R., & 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.

Treiman, D. J. (2013). Trends in educational attainment in China. *Chinese Sociological Review*, *45*(3), 3-25.

Tsui, M., & Rich, L. (2002). The only child and educational opportunity for girls in urban China. *Gender & Society*, *16*(1), 74-92.

Vogl, T. S. (2016). Differential fertility, human capital, and development. *The Review of Economic Studies,* 83(1), 365-401.

Warner, W. L., & Srole, L. (1945). The social systems of American ethnic groups.

Westoff, C. F. (1954). Differential fertility in the United States: 1900 to 1952. *American Sociological Review*, 19(5), 549-561.

Whyte, M. K., & Gu, S. Z. (1987). Popular response to China's fertility transition. *Population and Development Review*, 471-493.

Wietzke, F. B. (2020). Poverty, inequality, and fertility: the contribution of demographic change to global poverty reduction. *Population and Development Review*, *46*(1), 65-99.

Wu, X., Ye, H., & He, G. G. (2012). Fertility decline and women's empowerment in China. *International Center for Research on Women Fertility & Empowerment Working Paper Series*, 006-2012.

Wu, X., & 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., & Treiman, D. J. (2004). The household registration system and social stratification in China: 1955–1996. *Demography*, *41*(2), 363-384.

Wu, X., & 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.

Wu, X., & Zhang, Z. (2015). Population migration and children's school enrollments in China, 1990–2005. *Social Science Research*, *53*, 177-190.

Xie, Y., & Lu, P. (2015). The sampling design of the China family panel studies (CFPS). *Chinese Journal of Sociology*, *1*(4), 471-484.

Xu, H., & Xie, Y. (2015). The causal effects of rural-to-urban migration on children's well-being in China. *European Sociological Review*, *31*(4), 502-519.

Ye, H., & Wu, X. (2010). Fertility decline and educational gender inequality in China. *International Sociological Association-Research Committee on Social Stratification and Mobilit*y (ISA-RC28).

Yilong, Lu. "The Household Registration System-Control and Social Distinction (hu ji zhi du—kongzhiyu she hui cha bie)." (2003): 124-126.

Yount, K. M., Zureick-Brown, S., Halim, N., & LaVilla, K. (2014). Fertility decline, girls' well-being, and gender gaps in children's well-being in poor countries. *Demography*, *51*(2), 535-561.

Zhao, Z. (2015). Closing a sociodemographic chapter of Chinese history. *Population and Development Review*, *41*(4), 681-686.

Zhao, Z., & Zhang, G. (2018). Socioeconomic factors have been the major driving force of China's fertility changes since the mid-1990s. *Demography*, *55*(2), 733-742.

Zhang, J. (1990). Socioeconomic determinants of fertility in China. *Journal of Population Economics*, *3*(2), 105-123.

Appendix
Part I. Descriptive Statistics
Part II. Supplementary Results
Part III. Details of Casual Analysis with Instrumental Variables

Part I. Descriptive Statistics

Table A1. Descriptive Statistics for Cohorts

Cohort	Sample Size	Sibling Size	Average Schooling
1950	290	3.63	7.33
1951	263	3.64	7.30
1952	354	3.72	7.22
1953	367	3.69	7.47
1954	445	3.69	7.50
1955	437	3.58	7.89
1956	515	3.64	8.00
1957	570	3.61	8.22
1958	509	3.60	8.31
1959	435	3.47	8.63
1960	515	3.49	8.88
1961	509	3.48	8.72
1962	963	3.41	8.51
1963	1,111	3.36	8.34
1964	1,080	3.32	8.19
1965	1,105	3.26	7.98
1966	1,215	3.21	7.80
1967	1,146	3.07	7.87
1968	1,617	2.99	8.00
1969	1,455	2.85	8.10
1970	1,795	2.73	8.17
1971	1,704	2.62	8.15
1972	1,759	2.51	8.29
1973	1,785	2.42	8.31
1974	1,643	2.30	8.50
1965	1,574	2.17	8.67
1976	1,694	2.03	8.88
1977	1,579	1.90	9.09
1978	1,746	1.77	9.28
1979	1,703	1.67	9.53
1980	1,705	1.57	9.67
1981	1,778	1.47	9.89

1982	1,948	1.41	9.99
1983	1,687	1.40	10.18
1984	1,571	1.42	10.22
1985	1,715	1.44	10.33
1986	390	1.37	10.46
1987	463	1.22	11.06
1988	353	1.21	11.17
1989	373	1.18	11.73
1990	307	1.17	12.22
1991	294	1.10	12.92
1992	255	1.06	13.03
1993	205	1.02	13.12

Table A2. Descriptive Statistics by Groups

	Male		**Female**		**Urban**		**Rural**	
Cohort	Sibling Size	AYS	Sibling Size	AYS	Sibling Size	AYS	Sibling Size	AYS
1950	3.65	7.81	3.61	6.83	3.34	9.53	3.82	5.91
1951	3.63	7.87	3.66	6.68	3.26	9.52	3.87	5.96
1952	3.65	7.70	3.79	6.70	3.33	9.44	3.94	5.95
1953	3.59	8.03	3.80	6.88	3.28	9.48	3.94	6.25
1954	3.52	8.08	3.87	6.83	3.29	9.66	3.93	6.17
1955	3.48	8.48	3.68	7.27	3.12	10.10	3.89	6.41
1956	3.54	8.86	3.74	7.17	3.13	10.35	3.98	6.42
1957	3.52	8.99	3.68	7.54	3.08	10.32	3.97	6.76
1958	3.44	9.13	3.76	7.52	3.09	10.29	3.96	6.94
1959	3.30	9.20	3.65	8.06	2.95	10.34	3.87	7.35
1960	3.30	9.59	3.67	8.19	2.87	10.65	3.94	7.59
1961	3.32	9.36	3.64	8.08	2.88	10.57	3.85	7.58
1962	3.30	9.04	3.53	7.97	2.84	10.50	3.70	7.50
1963	3.30	8.74	3.43	7.92	2.83	10.46	3.61	7.35
1964	3.22	8.57	3.42	7.80	2.76	10.34	3.55	7.27
1965	3.15	8.36	3.37	7.57	2.70	10.52	3.47	7.00
1966	3.06	8.17	3.37	7.41	2.64	10.52	3.42	6.80
1967	2.94	8.20	3.21	7.52	2.50	10.81	3.29	6.74
1968	2.83	8.35	3.16	7.63	2.39	10.71	3.24	6.88
1969	2.68	8.52	3.03	7.64	2.28	10.79	3.07	7.02
1970	2.55	8.62	2.91	7.69	2.18	10.93	2.93	7.14
1971	2.45	8.51	2.82	7.77	2.13	11.05	2.80	7.12
1972	2.32	8.58	2.72	7.98	2.00	11.26	2.69	7.26
1973	2.26	8.56	2.60	8.05	1.89	11.29	2.60	7.32

Note: AYS means Average Year of Schooling.

1974	2.16	8.79	2.44	8.20	1.75	11.51	2.48	7.5
1975	2.07	8.91	2.28	8.43	1.63	11.66	2.34	7.7
1976	1.89	9.08	2.17	8.68	1.49	11.92	2.20	7.9
1977	1.75	9.32	2.07	8.85	1.31	12.13	2.09	8.1
1978	1.59	9.55	1.98	8.97	1.18	12.35	1.96	8.3
1979	1.53	9.78	1.84	9.23	0.99	12.52	1.88	8.6
1980	1.43	9.84	1.72	9.49	0.86	12.67	1.79	8.7
1981	1.35	10.02	1.60	9.76	0.70	12.86	1.73	8.9
1982	1.25	10.11	1.57	9.86	0.65	13.08	1.66	8.9
1983	1.26	10.26	1.55	10.08	0.62	13.33	1.66	9.1
1984	1.30	10.29	1.56	10.13	0.65	13.53	1.65	9.2
1985	1.33	10.41	1.56	10.23	0.62	13.72	1.66	9.4
1986	1.26	10.57	1.50	10.32	0.58	13.92	1.56	9.6
1987	1.07	11.25	1.38	10.86	0.40	14.51	1.41	10.
1988	1.07	11.16	1.34	11.18	0.46	14.52	1.37	10.
1989	1.07	11.72	1.28	11.74	0.52	14.86	1.32	11.
1990	1.04	12.22	1.30	12.22	0.47	15.10	1.34	11.
1991	0.96	13.03	1.24	12.80	0.43	15.28	1.28	12.
1992	0.93	13.03	1.21	13.03	0.35	15.20	1.25	12.
1993	0.91	13.06	1.14	13.19	0.33	15.00	1.20	12.

Part II. Supplementary Results

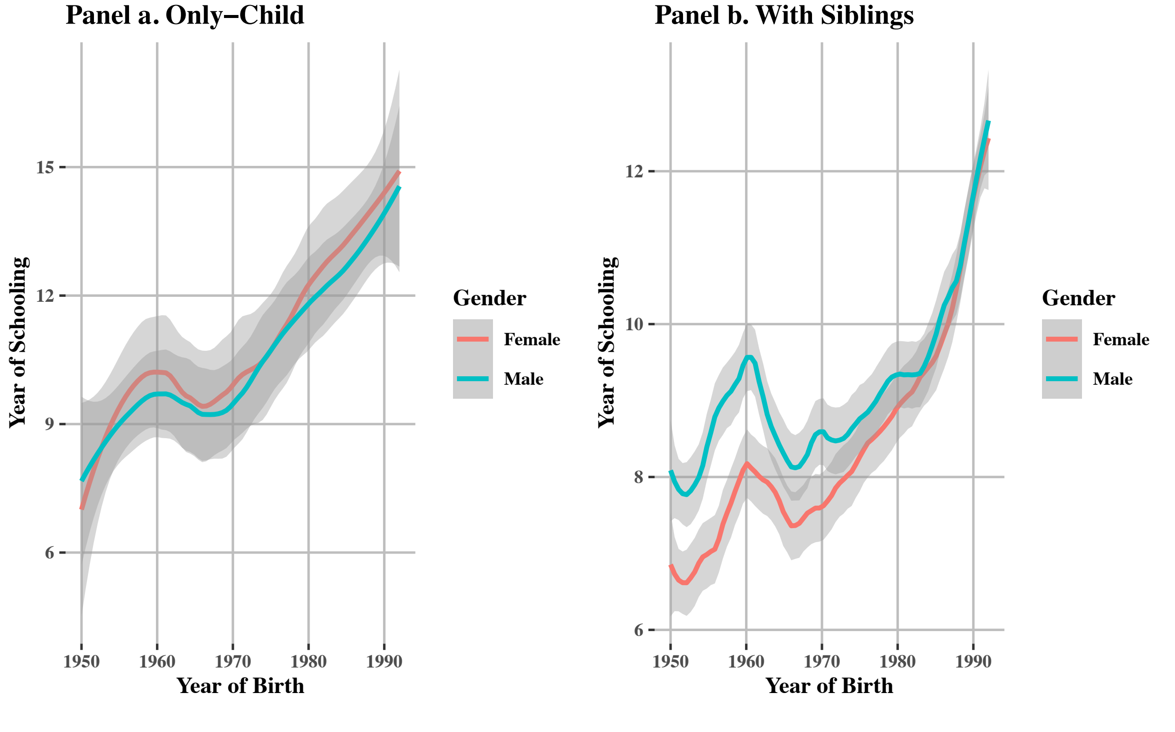

Fig.A1 Cohort Trends of Educational Attainment by Genders for Those Without Sibling (Panel a) and with Siblings (Panel b)

Part III. Details of Casual Analysis with Instrumental Variables

Data and Sample

To employ the local variations in fertility policy in rural areas as instrumental variables, I used the CFPS 2010 Data. It suits the research agenda for two reasons. First, information on the specific regulations of fertility policy is available at the community level. Compared with previous studies, it provides more refined policy variations. Second, it collected information on all children of the adult respondents, including those who did not reside in the same household. Thus, it is possible to examine the effect of sibling size on complete educational attainment at older ages.

Among those born after 1949 and older than 25 in 2010, a sub-sample is selected for the causal inference with the following criteria. 1) Mothers who are living in the same places as where their *hukou* is registered.; 2) Mothers who have the rural type of *hukou*. The first two criteria are set to exclude internal migrants. It responses to the concern mentioned earlier that people may migrant for fertility purpose to places where the local policy is more relaxed. Then, I restricted the sample to mothers whose 3) first child is older than 25 years old and 4) first birth is not twin birth. Next, following the practice in previous study (Liu, 2014), I further restricted the sample to 5) mothers who gave their first birth after 1976, four-year prior to the implementation of the "one-child" policy to make sure their fertility decisions were subjected to the policy restriction. The four-year interval is used as it was the recommended birth interval before the "one-child" policy.[4] 6) I exclude ethnic minority from the sample as they are exempted from the restriction. The final analytic sample contains 1,461 mothers, residing in 338 communities in 23 different provinces with complete information on the variables used in the analysis[5].

Measurements

Instrumental Variable

Two instrumental variables are employed as the exogenous variation in sibling size. They reflect the local fertility regulation. The information is collected from the officials in the community governments. The regulation is applied to all residences whose hukou is registered to the community.

[4] I tested the robustness of the results against the sample selection criteria. First, I selected mothers who were born after 1945, so they were younger than 35 in 1980 when the one-child policy is implemented. It is to make sure that mothers were at their prime fertility age at the time when one-child policy was implemented. Thus, their fertility behaviors were subjected to the variances of the policy. For the same purpose, I also conducted the analysis on a sample where mothers were younger than 30 years old in 1980 (N=2043). The results confirm the conclusions presented in the paper.

[5] Variables that suffer most from the missing values are the instrumental variables. 14.2% of missingness in IVa and 15.7% of missingness in IVb.

The first instrument (IVa) is a binary variable indicating whether the family is allowed to have more than one child. It is generated according to the local fertility regulation and the sex composition of the children in the family. Among communities that have valid information on fertility regulations, 10% has strict one-child policy. 57% allows for a second birth if there is no boy in the household. Another 33% allows for two or more births. Families living in the same community, they may subject to different regulations according to the sex composition of the children in the family. I generated a variable indicating how many births the family is allowed to have according to the local policy regulation and the sex composition of the children. In the sample, 46.11% of the cases is subjected to the one-child policy, 52.98% is allowed to have two births, and 0.91% is allowed to have three births. I further recoded it into a binary variable indicating whether one-child limit is rule (Yes=1; No= 0). It is expected to be negatively associated with the sibling size.

The second instrument (IVb) is a continuous variable measuring the minimum amount of fine charged for excess births. It ranges from 0 to 10,000 RMB. Two pieces of information can help understand the severity of punishment. First, the medium annual household income for families in the sample is 23,961 RMB. Second, I also computed the share of fine of the annual household income. It ranges from 0 to 20 times the annual household income with a mean of .83 and a standard deviation of 2.08.[6] It is expected to be negatively associated with the sibling size.

Other Variables

[6] The share of fine of family income is not used as the instrument for the following reasons. The household income is measured in 2009. It is less clear to what extent it can reflect the level of family resources when the child grown up. Furthermore, if the adult children were in co-residency with the parents, they also contribute to the household income, which may result in reverse causation. I also conducted analysis with the proportion as the instrument, the results suggest the same conclusion as presented.

Education attainment is measured as the year of education completed. The main explanatory variable is the sibling size of the family. It equals to the number of children born to the mother minus one. Throughout the analysis, I controlled for four sets of variables when assessing the relationship between sibling size and educational attainments: 1) the demographic features of the child: gender and age gap between the oldest and youngest children in the family; 2) the characteristics of the mother: her age at first birth and year of schooling; 3) father's year of schooling and the logarithm of annual net family income.

The Validity of the Instruments

Why IV yields a more sizable negative estimate than OLS? I explore two possibilities: 1) the violation of exclusion restriction, and 2) the IV estimates the Local Average Treatment Effect (LATE) for compliers whose behaviors are affected by the policy regulations. The differences between the OLS estimate and IV estimate indicate the heterogeneity in the treatment effect. I conducted the following analysis to provide evidence for the exclusion restriction assumption and see who the compliers might be.

Exclusion Restriction Assumption

First, as mentioned, the exclusion restriction assumption can be tested statistically when the instrumental variable is not precisely identified. I perform the Sargan-Hansen test (1988) for over-identification, and results show that there lacks sufficient evidence to reject the null hypothesis that the IVs are exogenous (p=.982).

Second, I utilize the fact that the ethnic minorities are exempted from the fertility control to conduct a placebo test. The idea is that since the ethnic minorities are exempted from the fertility regulation, the instruments should not be correlated with the fertility outcomes of this group. Therefore, it provides a chance to test whether the instruments are correlated with the education outcomes after controlling for the sibling size. Among families where the first child aged 25 to 35, there are 189 cases where the mothers belong to the ethnic minority groups, and they resided in 54 different communities. First, I regressed their fertility outcomes on the instruments. The results show that fertility regulations and punishments are not associated with sibling size for ethnic minorities. Then, I regressed the educational attainment on the instruments controlling for the sibling size. The parameters for the instruments are not statistically significant (p=.655 for IVa and p=.525 for IVb). Therefore, the placebo test suggests that the instruments are not correlated with the education outcome. It can be perceived as evidence supporting the validity of the instruments.

Third, measurement errors in the instrumental variable could bias the estimates if the errors are generated systematically. A potential source of bias comes from the fact that local fertility policy may change over time. The regulations that the families faced when making fertility decisions may not be the same as they were measured in 2010. Thus, if the temporal changes in the regulations are associated with the outcome, for instance, places where the investment in education increased over time, are more likely to increase punishment on excess fertility, the measurement errors in the instruments can lead to the violation of the exclusion restriction and bias the results. To assess the robustness of the IV results, I computed the temporal variations in the fine level between 2010 and the child's birth year. The data on historical changes in the fine level comes from Ebenstein (2010). It is a provincial-level measurement of the monetary punishments for excess fertility in China from 1979-2000. The fine level is measured as the share of the fine of the average household annual income, ranging from 0.3 to 5. The fine rate at the year the child was born is compared with the fine rate in 2010. The error is computed as the difference between them as,

$$IV_{error} = \frac{Community\ Fine\ Rate\ (2010)}{Household\ Annual\ Income\ (2010)} - \frac{Provincial\ Fine\ Rate\ (Year\ of\ Birth)}{Average\ Household\ Income\ (Year\ of\ Birth)}$$

Then, I regressed the education outcome on the computed variable. The results show that they are not correlated, which suggests that the IV estimator is robust against the potential measurement errors.