Son Preference, Intrahousehold Discrimination, and the Gender Gap in Education in China

Chenxu Hu, Hao Guo, and Xiaozhou Ding*

This paper investigates the gender gap in education by dividing it into two parts: the part that comes from intrahousehold discrimination and the part that comes from outside of the family. We develop a novel approach to measure the gender gap in education due to intrahousehold discrimination. Using China Household Income Project (2013) survey data, we find that intrahousehold discrimination accounts for a large part of the gender gap in education. The gap is large and persistent over time in both rural and urban regions, although the overall gender gap in education has declined significantly over time.

Keywords: education; son preference; gender gap; intrahousehold discrimination

---

* Hu: School of Economics, Nanjing University of Posts and Telecommunications, Nanjing, China; Guo (corresponding author): Advanced Institute of Finance and Economics, Liaoning University, Shenyang, China; Ding: Department of Economics, Dickinson College, Carlisle, US. For helpful comments and suggestions, we thank Samuel Ingram, Sherry Yu, Miaomiao Zou and seminar participants at Nanjing Audit University, Chinese Economist Society North America Annual Conference, Southern Economic Association Annual Meetings, and University of Kentucky Labor Group Meeting.

1. Introduction

The gender gap in education has declined significantly in China during the past 30 years, but it still exists, especially in rural regions (Zeng et al., 2014). While many studies have provided explanations for the gender gap in education (Jayachandran, 2015; Bordalo et al., 2019), few have distinguished the gender gap in education from within the family and from outside of the family. In this paper, we argue that certain cultures such as son preference, lead to unfair sibling rivalry and unequal intrahousehold resource allocation, thus creating the gender gap in education. This type of discrimination against women is within the family and is related to culture and traditions. Therefore, it is hard to erase, and its effect can persist even if the economic conditions for discrimination have disappeared.

Understanding how intrahousehold discrimination contributes to the gender gap in education has important implications. Since there are biases against women from both society and the family, the overall gender gap in education reflects both discrimination from society and unequal treatment within the family. To what extent intrahousehold discrimination contributes to the gender gap in education is a separate question from measuring the overall gender gap and it is less known to researchers. Measuring the gender gap in education due to intrahousehold discrimination can help us understand the role of the household in gender discrimination.

In this paper, we develop a novel approach to measure the part of the gender gap in education attributed to intrahousehold discrimination in China. We define two types of households: mixed-sex households and same-sex households. Mixed-sex households refer to families that have both male and female children, while same-sex households denote families with children of the same gender, either male or female. If gender biases against women are prevalent in society and rooted among households, women who grow up in families with brothers face not only discrimination from society but also unequal treatment from their parents. In contrast, women who grow up with female siblings face only discrimination from society. By comparing the gender gap in education between families with only sons or daughters and families with both sons and daughters, we can measure the effect of intrahousehold discrimination on the gender gap in education.

We use data from the China Household Income Project (CHIP) 2013 survey to examine the gender gap in education due to intrahousehold discrimination. We find that the gender gap in education exists in both urban and rural China, and intrahousehold discrimination can explain a large part of it. In the urban region, the overall gender gap in education is 0.428 years, and intrahousehold discrimination accounts for 104% (0.446 years) of the gap; in the rural region, the gender gap in education is 1.121 years, and intrahousehold discrimination explains 71% (0.798 years) of the gap. More importantly, the overall gender gap in education has declined continuously from the cohort born in the 1950s to the cohort born in the 1980s in both urban and rural regions. However, the part due to intrahousehold discrimination has not experienced the same steady decline. It first decreased, and then increased for the cohort born after China's economic reform (the 1980s cohort). The results indicate that intrahousehold discrimination induced by son preference still carries effects today, and market power by itself is unlikely to eradicate gender inequality.

The results we found are robust to different samples and specifications. One concern is that some urban residents who changed household registration (hukou) status might have grown up in rural regions. In the robustness check, we restrict the urban sample to those who did not change hukou status from rural to urban. This barely changes our results. The other concern is that people in mixed-sex households may grow up with more siblings than people in same-sex households. Thus, they face more within-household competitions and show a larger gender gap in education. In the robustness check, we run regressions for families with an equal number of children separately. It shows that our main results are not driven by the discrepancy of sibling size between two types of families.

Next, we show that parental education and birth order can affect the gender gap in education induced by intrahousehold discrimination. Specifically, we find that mother's education helps to reduce intrahousehold discrimination significantly, while father's education only has a small and insignificant effect. Also, we find that birth order matters for the gender gap in education. We find that parents invest more in their first child, no matter son or daughter. Thus, in mixed-sex families, if the first child is a girl, the gender gap is smaller; if the first child is a boy, the gender gap is larger.

Finally, we examine at what stage the gender gap in education occurs. By analyzing school dropout rates at different stages of education, we find that intrahousehold discrimination causes girls more likely to not attend school, and more likely to drop out of middle school (less likely to attend high school) in urban regions. In rural regions, intrahousehold discrimination leads girls more likely to not attend school, and more likely to drop out both in primary and middle school.

Our paper contributes to the study of the gender gap in education. Existing literature usually treats the gender gap in education as the investment outcome of parents and does not distinguish the part of the gender gap from society and the part from the intrahousehold discrimination. The underlying assumption is that there will be no gender gap in education if parents treat their sons and daughters equally. However, many studies in developed countries, where son preference is less of a concern and parents invest in their sons and daughters equally, have found that women attain higher levels of education than men (Jacob, 2002; Cobb-Clark and Moschion, 2017). For example, Jacob (2002) finds that women in the US have a higher college attendance rate than men and argues that women's higher cognitive skills explain the "reverse gender gap". Buchmann and DiPrete (2006) find that the higher college completion rate of women in the US is related to women's superior academic performance. The advantages of women in education are not limited to higher education. From kindergarten to high school, girls have higher grades, higher reading skills, better social skills, and better classroom behavior. Girls even outpace boys in math and science, which are traditionally considered boys' advantageous fields (Buchmann, DiPrete and McDaniel, 2008). These findings mean that the gender gap in education usually found in societies with son preference underestimates the bias against women (Chen, Chen, and Liu, 2019). The contribution of this paper is that we carefully choose our comparison group so that the gender gap in education when no sibling gender rivalry takes place is controlled. Thus, our estimate measures the gender gap in education that is due to intrahousehold discrimination.

This paper is also related to a large strand of literature studying the effect of sibling rivalry and gender composition on education attainments (Garg and Morduch, 1998; Bauer and Gang, 2001; Shrestha and Palaniswamy, 2017). In an intrahousehold resource allocation framework such as Behrman et al. (1986), parents invest in their children based on different preferences for sons and daughters and the relative return to education for men and women. Theory predicts that parents invest more in their sons because of the son preference and higher return to education for men, thus girls with brothers would receive less investment from parents and have lower educational attainment. However, empirical works do not show a uniform result, and the result depends on culture, institutions, and levels of economic development. Studies based on developed countries have found mixed results. Butcher and Case (1994) find that women raised with only brothers have received, on average, significantly more education than women raised with any sisters in the US. However, Kaestner (1997) uses more recent data in the US and finds no significant effect of sibling sex composition on women's education outcome. Amin (2009) uses British data and finds no significant impact of sibling sex composition. Evidence from Germany shows that women with older brothers are less likely to finish tertiary education (Bauer and Gang, 2001). Studies using developing country data usually find that women who grow up with more brothers receive less education. For example, Lei et al. (2017) and Zheng (2015) find that in China, an increase in the proportion of female siblings increases the educational attainment of men, and to a lesser extent, the educational attainment of women. Morduch (2000) estimates that in Tanzania, moving from an all-brothers to all-sisters scenario raises a child's completed years of schooling by 0.44 years. Our work based on the case of China also shows that girls with brothers have gone through intrahousehold discrimination and attain less education.

The rest paper proceeds as follows. Section 2 discusses son preference and the gender gap in education. Section 3 shows the data and summary statistics. Section 4 presents our empirical approach of estimating the gender gap in education due to intrahousehold discrimination, and Section 5 presents the results and the trend of the gender gap in education due to intrahousehold discrimination. Section 6 did some robustness checks. Section 7 examines how the effect of intrahousehold discrimination on the gender gap in education is affected by parents' education and birth order, and at what education stage the gender gap occurs. Section 8 concludes.

2. Son Preference and the Gender Gap in Education

2.1 The Origins of Son Preference

Son preference exists in many countries and regions. An extensive strand of literature has examined why parents prefer sons over daughters. An economic explanation is that men have higher productivity in agricultural production, which is brawn-based (Rendall, 2017). In a traditional agricultural economy, parents prefer sons since the returns of investing in sons are higher. However, even though agriculture has become a small part of the economy and women's productivity in agriculture has increased due to technological improvements, the gender gap still exists. Moreover, people who migrate out of their original country and move into society without son preference can still maintain their preference for sons. For example, Almond and Edlund (2008) use the 2000 US census and find male-biased sex ratios among U.S.-born children of Chinese, Korean, and Asian Indian parents. Almond, Edlund, and Milligan (2013) examine immigrants from South and East Asian countries in Canada and find those sex ratios are substantially elevated at higher parities if the previous children were all girls. However, a biased sex ratio is not found in families with Christian or Muslim families. All the above evidence suggests that in addition to economic incentives, there must be some cultural and institutional reasons that impede gender equality.

One important reason for son preference is for old-age support. In many cultures, people traditionally depend on their sons for old-age support. The oldest son has the greatest responsibility to support his parents. This mechanism still acts as a substitute for pension in places where a national pension system does not function well. In China, many old people still co-reside with and depend on their sons (Rosenzweig and Zhang, 2014), especially in the rural region. On the other hand, a daughter joins her husband's family and can provide limited support to their birth parents after marriage. Thus, the value of raising a daughter is lower than the value of raising a son. There is a proverb in China that states "married daughters are like poured water," which describes the view of parents on raising a daughter.

Other cultural reasons are also at work. Das Gupta et al. (2003) argue that the persistence of son preference in East and South Asia is due to the rigid patrilineal family system in these societies. In these regions, men constitute the social order and pass down the family lineage, and women are in subordinate positions. Only sons can inherit family assets, such as land and houses. They also find that even if old women in these societies have larger power within the family, the increase in women's status depends on having grown sons. This is another force for son preference and could lead women to have a stronger son preference than their husbands.

Ancestor worship also plays an important role in son preference in China and other Confucius societies (Chung and Gupta, 2007). Confucius culture considers it vital to pass down the family lineage. Since only sons can bear the responsibility of passing down the family linage, not having a son constitutes a major dereliction of duty. The fear of ending the family lineage gives parents, especially the father, a strong incentive to have and invest in sons.

2.2 Evolution of the Son Preference

The preference for sons is rooted in economic and cultural reasons, and it changes with economic development, demographic changes, and changes in social institutions. As the economy develops, the returns to education for women increase. Economic development presumably will ease the family budget constraints and mitigate the gender gap in education. That is, the large gender gap in many developing countries should be explained by underdevelopment itself.

Demographic changes also affect son preference and the gender gap in education. Research has found that China's one-child policy has led to more biased male-female birth ratios (Ebenstein, 2010). However, the forced fertility decline did increase the family resources allocated to daughters, and in cases of daughter-only families, girls face no sibling rivalry from brothers. Urban daughters in China have benefited from the demographic change brought by China's one-child policy (Fong, 2002). The low fertility rate induced by the one-child policy empowers women in two ways. First, it enables mothers to go to work and gain the ability to support their own parents, which gives daughters the ability to provide old-age support for parents. Second, singleton daughters have no brothers for their parents to favor, so they receive all investment from their parents and therefore have higher educational attainment compared to earlier cohorts not affected by fertility policies.

Changes in social institutions also influence son preference and the gender gap. One reason for son preference is that parents rely on sons for old-age support. The development of a national pension system and public and private insurance systems can help reduce such reliance. For example, Ebenstein and Leung (2010) study a voluntary pension program in rural China and find that the increase in pension participation is associated with a less skewed sex ratio at birth. Implementation of the compulsory education system and laws that give daughters equal rights to inherit family assets also help to decrease the bias against women within the family.

In sum, although economic development can change society's attitude toward son preference (Chung and Gupta, 2007), the process is slow since son preference is shaped by culture and social institutions. The culture and social norms caused by son preference in the past still exist. Therefore, in the following sections, we examine how the gender gap in education can be explained by intrahousehold discrimination, how it differs between urban and rural regions, and how it changes over time.

3. Data and Descriptive Evidence

3.1 Data Source

We use data from the China Household Income Project (CHIP). It is conducted every few years and is a large, representative national survey containing both urban and rural households. The most recent wave is 2013. The CHIP 2013 survey covered 234 counties in 15 provinces and contains information on each respondent's gender, educational attainment, working status, income, and many other personal characteristics. As a household survey, it also contains information on the family structure. Moreover, it selected one person (usually the household head or spouse) from the surveyed household and asked for information about his/her siblings' gender, birth year, education, etc[1]. Based on this information, we can infer the sibling structure of the household head or spouse and link their sibling structure with their educational attainment, current working status, and income. Most household surveys in China only contain information on younger children, and they do not trace children's income and education attainment after the children moved out of their original families. Thus, they cannot be used to analyze how sibling structure affects adults' education attainment. The unique feature of the CHIP 2013 allows us to obtain the sibling structure of adults who have finished education and have probably moved away from their original family. Therefore, we can study the effect of sibling structure on the gender gap in education.

The CHIP 2013 contains 10,490 households and 39,065 individuals in rural areas and 6,674 households and 19,887 individuals in urban regions. In our analysis, we use only household heads or spouses who are surveyed about their siblings' information. For each respondent, we expand each of their siblings as a new observation and append it with the main data set. We exclude people whose age is over 60 or below 24 because our primary focus in this paper is prime age workers who have completed their education[2]. We also drop individuals whose sibling structure cannot be identified from the data. For our main analysis, our final sample contains 15,018 observations in the urban sample and 27,703 observations in the rural sample.

3.2 The Gender Gap in Education in the Urban and Rural Regions

The labor market is segregated in China and it is mainly due to the household registration system (hukou) that separates the urban and rural labor markets (Meng, 2012). The urban region has been built as the production center of industrial goods, and urban residents who work in public sectors are eligible to receive a pension from their working unit after retirement. However, in rural regions, agricultural people are engaged in farming, and there is no pension system for them. The hukou system causes a large urban-rural divide not only in economic disparities but also in educational inequalities (Knight and Li, 1996; Liu, 2005). Therefore, in this paper, we keep the dichotomy and divide the data into the urban sample and the rural sample.

Table 1 shows the summary statistics for male and female individuals in same-sex and mixed-sex families in the urban area. On average, men have 0.345 years more education than women and the difference is significant. However, in same-sex families, women have 0.317 years more education than men. In mixed-sex families, women have 0.444 years less education. The data show a clear pattern that the women who grow up with no brothers have a higher level of education than men with no sisters, but women who grow up with brothers have less education than men with sisters. Of course, men and women are different in many other dimensions. Men are older and have fewer siblings in both same-sex and mixed-sex families. In general, women also grow up in families with higher parents' education, and this is more pronounced in same-sex families. The reason is that parents with higher education generally have lower levels of son preference and are less likely to select their children's sex. Thus, they are more likely to have daughters and we see girls' parents have higher education from the data[3]. All these differences could affect the gender gap in education differently between same-sex and mixed-sex families. We control for these differences in our regression analysis.

In the rural sample, as shown in Table 2, the gender gap in education is larger. The overall gender gap in education in rural China is 1.086 years. In mixed-sex families, men obtain 1.156 years more education than women. In same-sex families, men obtain only 0.102 years more education, and the difference is not significant. Again, the gender gap in education is more pronounced in mixed-sex families. This implies that much of the discrimination against women happens within the family. In the rural sample, we also find that women are younger and have more siblings than men in same-sex families. In mixed-sex families, the age gap is smaller and the difference in the number of siblings is insignificant.

4. Empirical Approach

For the main analysis, we incorporate the basic models for studying the gender gap (Altonji and Blank, 1999; Yu and Su, 2006; Autor et al., 2016) and sibling gender composition effects (Butcher and Case, 1994; Amin, 2009; Zheng 2015) by including an interaction term of the gender dummy and family type. Our regression model is specified as:

$$y_{ij} = \beta_0 + \beta_1 Mixed_j + \beta_2 Female_i + \beta_3 Mixed_j \times Female_i + Z\gamma + \theta_p + \varepsilon_{ij} \quad (1)$$

where y_{ij} is the years of education for individual i from household j. $Mixed_j$ equals one if the individual is from a mixed-sex family, and zero if the person comes from a same-sex family. $Female_i$ is a gender dummy that equals one for women. Z contains a set of control variables at the individual and household level, including birth order, number of siblings, age, age squared, and parental education. We also include a set of provincial level fixed effects (θ_p) to account for unobserved heterogeneity affecting education attainment across geographical units[4]. We include these variables to control for the differences between men and women in mixed-sex and same-sex families. Finally, we cluster the error term at the household level for all regressions to account for the potential common unmeasured household effect (Parish and Willis, 1993).

Our measure of the gender gap in education due to intrahousehold discrimination is β_3. Women who grow up in mixed-sex families face both discriminations from society and unequal treatment from their parents because of son preference. In contrast, women who grow up with only female siblings face no intrahousehold discrimination. Thus, by comparing the gender gap in education between mixed-sex and same-sex families, the coefficient of the interaction term β_3 measures the gender gap in education due to intrahousehold discrimination.

Note that the gender composition of a family is not determined randomly because parents may partially manipulate children's gender composition through prenatal selection (such as ultrasound detection and abortion) or postnatal selection (such as abandoning infants). Appendix Table A.1 presents evidence for non-random gender composition and reveals that son preference is the driving force. In both urban and rural samples, the shares of mixed-sex families are higher than the theoretical shares for families with up to 4 children. In addition, the shares of all-daughter families are lower than the shares of all-son families, especially in the rural region. Thus, it is plausible that the level of the son preference in mixed-sex families does not represent the average level of the population and β_3 does not identify the gender gap in education caused by differences in sibling gender composition.

However, the goal of this paper is to measure the part of the gender gap in education due to intrahousehold discrimination, and it does not rely on the exogeneity of sibling gender composition. In same-sex families, girls do not face sibling gender rivalry and there is no intrahousehold discrimination based on son preference. In mixed-sex families, son preference leads to intrahousehold discrimination against girls. Estimate of β_3 measures the average level of intrahousehold gender discrimination in education imposed by those mixed-sex families. Since only mixed-sex families impose intrahousehold discrimination based on children's gender, β_3 measures the average level of intrahousehold discrimination in society.

5. Results

5.1 The Gender Gap in Education Due to Intrahousehold Discrimination

Table 3 reports the regression results for both the urban and rural samples. Columns (1) and (2) show the gender gap in education for same-sex families and mixed-sex families in the urban region. Controlling for birth order, number of siblings, age, age squared, and province fixed effects, women from same-sex families attain only 0.018 fewer years of education than men from same-sex families and the difference is statistically insignificant. It shows that when intrahousehold discrimination does not exist, girls can reach the same educational attainment as boys in the urban region. However, women from mixed-sex families have 0.491 fewer years of education and the result is significant. It shows that girls with brothers attain less education than their brothers due to intrahousehold discrimination. Results in column 3 show that the overall gender gap in education is 0.428 years, while results in column 4 show that the gender gap caused by intrahousehold discrimination is 0.446 years. In the urban regions, intrahousehold discrimination can explain 104% of the overall gender gap.

The story in the rural sample is different. In the rural region, girls receive 0.445 fewer years of education than men even if they do not face intrahousehold discrimination, as shown in column 5. This could be explained by the following reasons. First, women traditionally do not work in rural regions and the returns to education for women is low. Second, people in rural regions rely heavily on sons for old-age support, and people who do not have sons may choose other ways than investing in their daughters for old-age support. Third, society has lower expectations of women's education. Thus, girls attain less education in the rural region even when they do not face gender rivalry from brothers. When girls face intrahousehold discrimination, the gender gap in education increases to 1.211 years, as shown in column 6. Results in column 7 show that the overall gender gap in education in the rural region is 1.121 years. The estimate of the interaction term in column 8 shows that intrahousehold discrimination cause women to have 0.798 fewer years of education than men in rural China. Thus, in the rural region, intrahousehold discrimination can explain about 72% of the gender gap in education.

Since our method to measure the gender gap in education due to intrahousehold discrimination is unique, few comparable results exist. Existing literature usually examines the effect of having one more male sibling on other children's educational attainment. (Zheng, 2015, Lei et al. 2017). While our results cannot compare with theirs directly, our results do corroborate their findings that having more male siblings is detrimental to one's educational attainment.

5.2 Trends of Intrahousehold Discrimination

Previous results show that intrahousehold discrimination explains a large part of the overall gender gap in education, both in urban and rural regions. How does the gender gap in education due to intrahousehold discrimination change over time? In this section, we first show the declining trend of the gender gap in education and then study how the part due to intrahousehold discrimination evolves over time. If the gender gap in education is caused by underdevelopment, we should see declining trends of both the overall and the intrahousehold part of the gender gap. However, if the gender gap is mainly caused by intrahousehold discrimination that is rooted in the culture of son preference, then development itself is unlikely to erase the gender gap in education.

We first present the overall gender gap by birth cohort. The sample is divided by birth cohort; specifically, we split the sample into people who were born in the 1950s, 1960s, 1970s, and 1980s. We first run a simple regression model:

$$y_i = \beta_0 + \beta_1 Female_i + Z_i'\gamma + \theta_p + \varepsilon_i, \qquad (2)$$

where y_i denotes the years of education, and $Female_i = 1$ if the individual is a woman. We include the same set of covariates as in Equation 1, and control for province fixed effect θ_p. The coefficient β_1 measures the gender gap in education. We run the regression for each birth cohort for both urban and rural samples.

Figure 1 depicts the trends of the gender gap in education for both urban and rural regions. The blue line with square dots depicts the trend in the urban region, and the red line with circle dots shows the trend in the rural region. In both rural and urban regions, the gender gap in education has declined significantly for people born in later cohorts. In the urban region, the gender gap in education has declined from 0.822 years to negative 0.104

years, in favor of women, although the effect for the 1980s cohort is not significant. In the rural region, the gender gap in education has declined from 1.953 years to merely 0.151 years, and the gap has become statistically insignificant. The diminishing overall gender gap in education has been well-documented with different datasets. (Treiman, 2013; Zeng et al. 2013).

We next estimate the gender gap due to intrahousehold discrimination for each decade in both the urban and rural regions using Equation 1. We then plot the coefficient of $Female_i \times Mixed\_sex_j$ against birth cohort. The results are shown in Figure 2. The blue line with square dots shows the trend in the urban region, and the red line with circle dots shows the trend in the rural region.

In the urban sample, the gender gap in education due to intrahousehold discrimination is 1.349 years for the 1950s cohort. However, for people born in the 1960s and later, the effect is smaller and insignificant. This could be because, after the founding of the People's Republic of China, urban regions were organized as production centers of industrial products, workers were assigned to work units, and the Chinese government established the pension and public health insurance systems in the urban region. Thus, urban parents do not have to rely on sons for old-age support and the incentives for them to prefer and invest more on sons have also decreased.

Things are very different in rural regions. As shown in the figure, the gender gap in education due to intrahousehold discrimination is 1.134 for the 1950s cohort, but the gap does not decrease continuously with time. The gap declined for those born in the 1960s and 1970s but rise back again to 1.1 years for those born in the 1980s. One important reason why the effect of intrahousehold discrimination does not diminish with economic development in rural China is that people rely on sons for old-age support[5] (Ebenstein and Leung, 2010; Das Gupta et al., 2003), and the expectation of the return to education affect how much parents invest in their son's education. In the 1960s and 1970s, especially during the cultural revolution, the return to education is low, so investing more education on sons is not that profitable. After China's economic reform in the 1980s, the return to education increased significantly. Thus, parents would invest more in their son's education. This investment incentive enlarges the gender gap in education.

6. Robustness

6.1 Redefine Urban Sample

As there are internal migrations from rural to urban regions, our urban sample contains people who might have grown up in rural regions.[6] Unfortunately, the survey does not ask where the respondent grew up. We use the change of hukou status to infer if an urban resident grew up in rural regions. If a respondent reports ever changed hukou status from rural to urban, we consider her grew up in rural regions. In our expanded urban sample, about 34% of observations have reported changing hukou status from rural to urban. We delete these observations and run regression again for this restricted urban sample. The results are shown in Table 4.

After dropping those who potentially grew up in rural regions, the regression results are quantitively similar to our main results. In same-sex families, there is a small and insignificant reverse gender gap in education (0.03 years); in mixed-sex families, the gender gap in education is 0.352 years. The overall gender gap in education is 0.3 years and the part due to intrahousehold discrimination is 0.348 years. Therefore, intrahousehold discrimination can account for 116% of the overall gender gap.

6.2 Controlling for Number of Siblings

As indicated by Table A.1, the share of respondents from mixed-sex families increases as the number of siblings increases. For example, when there are two siblings, the share of people from mixed-sex families is 57.8% in the urban region and 51.6% in the rural region; when there are three siblings, the share is 82.4% in the urban region and 77.5% in the rural region; when there are four siblings, the share is 90.7% in the urban region and 87.5% in the rural region; when there are five siblings, the share reaches to 93% in the urban region and 92.1% in the rural region. This means that people from mixed-sex families, on average, grow up with more siblings and face fiercer sibling rivalry. This could aggravate the gender gap in education. Thus, in this section, we examine the effect of intrahousehold discrimination for individuals with the same number of siblings. We run our regression Equation (1) for families with two to five siblings separately. We leave out observations with more than five siblings because the share of individuals from same-sex families is too small and the comparison between same-sex families and mixed-sex families is likely to be contaminated by the difference in the number of siblings and the variation of the error term.

Table 5 shows the results for observations with two to five siblings separately. The first four columns present the results for the urban sample, and the remaining four columns show the results for the rural sample. In the urban region, the effects of intrahousehold discrimination are small and insignificant irrespective of the number of siblings. In the rural region, the gender gap in education due to intrahousehold discrimination is 0.657 years for families with two children, 0.574 years for families with four children, and 1.294 years for families with five children. The effect is not significant for families with three children. Overall, the results show that even when using families with the same number of siblings, a large part of the gender gap in education can still be explained by intrahousehold discrimination. This verifies that our main findings in Table 3 are not driven by differences in the number of siblings between same-sex and mixed-sex families.

7. Heterogeneity

7.1 Parental Education and Intrahousehold Discrimination

Children's educational attainment is correlated with their parents' education (Lillard and Willis, 1994; Currie and Moretti, 2003; Black, Devereux, and Salvanes, 2005). When parents are more educated, they are better aware of the returns to education and are more likely to invest in their daughters. In this section, we examine whether parents' education affects intrahousehold discrimination and the gender gap in education. We divide both the urban and rural samples into two subsamples depending on whether the observed individual's father or mother has attended school[7] and test whether parental education helps alleviate the gender gap in education caused by intrahousehold discrimination.

In Table 6, we examine whether the gender gap in education due to intrahousehold discrimination is affected by the mother's education. In the urban region, if the mother has never attended school, the gender gap in education due to intrahousehold discrimination is 0.791 years. However, if the mother has received any education, the effect of intrahousehold discrimination becomes small and insignificant. The case is similar in the rural region, where the gender gap in education due to intrahousehold discrimination is 1.016 years if the mother has never attended school, but only 0.337 years and insignificant if the mother has received any education.

Table 7 shows the results for the father's education. Similar to previous results, the part of the gender gap due to intrahousehold discrimination is smaller for families in which fathers have any level of education in the urban area. However, in the rural region, the gap is still large and significant regardless of whether the father attended school or not, although it is smaller if the father has received some education.

The results in Table 6 and Table 7 show that mother's education help to reduce intrahousehold discrimination more than father's education. This corroborates the findings that mothers care more about daughters than fathers do, and an increase in the mother's education level or income helps alleviate the gender inequality (Thomas, 1990; Qian, 2008).

7.2Firstborn Son and Firstborn Daughter Effect

Literature has shown that the firstborn child receives more attention from parents. In societies with son preference, the firstborn son receives even more care and investments from parents. In China, especially in rural areas, parents usually live with their eldest son for old-age support. The eldest son also bears the responsibility of carrying down the family lineage. Thus, parents have higher expectations and invest more resources in their eldest son. However, if the firstborn is a daughter, parents may be reluctant to invest in her and often divert resources to her younger brothers. For example, Parish and Willis (1993) find that in Taiwan, having older sisters increases the educational attainment of younger siblings of both sexes. Yu and Su (2006) also show that in Taiwan, firstborn sons have additional advantages in sibling rivalry, while firstborn daughters do not enjoy the same advantage. Thus, firstborn daughters may face significant disadvantages within the family.

In this section, we divide the mixed-sex families into those in which the firstborn child is a son and those in which the firstborn child is a daughter. We calculate the gender gap in education between the firstborn son and their younger sisters and compare this gap with the gender gap for same-sex families[8]. We do similar estimations for the firstborn daughter. The results are shown in Table 8.

In the urban region, for families with firstborn sons, intrahousehold discrimination due to the firstborn son is 0.592 years. This estimate is quantitatively larger than our baseline estimate in Table 3, indicating that the firstborn son may enjoy more resources than average.

For mixed-sex families with firstborn daughters, the gender gap in education due to intrahousehold discrimination for the firstborn daughter is only 0.25 years and it is statistically insignificant. The number is smaller than the baseline estimate. This shows that being the firstborn gives the daughter certain advantages to mitigate intrahousehold discrimination.

The results are similar for the rural sample. Compared with the baseline estimates, a firstborn son's advantage in education due to intrahousehold discrimination is larger; for firstborn daughters in mixed-sex families, the gender gap in education due to intrahousehold discrimination is smaller. The above results imply that the firstborn child receives more resources within the family. Thus, for men, being the firstborn magnifies his advantage and leads to a higher gender gap in education; for women, being the firstborn alleviates the disadvantage and decreases the gender gap in education.

7.3 Intrahousehold Discrimination and School Dropout

When family resources are limited and parents invest more in their sons, at some point, parents will need to decide whether to stop investing in their daughter's education. At what stage do parents decide to stop investing in their daughters' education? In this section, we examine how intrahousehold discrimination affects the gender gap in school dropout at various education stages. We calculate the share of men and women who never attended school in both urban and rural regions. Upon attending school, we calculate the share of people whose highest level of education is primary school. We define it as the primary school dropout rate. Note that we define a person as having dropped out of primary school when he/she attended primary school and his/her highest education level is primary school, regardless of whether the person finished primary school education or not. Upon completing primary school, we also calculate the middle school dropout rate using the same procedure. Table 9 shows dropout rates at different stages of education.

The first three columns show the results for the urban sample. The share of people who have never attended school is very small: 3.1% for women and 1.7% for men. The primary school dropout rate is 12.2% for women and 9.7% for men. The middle school dropout rate is large, 42.6% for women and 42.9% for men. The large middle school dropout rate is mainly because high school is not included in the compulsory education system in China.

The next three columns show that the dropout rates at all levels of education are higher in the rural region than in the urban region for both men and women. In the rural region, 11.9% of the women never attend school; the number is only 5.2% for men. The primary school dropout rate is 40.2% for women and only 28.1% for men. For women, the middle school dropout rate is as high as 80.5%; the number is slightly lower for men, at 76%.

In both the urban and the rural regions, girls are less likely to attend school and are more likely to drop out of school. To examine how intrahousehold discrimination affects the gender gap in school dropout rates, we run Equation (1) but change the dependent variable to the dummy variable indicating whether a person has never attended school, whether a person dropped out of primary school, and whether a person dropped out of middle school. The first three columns in Table 10 show the results for the urban sample. Girls are more likely not to attend any school than boys in mixed-sex families when compared with same-sex families. The gender gap in the primary school dropout rate is not significantly different between same-sex families and mixed-sex families, but the gender gap in the middle school dropout rate is much higher in mixed-sex families. Women from mixed-sex families are 9.19% less likely to go to high school than their brothers compared to women from same-sex families. Columns (4)-(6) report the results for the rural region. In rural regions, intrahousehold discrimination leads girls to be more likely to drop out at all levels of education. Girls are 5.45% less likely to attend school, 6.22% less likely to go to middle school, and 6.59% less likely to attend high school.

The different effects of intrahousehold discrimination on the gender gap in dropout rates between urban and rural regions may be due to the differences in access to primary and middle school. In the urban region, household income is higher and the urban system offers more educational opportunities, so it is less urgent to let daughters drop out of school at an early stage. However, because high school admission is competitive, parents tend to support their sons rather than their daughters to go to high school. In the rural region, family income is low, and education resources are scarce. Parents face the choice of which child to support in education. Thus, we see a significant effect of intrahousehold discrimination at all levels of education.

Of course, job opportunities and social expectations could also play a role. In the urban region, women can become workers when they grow up. Education is rewarded at least at the middle school level. Thus, parents may still do their best to help their daughters finish middle school. In the rural regions, non-agricultural employment is rare for women, and the return to education in agricultural work is low. Thus, women's education is often considered futile. When a family has sons, parents tend to divert resources to their sons' education.

8. Conclusion

In this paper, we develop a novel approach to measure the gender gap in education due to intrahousehold discrimination. Using the CHIP 2013 data, we find that in both the urban and rural regions, intrahousehold discrimination accounts for a large share of the gender gap in education. While the overall gender gap in education has declined steadily with economic development, the part due to intrahousehold discrimination has persisted over the years, especially in rural regions. We argue that this persistent effect implies that the culture of son preference has not disappeared with economic development, and it still causes intrahousehold discrimination and leads to a gender gap in education.

Our results are robust to a series of checks. Given that some urban residents grew up in rural regions, we delete those who have changed hukou status from rural to urban and show that this does not change our main results. To account for the fact that same-sex families on average have fewer children than mixed-sex families, we run regressions for families with different sibling sizes separately, and we find that intrahousehold discrimination still exists and is significant for each sibling size in the rural sample.

We also find that the level of intrahousehold discrimination and the gender gap in education correlate with parents' education and birth order. We find that mother's education has a significant effect on closing the gender gap in education caused by intrahousehold discrimination, while father's education has a limited role. Also, being the firstborn provides additional advantages in sibling rivalry, meaning that the firstborn son receives more education and the gender gap in education is higher, while the firstborn daughter faces less severe intrahousehold discrimination and the gender gap in education is lower. Finally, we find that in the urban region, the gender gap in school dropout due to intrahousehold discrimination mostly occurs in middle school, while in the rural region, intrahousehold discrimination causes girls to have higher dropout rates at all levels of education.

Nowadays, as China liberalizes its one-child policy, family size may increase and intrahousehold discrimination could again be an important determinant of the gender gap in the future. Our paper is relevant to several policy implications for reducing the gender gap in education. The analysis shows that increasing the education of mothers may help to close the gender gap, and expanding compulsory education also help to reduce the gender gap. Our analysis also suggests that improving the public pension system can reduce parents' incentives to invest more in sons and lower the gender gap in education.

References

Almond, D., & Edlund, L. (2008). Son-Biased Sex Ratios in the 2000 United States Census. *Proceedings of the National Academy of Sciences*, 105 (15), 5681-5682.

Almond, D., & Milligan, K. (2013). Son Preference and the Persistence of Culture: Evidence from South and East Asian Immigrants to Canada. *Population and Development Review*, 39 (1), 75-95.

Altonji, J., & Blank, R. (1999) Race and Gender in the Labor Market, *Handbook of Labor Economics,* 3,3143-3259.

Amin, V. (2009). Sibling Sex Composition and Educational Outcomes: A Review of Theory and Evidence for the UK. *Labour*, 23 (1), 67-96.

Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016) School Quality and the Gender Gap in Educational Achievement. *American Economic Review*,106(5), 289-295.

Babiarz, K. S., Ma, P., Miller G., &Song S. (2018). The Limits (and Human Costs) of Population Policy: Fertility Decline and Sex Selection in China Under Mao, *NBER Working Paper No. 25130*

Bauer, T., & Gang, I. N. (2001). Sibling Rivalry in Educational Attainment: The German Case. *Labour*, 15 (2), 237-255.

Behrman, J. R., Pollak, R. A., & Taubman, P. (1986). Do Parents Favor Boys*? International Economic Review*, 33-54.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *The American Economic Review*, 95 (1), 437-449.

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2019). Beliefs about Gender. *American Economic Review*, 109 (3), 739-773.

Buchmann, C., & DiPrete, T. A. (2006). The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement. *American Sociological Review*, 71 (4), 515-541.

Buchmann, C., & McDaniel, A. (2008). Gender Inequalities in Education. *Annual Review of Sociology*, 34 (1), 319-337.

Butcher, K. F., & Case, A. (1994). The Effect of Sibling Sex Composition on Women's Education and Earnings. *The Quarterly Journal of Economics*, 109 (3), 531-563.

Chen, S. H., Chen, Y., & Liu, J. (2019). The Impact of Family Composition on Educational Achievement. *Journal of Human Resources*, 54 (1), 122-170.

Chung, W., & Gupta, D. G. (2007). The Decline of Son Preference in South Korea: The Roles of Development and Public Policy. *Population and Development Review*, 33 (4), 757-783.

Cobb-Clark, D. A., & Moschion, J. (2017). Gender Gaps in Early Educational Achievement. *Journal of Population Economics*, 30 (4), 1093-1134.

Currie, J., & Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*, 118 (4), 1495-1532.

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

Ebenstein, A., & Leung, S. (2010). Son Preference and Access to Social Insurance: Evidence from China's Rural Pension Program. *Population and Development Review*, 36 (1), 47-70

Fong, V. L. (2002), "China's One-Child Policy and the Empowerment of Urban Daughters. *American Anthropologist*, 104 (4), 1098-1109

Garg, A., and Morduch, J. (1998). Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana. *Journal of Population Economics,* 11 (4), 471-493

Gupta, M. D., Jiang, Z., Li B., Xie, Z., Chung W., & Bae, H. (2003). "Why is Son preference so Persistent in East and South Asia? A Cross-Country Study of China, India and the Republic of Korea. *The Journal of Development Studies*, 40 (2), 153-187.

Jacob, B. A. (2002), "Where the Boys Aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education. *Economics of Education review*, 21 (6), 589-598

Jayachandran, S. (2015), "The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 7 (1), 63-88

Kaestner, R. (1997). "Are Brothers Really Better? Sibling Sex Composition and Educational Achievement Revisited. *Journal of Human Resources*, 32 (2)

Knight, J., & Li (1996), Educational Attainment and the Rural-Urban Divide in China. *Oxford Bulletin of Economics and Statistics*, 58 (1)

Lei, X., Shen, Y., Smith, J. P., & Zhou, G. (2017). Sibling Gender Composition's Effect on Education: Evidence from China. *Journal of Population Economics*, 30 (2), 569-590

Lillard, L. A., & Willis, R. J. (1994). Intergenerational Educational Mobility: Effects of Family and State in Malaysia. *The Journal of Human Resources*, 29 (4), 1126-1166.

Liu, Z. (2005). Institution and Inequality: The Hukou System in China. *Journal of Comparative Economics*, 33 (1), 133-157

Meng, X. (2012). Labor Market Outcomes and Reforms in China. *The Journal of Economic Perspectives*, 26 (4), 75–102

Morduch, J. (2000). Sibling Rivalry in Africa. *American Economic Review*, 90 (2), 405-409.Parish,

Parish, W.L. & Willis, R. J. (1993). Daughters, Education, and Family Budgets Taiwan Experiences. *Journal of Human Resources*, 863-898

Qian, N. (2008). "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 2008, 123 (3), 1251-1285

Rendall, M. (2017). Brain versus brawn: the realization of women's comparative advantage. University of Zurich, Institute for Empirical Research in Economics, Working Paper (491).

Rosenzweig, M. R., & Zhang, J. (2014). Co-Residence, Life-Cycle Savings and Intergenerational Support in Urban China. *NBER Working Paper No. 20057*

Shrestha, S. A., & Palaniswamy, N. (2017). Sibling Rivalry and Gender Gap: Intrahousehold Substitution of Male and Female Educational Investments from Male Migration Prospects. *Journal of Population Economics*, 30 (4), 1355-1380.

Thomas, D. (1990). Intra-Household Resource Allocation: An Inferential Approach. *Journal of Human Resources*, 25 (4), 635–664.

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

Yu, W. H., & Su, K. H. (2006). Gender, Sibship Structure, and Educational Inequality in Taiwan: Son Preference Revisited. *Journal of Marriage and Family*, 68 (4), 1057-1068.

Zeng, J., Pang, X., Zhang, L., Medina, A., & Rozelle, S. (2014). Gender Inequality in Education in China: a Meta-Regression Analysis. *Contemporary Economic Policy*, 32(2), 474-491.

Zheng, L. (2015). Sibling Sex Composition, Intrahousehold Resource Allocation, and Educational Attainment in China. *The Journal of Chinese Sociology*, 2 (1)

---

[1] 2013 is the only year that CHIP contains this information. The previous waves of the data have no information on the household head's or spouse's siblings.

[2] Our sample contains only those between 24 and 60. We assume that people have finished education by age 24. The results are not driven by this particular lower bound. We have tested using age 22 as our lower bound and the results do not change. The results are not reported in the main paper and are available upon request.

[3] Using the same dataset, we can also find that families with higher parental education also have a high share of daughters. This indicates that more educated people have a lower level of son preference and are less likely to select children's gender. See details in appendix B.

[4] Presumably controlling for county fixed effects makes our regression more convincing. However, we know only the respondent's current county of living, not the county where they grew up and received education. Considering that many people migrate out of their birth county after they grow up, the current county of living may not be a good proxy for the birth county. Here, we control for province fixed effects because migrations among provinces are not usual for our sample period. Thus, we use the current province of living as a proxy for the province they grew up.

[5] Using the CHIP rural sample, we find that for people aged over 60 who live with their children, only 9.7% of them live with daughters. Most elderly people in the rural region still rely on sons for old-age support. See details in appendix C.

[6] Migrations from urban to rural regions are rare. The survey does not ask if a rural resident changed her hukou status from urban to rural.

[7] In the urban region, 30.18% of the respondents' fathers never attended school, and 46.73% of the respondents' mothers never attended school. In the rural region, 50.81% of the respondents' fathers never attended school, and 71.46% of the respondents' mothers never attended school.

[8] This is realized by regressing Equation 1 with data for individuals from same-sex families and individuals from mixed-sex families where the firstborn is a son but excluding younger sons in those families.

Table 1: Gender Gap in Education and Personal Characteristics: Urban Sample

	Same-sex			Mixed-sex			Total		
	Male	Female	Diff.	Male	Female	Diff.	Male	Female	Diff.
Years of Edu.	11.84	12.16	0.317*	10.68	10.24	-0.443***	10.91	10.56	-0.345***
	(3.448)	(3.402)	(2.42)	(3.300)	(3.501)	(-7.20)	(3.361)	(3.558)	(-6.09)
Age	43.38	41.37	-2.005***	47.28	46.84	-0.435**	46.51	45.92	-0.599***
	(9.415)	(9.290)	(-5.60)	(8.175)	(8.095)	(-2.96)	(8.572)	(8.559)	(-4.29)
No. siblings	1.592	1.711	0.119*	3.213	3.279	0.0661*	2.895	3.013	0.118***
	(1.252)	(1.486)	(2.28)	(1.493)	(1.522)	(2.43)	(1.585)	(1.626)	(4.48)
Birth order	1.716	1.803	0.0871*	2.561	2.605	0.0434	2.396	2.469	0.0727**
	(0.959)	(1.202)	(2.10)	(1.510)	(1.548)	(1.57)	(1.458)	(1.525)	(2.98)
No. male sib.	2.551	0	-2.551***	2.329	1.809	-0.520***	2.373	1.502	-0.871***
	(1.245)	(0)	(-74.00)	(1.126)	(0.959)	(-27.64)	(1.153)	(1.107)	(-47.21)
No. female sib.	0	2.660	2.660***	1.907	2.497	0.590***	1.534	2.524	0.991***
	(0)	(1.455)	(69.27)	(1.023)	(1.241)	(28.60)	(1.189)	(1.281)	(49.06)
Father attend school	0.746	0.859	0.112***	0.666	0.685	0.0198*	0.681	0.714	0.0329***
	(0.435)	(0.349)	(7.25)	(0.472)	(0.464)	(2.30)	(0.466)	(0.452)	(4.32)
Mother attend school	0.639	0.714	0.0746***	0.491	0.511	0.0196*	0.520	0.545	0.0248**
	(0.480)	(0.452)	(4.09)	(0.500)	(0.500)	(2.14)	(0.500)	(0.498)	(3.00)
Years of Father Edu.	6.638	7.726	1.088***	5.441	5.670	0.229**	5.675	6.013	0.338***
	(4.609)	(4.259)	(6.26)	(4.474)	(4.504)	(2.78)	(4.526)	(4.529)	(4.49)
Years of Mother Edu.	5.287	6.070	0.783***	3.743	3.932	0.189*	4.044	4.291	0.246***
	(4.526)	(4.510)	(4.44)	(4.156)	(4.225)	(2.46)	(4.275)	(4.348)	(3.45)
Observations	1435	1305	2740	5893	6385	12278	7328	7690	15018

Notes: The first three columns show the gender gap in education and other personal characteristics in families with only sons or daughters. The next three columns show the summary statistics for families with both sons and daughters. The last three columns show the summary statistics for all families. * indicates that the difference is significant at the 10% significance level; ** indicates that the difference is significant at the 5% significance level; *** indicates that the difference is significant at the 1% significance level. For the “Male" and “Female” columns, standard deviations are reported in the parentheses; for the “Diff” columns, t stat. are reported in the parentheses.

Table 2: Gender Gap in Education and Personal Characteristics: Rural Sample

	Same-sex			Mixed-sex			Total		
	Male	Female	Diff.	Male	Female	Diff.	Male	Female	Diff.
Years of Edu.	8.469	8.366	-0.102	8.311	7.155	-1.156***	8.343	7.257	-1.086***
	(2.879)	(3.076)	(-0.98)	(2.995)	(3.386)	(-27.75)	(2.972)	(3.378)	(-28.44)
Age	45.50	42.74	-2.764***	47.51	47.07	-0.436***	47.09	46.71	-0.387***
	(9.180)	(9.232)	(-8.41)	(7.888)	(7.956)	(-4.22)	(8.210)	(8.160)	(-3.91)
No. siblings	2.178	2.310	0.132**	3.741	3.716	-0.0250	3.419	3.598	0.179***
	(1.369)	(1.522)	(2.61)	(1.532)	(1.518)	(-1.26)	(1.627)	(1.567)	(9.24)
Birth order	1.999	2.116	0.118**	2.764	2.776	0.0128	2.606	2.721	0.115***
	(1.180)	(1.295)	(2.72)	(1.549)	(1.577)	(0.63)	(1.512)	(1.566)	(6.18)
No. male sib.	3.137	0	-3.137***	2.789	2.137	-0.652***	2.860	1.957	-0.903***
	(1.361)	(0)	(-74.32)	(1.254)	(1.101)	(-42.23)	(1.285)	(1.209)	(-59.74)
No. female sib.	0	3.229	3.229***	1.974	2.599	0.625***	1.568	2.652	1.085***
	(0)	(1.495)	(121.42)	(1.040)	(1.205)	(42.65)	(1.223)	(1.245)	(72.80)
Father attend school	0.542	0.563	0.0215	0.479	0.486	0.00745	0.492	0.492	0.000759
	(0.498)	(0.496)	(1.17)	(0.500)	(0.500)	(1.13)	(0.500)	(0.500)	(0.12)
Mother attend school	0.356	0.407	0.0509**	0.268	0.274	0.00652	0.286	0.285	-0.00085
	(0.479)	(0.491)	(2.87)	(0.443)	(0.446)	(1.11)	(0.452)	(0.451)	(-0.15)
Years of Father Edu.	3.895	4.023	0.128	3.352	3.416	0.0637	3.464	3.465	0.00133
	(3.868)	(3.799)	(0.90)	(3.732)	(3.754)	(1.28)	(3.767)	(3.761)	(0.03)
Years of Mother Edu.	2.398	2.832	0.434***	1.762	1.804	0.0420	1.892	1.887	-0.00574
	(3.341)	(3.633)	(3.45)	(3.005)	(3.021)	(1.05)	(3.088)	(3.087)	(-0.15)
Observations	3158	1040	4198	12172	11333	23505	15330	12373	27703

Notes: The first three columns show the gender gap in education and other personal characteristics in families with only sons or daughters. The next three columns show the summary statistics for families with both sons and daughters. The last three columns show the summary statistics for all families. * indicates that the difference is significant at the 10% significance level; ** indicates that the difference is significant at the 5% significance level; *** indicates that the difference is significant at the 1% significance level. For the “Male" and “Female” columns, standard deviations are reported in the parentheses; for the “Diff” columns, t stat. are reported in the parentheses.

Table 3: The Effect of Sibling Rivalry on the Gender Gap in Education

	Urban				Rural			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Same-sex	Mixed-sex	All	All	Same-sex	Mixed-sex	All	All
Female	-0.0180	-0.491***	-0.428***	-0.0458	-0.445***	-1.211***	-1.121***	-0.412***
	(0.159)	(0.0530)	(0.0511)	(0.159)	(0.142)	(0.0411)	(0.0397)	(0.136)
Mixed-sex				0.0490				0.247***
				(0.120)				(0.0794)
Female×Mixed-sex				-0.446***				-0.798***
				(0.168)				(0.142)
Birth order	0.0800	0.154***	0.146***	0.146***	0.0598	0.0699***	0.0675***	0.0701***
	(0.0797)	(0.0267)	(0.0251)	(0.0252)	(0.0481)	(0.0179)	(0.0167)	(0.0167)
Sibling size	-0.418***	-0.426***	-0.430***	-0.421***	0.0680	-0.162***	-0.130***	-0.137***
	(0.0957)	(0.0373)	(0.0338)	(0.0347)	(0.0581)	(0.0270)	(0.0239)	(0.0248)
Age	0.0185	-0.0747**	-0.0634**	-0.0596*	-0.122**	0.0509*	0.0216	0.0234
	(0.0717)	(0.0355)	(0.0316)	(0.0316)	(0.0514)	(0.0262)	(0.0232)	(0.0231)
Age squared	-0.0967	0.0220	0.00725	0.00372	0.0516	-0.138***	-0.107***	-0.108***
			(0.168)				(0.142)	
Observations	2174	12266	14440	14440	3900	23490	27390	27390
R^2	0.257	0.240	0.248	0.249	0.147	0.167	0.165	0.166

Notes: The dependent variable is the year of education. Columns 1 uses the same-sex families and column 2 uses the mixed-sex families in the urban sample. Columns 3 and 4 include all observations in the urban sample. Columns 5 and 6 use same-sex families and mixed-sex families in the rural sample, while columns 7 and 8 include all observations in the rural sample. All the regressions control for parents' education level and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4： Robustness: Controlling for the Change of Hukou Status

	(1) Same-sex	(2) Mixed sex	(3) All	(4) All
Female	0.0346	-0.352***	-0.300***	-0.00593
	(0.182)	(0.0601)	(0.0585)	(0.182)
Mixed-sex				0.0161
				(0.139)
Female×Mixed-sex				-0.348*
				(0.192)
Birth order	0.219**	0.147***	0.150***	0.150***
	(0.0926)	(0.0317)	(0.0298)	(0.0298)
Sibling size	-0.512***	-0.444***	-0.448***	-0.440***
	(0.109)	(0.0482)	(0.0433)	(0.0441)
Age	0.113	-0.00643	0.0129	0.0154
	(0.0856)	(0.0433)	(0.0385)	(0.0386)
Age squared	-0.179*	-0.0352	-0.0587	-0.0609
	(0.0984)	(0.0481)	(0.0432)	(0.0432)
Observations	1488	7931	9419	9419
R^2	0.299	0.243	0.254	0.255

Notes: The dependent variable is the year of education. Column 1 uses the same-sex families and column 2 uses the mixed-sex families in the urban sample. Columns 3 and 4 include all observations in the urban sample. All the regressions control for parents' education level and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 5: Robustness: Controlling for Sibling Size

	Urban				Rural			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sibsize=2	Sibsize=3	Sibsize=4	Sibsize=5	Sibsize=2	Sibsize=3	Sibsize=4	Sibsize=5
Female×Mixed-sex	-0.180	-0.202	-0.303	-0.585	-0.657^{***}	-0.0464	-0.574^{*}	-1.294^{***}
	(0.239)	(0.330)	(0.496)	(0.618)	(0.234)	(0.256)	(0.336)	(0.455)
Observations	2247	3657	3280	2607	2669	4999	6249	5987
R^2	0.244	0.219	0.175	0.172	0.258	0.187	0.130	0.131

Notes: This table presents estimates of the gender gap in education induced by intrahousehold discrimination using households with different numbers of siblings. In each column, we control for birth order, age, age squared, parents' education level, gender dummy variable, whether the person is from mixed-sex families, and provincial fixed effect. Robust standard errors are clustered at the household level. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

Table 6: Mother's Education and the Intrahousehold Discrimination

	Urban		Rural	
	(1)	(2)	(3)	(4)
	Never School	Any School	Never School	Any School
Female×Mixed-sex	-0.791***	-0.208	-1.016***	-0.337
	(0.296)	(0.213)	(0.195)	(0.214)
Observations	6698	7314	19088	7560
R^2	0.169	0.187	0.142	0.123

Notes: This table presents estimates of the measure of the gender gap in education due to intrahousehold discrimination when mothers received no education and when mothers attended any school. In each column, we control for birth order, number of siblings, age, age squared, father's education level, gender dummy variable, whether the person is from mixed-sex families, and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Father's Education and the Intrahousehold Discrimination

	Urban		Rural	
	(1)	(2)	(3)	(4)
	Never School	Any School	Never School	Any School
Female×Mixed-sex	-0.762*	-0.279	-1.081***	-0.485**
	(0.426)	(0.192)	(0.214)	(0.194)
Observations	4310	9648	13541	13071
R^2	0.116	0.177	0.132	0.118

Notes: This table presents estimates of the measure of the gender gap in education due to intrahousehold discrimination when fathers received no education and when fathers attended any school. In each column, we control for birth order, number of siblings, age, age squared, father's education level, gender dummy variable, whether the person is from mixed-sex families, and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Firstborn Son Effect and Firstborn Daughter Effect

	Urban		Rural	
	(1)	(2)	(3)	(4)
	Firstborn Son	Firstborn Daughter	Firstborn Son	Firstborn Daughter
Female×Mixed-sex	-0.592***	-0.250	-1.028***	-0.746***
	(0.199)	(0.199)	(0.165)	(0.166)
Observations	5988	6100	11049	10806
R^2	0.270	0.253	0.163	0.161

Notes: This table presents estimates of the gender gap in education due to intrahousehold discrimination for firstborn sons and daughters. In each column, we control for birth order, number of siblings, age, age squared, parents' education, gender dummy variable, whether the person is from mixed-sex families, and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: The Gender Gap in the School Dropout Rate

	Urban			Rural		
	No school	Primary school	Middle school	No school	Primary school	Middle school
Female	3.1%	12.2%	42.6%	11.9%	40.2%	80.5%
Male	1.7%	9.7%	42.9%	5.2%	28.1%	76.0%

Notes: This table shows the school dropout rate for urban and rural regions.

Table 10: Intrahousehold Discrimination and School Dropout

	Urban			Rural		
	(1)	(2)	(3)	(4)	(5)	(6)
	No School	Primary school	Middle school	No School	Primary school	Middle school
Female×Mixed-sex	0.0138**	-0.00577	0.0919***	0.0545***	0.0622***	0.0659***
	(0.00595)	(0.0149)	(0.0271)	(0.0100)	(0.0240)	(0.0239)
Obs.	14440	14077	12481	27390	25133	16739
R^2	0.054	0.101	0.124	0.104	0.111	0.043

Notes: This table presents the results of the gender gap in education using school dropout as the dependent variable. In each column, we control for birth order, number of siblings, age, age squared, parents' education, gender dummy variable, whether the person is from mixed-sex families, and provincial fixed effects. Robust standard errors are clustered at the household level. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figures

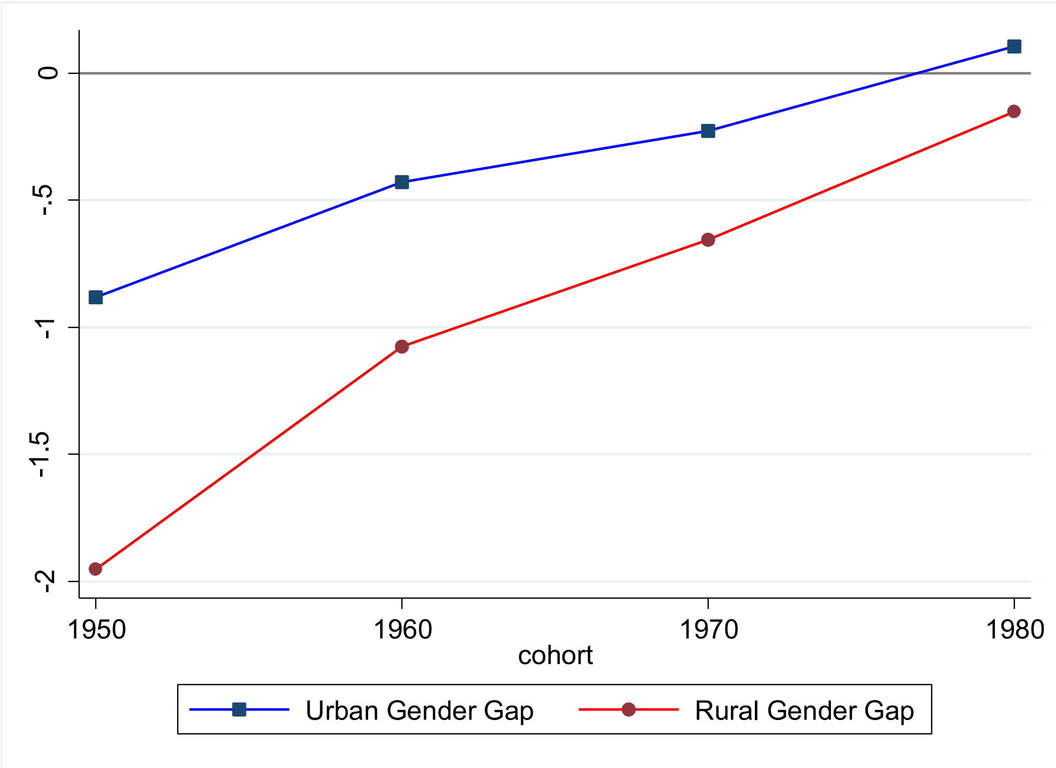

Figure 1: Trends of the Gender Gap in Education in Rural and Urban Regions

Notes: This figure draws the trends of the gender gap in education in urban and rural regions. The point estimates are obtained by running regression equation (2).

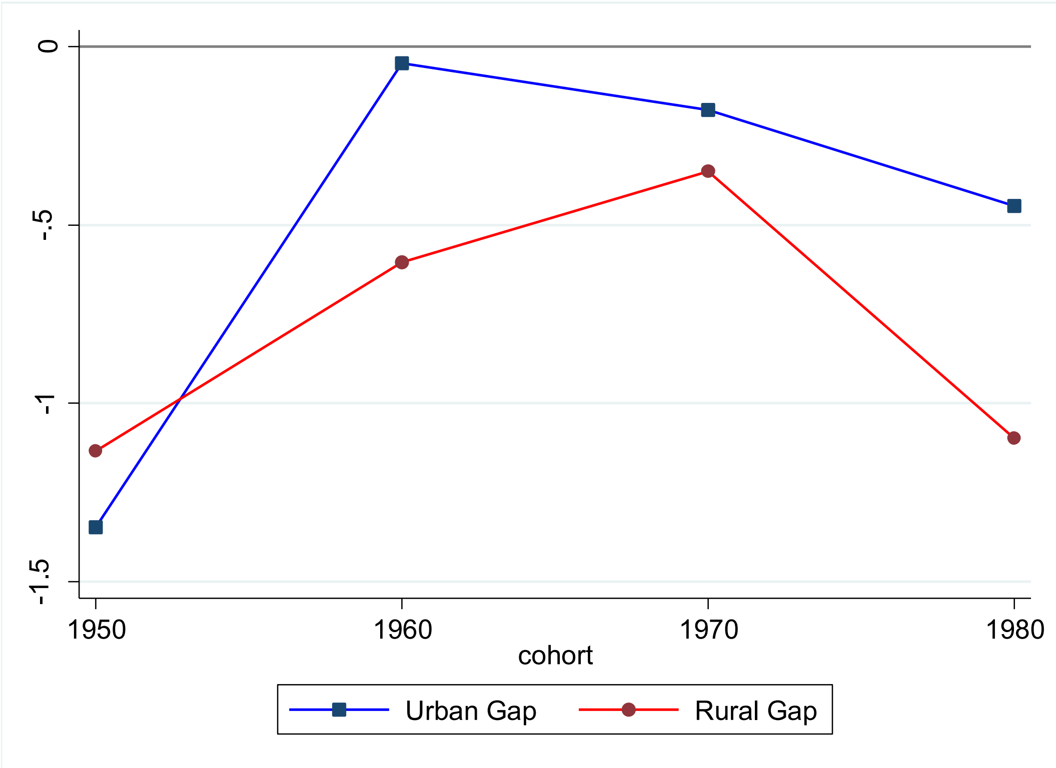

Figure 2: Trends of the Intrahousehold Part of the Gender Gap in Education in Rural and Urban Regions

Notes: This figure draws the trends of our measures of the gender gap in education due to intrahousehold discrimination in both the urban region and the rural region. The point estimates are obtained from regression equation (1).

Appendix A: Son Preference and the Share of Mixed-sex Families

If there is no son preference and sex selection of children, the share of people from mixed-sex families should follow some arithmetic rule. Suppose male and female infants are born with an equal chance and there is no postnatal selection such as abandoning female infants, mixed-sex families should account for 50% for families with 2 children, and 75% for families with 3 children, and so on. The theoretical share of mixed-sex families is shown in the last column of Table A.1.

However, son preference is indeed a concern in China, and it affects the share of mixed-sex families in two ways. The first is male-biased fertility stopping rules. Suppose parents prefer sons, and they would like to have one more child if they do not yet have sons. For two-children families, those with two girls will choose to have one more child and thus they leave the two-children families. This will decrease the share of all-daughter families and makes the share of mixed-sex families to be higher than the theoretical share for two-children families. Note that the male-stopping rule does not change the overall male-female ratio because there is no selection of which gender to be born. It only changes the distribution of mixed-sex families and all-daughter families. It makes the mixed-sex families more concentrated and all-daughter families less concentrated in families with low numbers of children.

The second is sex selection through prenatal or postnatal selections (Babiarz et al., 2018). Given that ultrasound detection was not readily available for most of the individuals in our sample period and sex detection was prohibited during the one-child policy period in China, most of the sex selections are through postnatal selections. Suppose parents prefer sons and sex selection are against girls, it will lead to a biased male-female ratio. The share of all-daughter families will be lower than the theoretical share and the share of all-son families will be higher for each sibling size. But its effect on the share of mixed-sex families is indeterminate.

Table A.1 shows shares of people from each type of family for families with one to eight children in both the urban region and the rural region. In the urban region, the shares of people from all-son families and all-daughter families are close, and all-son families on average have a smaller sibling size. For families with 2 to 4 children, the mixed-sex accounts for a larger share than the theory predicts. For families with 5 to 8 children, the share of mixed-sex families is lower than the theoretical share. The results are consistent with the predictions of the male stopping rule. In the rural region, the share of observations from all-son families is about three times the share from all-daughter families, and the share from all-son families is higher for each sibling size. For families with 2, 3, and 6 children, the share from mixed-sex families is higher than expected. For families with 5, 7, and 8 children, the share from mixed-sex families is lower than the theoretical share. For families with 4 children, the share from mixed-sex families is equal to the theoretical share. The results in the rural region indicate that both the male-stopping rule and sex selection have affected the gender composition of children.

Table A1: Share of Mixed-Sex Families and Theoretical Predictions (%)

	Urban			Rural			Theory
Sibling size	All-son	All-daughter	Mixed	All-son	All-daughter	Mixed	Mixed Share
1	52.3	47.7	0.0	70.2	29.8	0.0	0.0
2	21.9	20.3	57.8	36.5	11.9	51.6	50.0
3	9.8	7.8	82.4	17.8	4.8	77.5	75.0
4	5.2	4.1	90.7	9.8	2.7	87.5	87.5
5	3.4	3.6	93.0	5.3	2.6	92.1	93.8
6	1.5	2.1	96.3	2.0	1.0	97.0	96.9
7	0.2	3.0	96.8	1.6	0.9	97.4	98.4
8	1.0	2.2	96.8	3.4	0.7	95.9	99.2
Total	9.6	8.7	81.8	11.4	3.8	84.8	

Notes: This table presents the share of people from all-son, all-daughter and mixed-sex families in our sample and also the theoretical predictions if son preference does not exist and girls and boys are born with an equal probability. All numbers are in percentage.

Appendix B: Parental Education and the Share of Daughters in the Family

Table 1 shows that girls in urban regions generally grow up in families with higher parental education. It implies that more educated parents have a lower level of son preference and are less likely to select their children' gender. We use the urban sample to investigate whether high educated parents tend to have more daughters compared to son. In Table B1, we calculate the share of daughters in families separated by both father's and mother's education levels. From the first column, we can see that families with a higher level of father's education also have a higher share of daughters. The same pattern can also be found by mother's education. Thus, the data supports that highly educated people have a low level of son preference and are more likely to have daughters.

Table B1: Parental Education and the Share of Daughters in the Family in the Urban Region

	Father	**Mother**
Never schooled	0.48	0.49
Less than high school	0.51	0.52
High school	0.53	0.52
College and above	0.55	0.59

Note: The sample is the same as in Table 1 of the main text.

Appendix C: Son Preference and Old-age Support in Rural and Urban Regions

In Section 5 of the paper, we state that "*one important reason why the effect of intrahousehold discrimination does not diminish with economic development in rural China is that people rely on sons for old-age support*". While we do not have information on financial support from children to parents in our dataset, we can look at co-residence patterns of old adults to see whether sons or daughters provide more supports to their parents. If the elderly people live with sons, it is probably that the sons take the most responsibility of old-age support. If the elderly people live with daughters, then we can infer that the daughters give more support to their parents. Table C1 shows the co-residence patterns for the old (aged over 60). For the rural sample, about 60% of old adults live with children, and for those who live with children, only 9.7% of them live with daughters. The striking result shows that "relying on sons" still is a dominant way for old-age support in rural China. In the urban region, only 43.2% of the old adults live with children. For those who live with children, although more people choose to live with their sons, the share is less skewed towards sons as in the rural region.

Table C1: Co-residence Patterns of Elderly People

	Son	**Daughter**	**Neither**
Rural	54.6%	5.8%	39.6%
Urban	31.4%	11.8%	56.8%

Note: The sample consist of observations aged over 60 in CHIP 2013. The column "son" reports the percentage of old adults living with sons. "Daughter" reports the percentage living with daughters. "Neither" reports the percentage of old adults who live neither with sons nor daughters.