Social Norms and The Impact of Early Life Events on Gender Inequality[*]

Wei Luo Wei Huang Albert Park

Abstract

We study the influence of social norms in shaping the impact of early life exposure to China's Great Famine on gender inequality. We model how social norms interact with adverse shocks to affect male and female survival chances and influence subsequent human capital investments. We test these predictions empirically by using the 2000 China Population Census that has information on birthplace and estimate a difference-in-differences model that combines cohort and regional variation in exposure to the famine with regional variation in the culture of son preference. We find that son preference buffers the negative impact of intrauterine famine shocks on cohort male-to-female sex ratios but actually reduces famine's impact on gender inequality in health and education.

Keywords: Famine, Son preference, Sex ratios, Human capital investment

JEL Codes: J13, J16, I24, I26

[*] Luo: The University of Hong Kong. Email: weiluo.natalie@outlook.com. Huang: Emory University and IZA. Email: huangw@nber.org. Park: The Hong Kong University of Science and Technology, IZA, and CEPR. Email: albertpark@ust.hk. We thank Douglas Almond, Li Han, James Kai-sing Kung, Katherine Meckel, Adriana Lleras-Muney, Jin Wang, Lingwei Wu, and Jane Zhang, as well as seminar participants at UCLA, HKUST, Royal Economic Society Annual Meeting, and EEA Annual Congress for their helpful comments. All remaining errors are our own. This paper was previously titled "Early life adverse shocks, son preference and gender inequalities in later life".

1 Introduction

Previous literature has provided well-documented evidence on the long-run impact of prenatal events (Barker, 1992; Case et al., 2005; Almond, 2006; Maccini and Yang, 2009; Almond and Currie, 2011; Currie and Vogl, 2013; Nilsson, 2017). This link often is not gender-neutral, but empirical findings are mixed on which gender is more affected by prenatal events. Both biological and economic studies show that male fetuses and babies are more vulnerable to early life events (Kraemer, 2000; Cameron, 2004; Almond and Mazumder, 2011; Dinkelman, 2017; Nilsson, 2017), while others find that girls are more affected (Pathania, 2007; Maccini and Yang, 2009; Shi, 2011). Of interest, most research which finds girls are more responsive to early life events are in Asian settings (Pathania, 2007; Maccini and Yang, 2009; Shi, 2011; Cui et al., 2020).

Abundant literature provides evidence that cultural preferences favoring sons contribute to gender disparities in health outcomes and mortality rates in Asian countries, such as India and China (Gupta, 1987; Chen et al., 2007; Bhalotra et al., 2010; Jayachandran and Kuziemko, 2011; Jayachandran and Pande, 2017). Therefore, comparing gender differences in the long-run impact of prenatal events independent of the cultural setting may be problematic. Emerging literature points out that gender differences in the impact of early life events are context-specific. For example, Dinkelman (2017) mentioned that her finding that local environmental shocks have a more negative impact on males in South Africa is quite different from most findings in Asian settings, and it seems likely that differences in son preference across these continents could contribute to these differences. Besides, a growing literature recognizes that shocks, investments, and interventions can interact in complex ways (Almond et al., 2018; Duque et al., 2018). In this paper, we study the extent to which social norms influence gender disparities in the long-run impact of early life events.

We study the long-run impact of early life events and their interaction with social norms in the context of perhaps the most severe famine in human history, China's Great Famine of 1959-1961, which struck provinces in China unexpectedly and with large regional variation in death rates. During 1959-1961, agricultural production dropped sharply, and estimated daily available food energy fell below the minimum food energy requirement (Ashton et al., 1992; Lin and Yang, 2000). The prolonged famine caused an unprecedented number of deaths. National death rates were 14.6, 25.4, and 14.2 per thousand in 1959, 1960, and 1961, compared to 11.4, 10.8, and 12.0 per thousand in the previous three years (1956-1958).[1] The Chinese central government eventually recognized the severity of the famine and moderated its policies, including reducing the transfer of grain from rural areas to urban areas and sending millions of people back to the countryside to boost agricultural production (Li and Yang, 2005). By the end of 1961, death rates began to return to the pre-1959 level in over half of the provinces, and birth rates started to rebound.

Although China's Great Famine was national in scope, the famine intensity varied significantly across provinces (Ashton et al., 1992; Lin and Yang, 2000; Chen and Zhou, 2007; Meng et al., 2015). Figure 1 displays the geographical variation of famine intensity by mapping the provincial death rate in 1960, the worst year of the famine in terms of fatalities. Provinces in central China, such as Henan and Anhui, and some southwestern provinces, were severely affected by the famine, while the northeastern provinces were less severely impacted. In a word, the Great Famine generated plausibly exogenous adverse shocks of varying magnitude to people living in different regions.

According to rich literature, the culture of son preference is deeply rooted and persistent in China (Das Gupta et al., 2003; Guilmoto, 2009). This tradition emphasizes the importance of continuing the family line through the male offspring, thereby reinforcing male dominance within a household (Murphy et al., 2011). The culture of son preference has profoundly shaped childbearing and child-rearing behavior in China (Yi et al., 1993; Chen et al., 2007). One manifestation of son preference is sex-selection practices, which may be performed either prenatally or postnatally. For example, female infanticide, the neglect of baby girls, and the preferential allocation of household resources to sons all can be categorized as postnatal sex-selection strategies.

The cultural preference for sons is not only deeply rooted in the history of China but also characterized by obvious and remarkable regional differences. Male-to-female sex ratios of newborn babies or young cohorts are often used as a proxy for the culture of son preference (Arnold and Liu, 1986; Park and Cho, 1995; Edlund, 1999; Jayachandran and Kuziemko, 2011). In Figure 2, we describe the geographic distribution of male-to-female sex ratios of cohorts aged 0-10 in the 1953 China Population Census. The figure shows that southern provinces, such as Guangdong and Fujian, and provinces in central China, such as Anhui and Jiangxi, have higher levels of son

[1] The national death rate during the famine period is from the Statistical Yearbook published by the National Bureau of Statistics of China.

preference. Comparing this map with the geographic distribution of famine intensity reveals that the two are not significantly spatially correlated. In fact, the correlation coefficient between the provincial death rate in 1960 and sex ratios (males to females) of cohorts aged 0-10 in the 1953 China population census is only 0.116.

In this paper, we develop a unified model to show that: (1) Adverse circumstances, such as famine, negatively influence individuals' survival chances by increasing the threshold to survive. (2) Gender-specific postnatal investments (caused by son preference) buffer the negative impact of adverse shocks (by lowering the threshold to survive) and influence male and female chances differently. (3) Selective mortality caused by son preference combined with famine shocks alters the gender gap in the expected health of survivors; (4) If we assume that human capital investment is complementary to health, the gender gap in education among survivors is changed. Specifically, we argue that parents' unequal allocation of resources across children (as a result of son preference), especially under adverse circumstances, increases the relative survival chances of males versus females (reducing sex ratios), which further mitigates the effects of famine shocks on the gender gaps in health and educational attainment.

To test our predictions about the long-run consequences of intrauterine famine shock and its interacting effect with gender norms, we assemble data from several sources. We collect cohort information on gender inequalities from the 2000 China Population Census and the 2010 China Family Panel Survey. The advantage of using the 2000 China Population Census is that it provides information on individuals' birthplace, enabling us to precisely identify famine severity received in utero and alleviate the concern about migration. In addition, we gather detailed province-level death rates from the Comprehensive Statistical Data and Materials on 50 Years of New China which is compiled by the National Bureau of Statistics of China. At last, we collect information on the regional culture of son preference from pre-famine Census data (the 1953 China Population Census). Our analysis adopts a generalized difference-in-differences approach to estimate the causal linkages. That is, our identification strategy is based on regional and cohort variation in famine severity received when in utero together with pre-famine regional variation in the culture of son preference.

We find that even though intrauterine famine shocks contribute to a reduction in male-to-female sex ratios in gender-neutral areas, the reduction in sex ratios induced by the famine shock in provinces with son preference is much smaller compared to that in gender-neutral provinces, which is consistent with our theoretical predictions on survival chances. The difference in the impact of famine on sex ratios in areas with son preference and gender-neutral areas associated with a one standard deviation increase in famine severity accounts for 1.89% of the mean cohort sex ratio (or 15.26% of the standard deviation). Moreover, we observe that intrauterine famine shocks increase the gender gap in height in gender-neutral areas. However, the gender gap in height generated by the famine shock is narrower in areas with son preference, consistent with our theoretical predictions. Finally, we show that intrauterine famine shocks increase the gender gap in years of education in gender-neutral areas, but this is less true in areas with a culture of son preference. The change in the gender gap in years of education in response to a one standard deviation increase in famine severity in areas with son preference accounts for 6.41% of the mean (or 9.46% of the standard deviation). We further utilize birth month to capture the timing of pregnancy and show our results are not driven by selective fertility in response to the famine. Our analysis is robust to adding rich pre-famine regional controls. It passes several robustness tests, including using alternative sample restrictions, alternative measures of regional son preference, and alternative measures of famine exposure, controlling for the impacts of the Chinese Civil War, the Cultural Revolution, and famine exposure received the first year of life, and a placebo test.

Our research contributes to the literature on the long-run impact of early life events, especially empirical tests of the fetal origins hypothesis.[2] As noted earlier, prior research has emphasized the lasting impact of early life events, and there is well-documented evidence that the effect of prenatal events plays out differently by gender. However, much less known is about the extent to which this long-run effect of early life events is mitigated or exacerbated by social norms (especially gender norms). This paper thus complements, and adds to, the literature on the persisting impact of early life events by linking early life adverse shocks to social norms (in our case, the cultural preference for sons).

Our paper also relates to the literature on the lasting impact of China's Great Famine. Previous research has documented significant negative effects of early life exposure to famine on adult outcomes, in terms of educational attainment, labor market outcomes, and health (Luo et al., 2006; Almond et al., 2007; Chen and Zhou, 2007; Meng and Qian, 2009; Shi, 2011; Mu and Zhang, 2011; Cui et al., 2020). Even though prior work shows that the impact of early life famine exposure on

[2] For a general review of early life circumstances and adult outcomes, see Currie and Vogl (2013) and Almond et al. (2018).

some adult outcomes differs by gender (Almond et al., 2007; Shi, 2011; Mu and Zhang, 2011; Cui et al., 2020),[3] few papers systematically investigate the impact of early life famine exposure on gender inequality (Mu and Zhang, 2011).[4] Our paper fills the gap by presenting causal evidence that intrauterine famine exposure shapes gender inequalities in survival chances, health outcomes, and educational attainment in later life. Furthermore, to the best of our knowledge, this is the first study to relate the impact of famine exposure on gender inequalities to the preference for sons and quantify the role of son preference in mitigating/exacerbating the long-run impact of the famine on gender gaps in China.

Our contribution can also be situated relative to the literature on son preference and the "missing girls" phenomenon (Gupta, 1987; Sen, 1990; Yi et al., 1993; Das Gupta et al., 2003; Ebenstein, 2010). Our findings extend the literature on the impact of son preference on gender disparities in mortality (Chen et al., 2007; Bhalotra et al., 2010) and in health outcomes (Jayachandran and Pande, 2017) by showing that this impact is exacerbated by adverse shocks and results in unintended consequences for later-life gender inequality.

The rest of this paper is structured as follows. The next section presents the theoretical framework. Section 3 describes the data and variables. Section 4 introduces our identification strategy. In section 5, we present our empirical findings. Section 6 presents robustness checks. Finally, we conclude in section 7.

2 Theoretical framework

2.1 Overview

We develop a simple model to show that adverse shocks negatively influence individuals' survival chances. Son preference mitigates the effect of adverse shocks on male survival chances (leading to gender asymmetry in survival chances) and alters the effect of adverse shocks on gender gaps in health and educational attainment in later life. Our model differentiates the intrauterine period

[3] Their findings are mixed. For example, Almond et al. (2007) find that fetal famine exposure reduced the sex ratio (males to females). Meanwhile, Shi (2011) finds no significant impact of famine intensity experienced in the first year of life on the proportion of women in their sample.

[4] Mu and Zhang (2011) use post-famine ethnic group variations in the culture of son preference to show the observed gender difference in the impact of famine on illiteracy rate is probably better explained by the culture of son preference.

and the postnatal period. Children are exposed to famine shocks in utero. As sex selection technology is not available during that period,[5] parents make gender differential investments only after observing the child's sex (postnatal period).

Our theoretical predictions are based on the position of survival thresholds relative to health distributions. Therefore, modeling gender differences as different health distributions relative to a common threshold or as different survival thresholds relative to a common health distribution will not alter our theoretical predictions. To make our analysis simple, we make two assumptions here: 1) health distributions of boys lie to the left of that of girls (Catalano and Bruckner, 2006; Mu and Zhang, 2011);[6] 2) famine increases the survival thresholds of both boys and girls.[7]

2.2 Setup of the model

Initial endowment distributions. Initial health endowments of each gender i of each cohort are assumed to be normally distributed.[8] We use subscripts to denote boys (b) and girls (g). The distribution of boys' endowments lies slightly to the left of that of girls,[9] $\mu_b < \mu_g$. For simplicity, we assume that the standard deviations of initial endowment draws are the same for boys and girls, or $\sigma_b = \sigma_g = \sigma$.

Postnatal investment and son preference. Gender-specific postnatal investment is purely determined by son preference θ_i and assumed to decrease survival thresholds. We use superscripts n and p to denote two social norms-gender-neutral (n) and son preference (p). In gender-neutral

[5] Ultrasound did not become prevalent in China until the 1970s (Chen et al., 2013).

[6] Evidence shows that the infant mortality of boys is higher than girls. This phenomenon can be modeled in two ways: 1) the health endowment of boys and girls are different (weaker boys), and the survival threshold is uniform; or 2) health distributions are uniform, but boys have a higher survival threshold. Previous studies in biology and economics document that females may have better health endowments than males (Catalano and Bruckner, 2006; Mu and Zhang, 2011).

[7] Valente (2015) provides a comprehensive discussion about how adverse circumstances influence mortality rate. It may reduce health parameters (scarring effect) or increase the threshold to survive (culling effect). The main difference is that the scarring effect may not improve health among survivors (Almond, 2006), but the culling effect will improve health endowment.

[8] Rich empirical evidence in economics and biology supports the normal distribution of human height and weight (Tanner and Tanner, 1981; Wachter and Trussell, 1982; Steckel, 1995).

[9] It is consistent with the biological literature that finds that boys are naturally weaker and more vulnerable than girls (Andersson and Bergstrom, 1998; Kraemer, 2000).

areas, parents treat boys and girls equally ($\theta_b^n = \theta_g^n$). In areas with son preference, parents care more about boys than girls ($\theta_b^p > \theta_g^p$).[10]

Famine shocks. Intrauterine exposure to famine shocks is assumed to increase survival thresholds (Δμ) of both genders.

Survival chances. Individuals survive if the endowment draw is high enough given the survival threshold Z (equation 1), where Φ(.) denotes the standard normal cumulative distribution function (henceforth, CDF). $Z = z + \Delta\mu - \theta_i$, where z is the survival threshold under natural conditions (uniform to both genders). We define $z_i = z - \theta_i$ as gender-specific survival thresholds accounting for postnatal investments.[11]

$$Sr(E_i > Z) = 1 - \Phi\left(\frac{z - \theta_i + \Delta\mu - \mu_i}{\sigma}\right) \quad (1)$$

Observed health of survivors. Survivors' expected health can be viewed as the expectation of a normal distribution of health truncated by the survival threshold (equation 2). We use $\lambda(.)$ to denote the inverse Mills ratio, representing the selection effect.

$$Q(E_i > Z) = \mu_i + \sigma\lambda\left(\frac{z - \theta_i + \Delta\mu - \mu_i}{\sigma}\right) \quad (2)$$

Schooling choice of survivors. In a two-child family (one boy and one girl),[12] parents choose children's schooling, consumption, and savings to maximize their utility (equation 3) subject to the life-time budget constraint (equation 4) and the determinants of children's earnings (equation 5). Parents' utility depends on their consumption in both periods, C_1 and C_2, and the utility getting from their children's income, where β is the discount factor and θ is the degree of altruism

[10] Although we recognize that health endowments will influence health investments as well as educational investments, we do not include health investments in the model explicitly because doing so would not change the qualitative predictions of the model as long as gender differences in the negative impact of the famine on health endowments are not fully offset by compensating health investments, but adding health investments would make the model much less tractable.

[11] The survival probabilities for non-exposed cohorts can be obtained by equating $\Delta\mu$ to 0.

[12] This model is also suitable to a single-child household.

(preference) toward children.[13] Parental resources equal Y (earnings in period 1) as we assume that children do not transfer resources to their parents.[14] Parental consumption in period 2 is determined by the budget constraint $C_2 = RX_1$, where R denotes the rate of return on X_1 and is assumed to be exogenously determined by the competitive market. X_1 is saving in period 1 and p is the price of per unit schooling. We model children's earnings as the product of human capital that depends on schooling years s_i and children's health Q_i, and w_i (the market-determined wage for a unit of human capital), shown in equation 5. The key assumption is that parents' investment in schooling is complementary to children's health.

$$Max\ E\{\ln C_1 + \beta \ln C_2 + \beta\theta_b I_b + \beta\theta_g I_g\} \tag{3}$$

$$C_1 + \frac{C_2}{R} + ps_b + ps_g = Y \tag{4}$$

$$I_i = w_i Q_i s_i^{\gamma}, 0 < \gamma < 1 \tag{5}$$

2.3 Predictions of the model

2.3.1 Survival chances

The impact of famine on survival chances is the difference in the survival probability of cohorts exposed to the famine and those who were not, equaling the difference between two CDFs (equation 6). The standard normal CDF has the property that $\Phi'(x) > 0$ for any $x < 0$. Intuitively, if x_1 and x_2 are two possible survival thresholds lying on the left tail of the standard normal distribution ($x_1 < x_2$), a slight rightward shift (Δv) of x_2 will generate larger changes in the cumulative distribution function than that of x_1.[15]

[13] As long as parents have some degree of altruism (or preference) toward their children, and $\theta > 0$, they invest positive amount in children's schooling, since the marginal rate of these investments is very high for small investments, which satisfies the Inada conditions.

[14] The budget constraint of parents in period 1 is $ps_b + ps_g + C_1 + X_1$=Y.

[15] This statement is true only for $x_1 + \Delta v < x_2 + \Delta v < 0$. See details in Lemma 1 in the Appendix.

$$Sr(E_i > z_i' + \Delta\mu) - Sr(E_i > z_i') = \Phi\left(\frac{z-\theta_i-\mu_i}{\sigma}\right) - \Phi\left(\frac{z-\theta_i+\Delta\mu-\mu_i}{\sigma}\right) \tag{6}$$

First, we discuss how famine influences survival probabilities in gender-neutral areas. As $\mu_b < \mu_g$, the position of survival threshold relative to boys' health distribution lie right to that of girls (Figure 5). According to the property mentioned above, the impact of famine on survival chance reduction is greater for boys than girls in gender-neutral areas, hence, reducing male-to-female sex ratios.

As the preference for boys is higher in areas with son preference, the boys' survival threshold is smaller in areas with son preference than that in gender-neutral areas ($z'^n_b > z'^p_b$). By the same logic, the famine generates a larger survival probability reduction for boys in gender-neutral areas than in areas with son preference (equation 7). Similarly, the preference for girls is higher in gender-neutral areas, and the girls' survival threshold is greater in areas with son preference compared to gender-neutral areas ($z'^p_g > z'^n_g$). Thereby, the famine will lead to a larger reduction in survival probabilities for girls in areas with son preference than in gender-neutral areas (equation 8).

$$Sr(E_b > z_b'^n) - Sr(E_b > z_b'^n + \Delta\mu) > Sr(E_b > z_b'^p) - Sr(E_b > z_b'^p + \Delta\mu) \tag{7}$$

$$Sr(E_g > z_g'^p) - Sr(E_g > z_g'^p + \Delta\mu) > Sr(E_g > z_g'^n) - Sr(E_g > z_g'^n + \Delta\mu) \tag{8}$$

From these two inequalities, we know that the gender difference in the impact of famine on survival probability reduction is more significant in gender-neutral areas compared to areas with son preference. We thus put forward Proposition 1.

Proposition 1. *The reduction in sex ratios due to the famine is greater in gender-neutral areas than in areas with son preference.*[16]

[16] See details in proof of Proposition 1 in the Appendix.

2.3.2 The expected health of survivors

The impact of famine on survivors' expected health is the difference in expected health of famine-exposed and non-exposed cohorts, determined by the change in the inverse Mills ratios, or the selection effect (equation 9). The inverse Mills ratios have the property that $\lambda''(x) > 0$. Suppose that x_1 and x_2 are two different survival thresholds on the left tail of the distribution ($x_1 < x_2$). A rightward shift of x_2 will generate a greater change in the inverse mills ratios than x_1.

$$Q(E_i > z_i' + \Delta\mu) - Q(E_i > z_i') = \sigma\lambda\left(\frac{z + \Delta\mu - \theta_i - \mu_i}{\sigma}\right) - \sigma\lambda\left(\frac{z - \theta_i - \mu_i}{\sigma}\right) \quad (9)$$

First, we consider these changes in gender-neutral areas. The survival threshold for girls is lower than for boys (($z'^n_g > z'^n_b$). According to the property of the inverse Mills ratios, the famine generates a larger selection effect for boys than girls. That is, famine increases the gender gap in health.

Since $\theta_b^p > \theta_b^n$, the boys' survival line is lower in areas with son preference than in gender-neutral areas ($z'^n_b > z'^p_b$). The famine leads to a larger change in expected health for boys in gender-neutral areas compared to areas with son preference (equation 10). Meanwhile, $\theta_g^n > \theta_g^p$, the girls' survival threshold is greater in areas with son preference than in gender-neutral areas ($z'^p_g > z'^n_g$). The famine causes a larger impact on expected health for girls in areas with son preference compared to in gender-neutral areas (equation 11).

$$Q(E_b > z_b'^n + \Delta\mu) - Q(E_b > z_b'^n) > Q(E_b > z_b'^p + \Delta\mu) - Q(E_b > z_b'^p) \quad (10)$$

$$Q(E_g > z_g'^p + \Delta\mu) - Q(E_g > z_g'^p) > Q(E_g > z_g'^n + \Delta\mu) - Q(E_g > z_g'^n) \quad (11)$$

From the inequalities in equations 10 and 11, we know that the gender difference (boys versus girls) in the impact of the famine on expected health is greater in gender-neutral areas than in areas with son preference. Hence, we put forward Proposition 2.

Proposition 2. *The impact of famine on survivors' gender gap in health is larger in gender-neutral areas compared to areas with son preference.*[17]

2.3.3 Schooling choice for survivors

Solving for the optimal schooling choices for boys and girls based on equations 3-5 yields equations 12 and 13. Here, the Lagrangian multiplier for the lifetime budget constraint is τ.

$$s_b = \left(\frac{\theta_b w_b Q_b \gamma\beta}{p\tau}\right)^{\frac{1}{1-\gamma}} \tag{12}$$

$$s_g = \left(\frac{\theta_g w_g Q_g \gamma\beta}{p\tau}\right)^{\frac{1}{1-\gamma}} \tag{13}$$

If we take the logarithm of schooling, the impact of famine on schooling can be expressed as the difference in schooling between famine-exposed and non-exposed cohorts (equation 14). It equals the effect of famine on the expected health multiplied by the human capital production parameter.

$$\ln s\left(E_i > z_i' + \Delta\mu\right) - \ln s\left(E_i > z_i'\right) = \frac{1}{1-\gamma}\left(\ln Q\left(E_i > z_i' + \Delta\mu\right) - \ln Q\left(E_i > z_i'\right)\right) \tag{14}$$

The gender difference in the impact of the famine on schooling in gender-neutral areas (ds^n) and that in areas with son preference (ds^p) can be defined as follows.

$$ds^n = \frac{1}{1-\gamma}\left(\ln\frac{Q\left(E_b > z_b'^n + \Delta\mu\right)}{Q\left(E_b > z_b'^n\right)} - \ln\frac{Q\left(E_g > z_g'^n + \Delta\mu\right)}{Q\left(E_g > z_g'^n\right)}\right) \tag{15}$$

$$ds^p = \frac{1}{1-\gamma}\left(\ln\frac{Q\left(E_b > z_b'^p + \Delta\mu\right)}{Q\left(E_b > z_b'^p\right)} - \ln\frac{Q\left(E_g > z_g'^p + \Delta\mu\right)}{Q\left(E_g > z_g'^p\right)}\right) \tag{16}$$

[17] See details in proof of Proposition 2 and Lemma 2 in the Appendix.

What we are interested in is the sign of $ds^p - ds^n$. The expected health of non-exposed girls in gender-neutral areas is smaller than in areas with son preference ($Q\left(E_g > z'^{n}_{g}\right) < Q\left(E_g > z'^{p}_{g}\right)$), and the expected health of non-exposed boys in gender-neutral areas is greater than in areas with son preference ($Q\left(E_b > z'^{n}_{b}\right) > Q\left(E_g > z'^{p}_{b}\right)$). Given the inequalities expressed in equations 10 and 11, we can then prove that the sign of $ds^p - ds^n$ is negative. Thus, we put forward Proposition 3.

Proposition 3. *The impact of famine on survivors' gender gap in schooling is larger in gender-neutral areas compared to areas with son preference.*[18]

3 Data and descriptive evidence

3.1 Data and measures

To empirically test our theoretical predictions, we combine data from several sources. First, we test our predictions about gender gaps in survival rates and years of education using data from the 2000 China Population Census. Our main empirical analysis is based on a sample that includes cohorts who were born between 1954 and 1966.[19] The main advantage of using this data set rather than other household data is that the 2000 China Population Census provides a large, representative sample across all 30 provinces, including a large number of persons who experienced the famine. This allows us to focus our analysis on cohorts with famine exposure in regions with different degrees of son preference while retaining a sufficient sample size in each birth cohort and birth province. In addition, the 2000 China Population Census provides information on respondents' birth province and enables us to contract an exact measure of famine severity received in utero. Cross province migration was not very prevalent in 2000. Based on our calculation, the inter-

[18] See detailed proof in the Appendix.

[19] The 2000 China Population Census only has information on current Hukou status and does not know the Hukou status when they were born or in utero. Here we do not divide the sample based on current Hukou status as the change of Hukou type during individuals' life courses is endogenous and highly selective.

provincial migration rate was only 3.04% in 2000 for the whole population and 1.92% for the sample that we use. It also has basic demographic and education information, which can be used to construct two dependent variables of interest, male-female sex ratios and the gender gap in educational attainment. We combine the 2000 China Population Census data with additional data on the provincial death rate by year and pre-famine sex ratios as a measure of son preference from the 1953 China Population Census.[20]

To test our theoretical predictions on health outcomes, we use data from the 2010 China Family Panel Survey (CFPS), a nationally representative dataset of Chinese households and individuals with a wealth of information on health and individual background variables.

Two key explanatory variables in our empirical analysis are intrauterine exposure to famine and regional level of son preference. The construction of famine severity measure follows the approach in Almond et al. (2007) and is shown in equations 17 and 18. Famine severity received when in utero is constructed at province birth year-month level and is defined as the province-level excess death rate in their birth year and the year before weighted by the share of intrauterine months in these two years. EDR_j denotes the average excess death rate during the famine period, which is the difference between the average death rate during the famine (1958-1961) and the average death rate in normal years (1972-1976).

$$FS_{jmt} = [EDR_{jt} \times max(9 - birthm, 0) + EDR_{jt-1} \times min(9, birthm)]/9 \tag{17}$$

$$EDR_{jt} = \begin{cases} EDR_i, & 1957 < t < 1962 \\ 0, & otherwise \end{cases} \tag{18}$$

We construct a proxy for the culture of son preference using the male-to-female sex ratios of young cohorts (aged 0-10) from pre-famine Census data (the 1953 China Population Census). This exercise enables our measure of son preference not to be influenced by shocks to sex ratios caused by the famine. In Figure 3, we plot the cohort sex ratios of individuals aged 0-10 in the 1953 Census against sex ratios of the same age cohort in the 1982 and 1990 Censuses. A significant correlation provides evidence that son preference culture is persistent and also confirms the validity of our son preference measure. We also check whether son preference culture is correlated with regional pre-

[20] It comes from the Comprehensive Statistical Data and Materials on 50 Years of New China which is compiled by the Department of Comprehensive Statistics of the National Bureau of Statistics of China.

famine conditions in Figure A1. We find that the correlation coefficients between son preference measure and regional GDP per capita, number of health facilities, and number of high school students in 1954 are very low. As the biologically normal sex ratio (males to females) is 1.05, we define a dummy variable for son preference to equal 1 if the provincial male-to-female sex ratio of cohorts aged from 0 to 10 in the 1953 China Population Census is above 1.07 following the method used by (Mu and Zhang, 2011).

Pre-famine regional characteristics are included in our analysis. Due to the skewed distributions of these characteristics, we subtract the median values of these variables. We control for GDP per capita, number of health facilities, and number of high school students in 1954 to capture pre-famine economic conditions, development of public health services, and educational institutions.[21] We report the summary statistics of these variables in Table A1 in the Appendix.

3.2 Descriptive evidence

We first provide descriptive evidence on how the cultural preference for sons buffers the negative impact of famine shock on cohort sex ratios (males to females). In Figure 4, we plot male-to-female sex ratios of each birth year cohort by regional famine severity and the degree of preference for sons. It is clear that cohort sex ratios experienced a sharp decrease in 1960. This pattern is consistent with the argument in the biological literature that boys are physically more fragile and thus more likely to die in the face of severe adverse shocks (Kraemer, 2000; Almond et al., 2007). Interestingly, the differences between these patterns in areas with a high degree of historical son preference and areas with a low degree of historical son preference when controlling for famine severity. We find that cohort sex ratios dropped dramatically in areas with lower son preference compared to areas with high son preference.[22] In a society lacking access to prenatal sex selection technology, the culture of son preference would strongly affect the survival chances of young girls and alleviate the impact of adverse shocks on young boys. Taken together, Figure 4 suggests that son preference plays an important role in buffering the impact of the Great Famine on male survival

[21] The information in 1954 is used to construct pre-famine controls is because it is the earliest data available.

[22] The impact of famine on cohort sex ratios is defined as the difference in cohort sex ratios between cohorts born between 1954 and 1957 and those born between 1958 and 1961.

chance (or cohort male-to-female sex ratios). However, this simple comparison in Figure 4 is naive, and we next present more rigorous causal evidence.

4 Identification strategy

4.1 Average treatment effect model

Our identification strategy is based on quasi-random variation in famine severity across provinces and birth cohorts together with regional variation in the culture of son preference. To be specific, we compare outcomes of interest across famine exposed cohorts and non-exposed cohorts, across regions with high famine severity and low famine intensity, and across regions that have a culture of son preference and those that do not, controlling for pre-famine economic conditions, development of the public health system, and development of the education system. To fix ideas, our specification is based on a generalized difference-in-differences model:

$$Y_{jmt} = \alpha_0 + \alpha_1 FS_{jmt} + \alpha_2 FS_{jmt} \times SP_j + +\alpha_3 FS_{jmt} \times X_j + \gamma_j + \delta_{mt} + \epsilon_{jmt} \quad (19)$$

Here, Y_{jmt} denotes the outcome variables of interest, which is the mean outcome for cohorts born in province j in month m of year t. FS_{jmt} measures intrauterine exposure to famine severity, constructed using province-level excess death in one's birth year (t) and the year before ($t-1$) weighted by the share of intrauterine months in these two years.[23] SP_j is a dummy variable that equals 1 if provincial male-to-female sex ratios of cohorts aged 0-10 in the 1953 China Population Census are above 1.07.

We also include pre-famine regional controls to alleviate the concern that famine exposure and local gender inequalities may be correlated with some unobserved factors. For example, if rich provinces (with good economic development) are less likely to suffer from famine and also more gender equal, then our estimated impact of famine shock and its interacting effect with son preference would have a downward bias. X_j denotes a series of pre-famine provincial controls, including provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). As X_j are cross-sectional variables that

[23] Continuous measure of famine intensity/treatment is used and thereby more variation in the data can be captured.

would be absorbed by the province fixed effects, we interact the FS_{jmt} with pre-famine regional characteristics X_j to control for factors that may be correlated with son preference and interact with famine severity to affect outcomes of interest. We control for province fixed effects γ_j and birth year-month fixed effects δ_{mt}. All models are weighted by the sample size of each cohort. We cluster standard errors at the province level.

It is worth noting that results without pre-regional conditions can be interpreted as the estimates for the full sample. In addition, results with pre-famine controls can be interpreted as the impacts when these pre-famine conditions are at the median levels. In equation 19, the coefficient α_1 indicates the effect of famine in gender-neutral areas when pre-famine conditions were at the median levels. The coefficient of interest, α_2, indicates the difference in the impact of famine between areas with son preference and gender-neutral areas when pre-famine conditions were at the median levels. The estimation strategy has all of the advantages and potential pitfalls of the standard DID approach. Province fixed effects control for all time-invariant factors that differ between provinces. Birth year-month fixed effects control for patterns in sex ratios or gender gaps in educational attainment or health that differ by cohorts. Our identification strategy relies on the assumption that there are no other contemporaneous shocks affecting gender gaps in outcome variables in gender-neutral areas and areas with son preference beyond the famine shocks (Bertrand et al., 2004). One potential destructive concern is that fertility decisions may be postponed during the famine, and selective fertility could influence our outcome variables of interest. We will address concerns about selective fertility in detail in the results section.

4.2 Flexible estimates

A fully flexible specification is used to examine pre-trends of outcome variables of interest and to examine the timing of the impact of the famine shock interacted with regional son preference (equation 20). We restrict our sample to cohorts born between 1954 and 1966 in the 2000 China Population Census.

$$Y_{jmt} = \beta_0 + \sum_{t=1955}^{1964} \left(\beta_t \times EDR_j \times SP_j \times Year_t\right) + \sum_{t=1955}^{1964} \left(EDR_j \times X_j \times Year_t\right) + \gamma_j + \delta_{mt} + \epsilon_{jt} \quad (20)$$

Here, $Year_t$ denotes the birth year, which is equal to 1 if individuals are born in the year t. All other variables are defined as before. We interact regional excess death rate during the famine with regional son preference and each of the birth year dummies. The estimated vectors of $\beta_t s$ reveal the correlation between famine severity, regional son preference, and the outcome variables of interest for each cohort. We would expect the estimated $\beta_t s$ to be close to zero and statistically insignificant over time before the famine and for the magnitude of the coefficients on the triple interaction term to be larger for the famine exposed cohorts.

5 Empirical results

5.1 The impact on cohort sex ratios

In Table 1, we investigate the impact of intrauterine famine exposure and its interacting effect with regional son preference on cohort male-to-female sex ratios. We multiply sex ratios by 100 to make it feasible to interpret the estimated coefficients. In columns 1-3, the dependent variable is cohort male-to-female sex ratios at province birth year and month level. Column 1 shows that intrauterine famine exposure reduces sex ratios, but the interaction term between famine exposure and son preference is positive, indicating the negative effect of famine exposure on cohort male-to-female sex ratios is mitigated by regional son preference. Results in column 2 hold to including pre-famine regional controls, which implies these estimates remain robust when all pre-famine regional conditions were at the median levels. Column 2 shows, on average, the difference in the impact of famine on sex ratios in areas with son preference and gender-neutral areas associated with a one standard deviation increase in famine severity is 1.98 percentage points, accounting for 1.89% of the mean cohort sex ratio (or 15.26% of the standard deviation), consistent with Proposition 1.

One potential concern about our identifying assumption is that fertility behavior may be selective during the famine, which will confound our analysis. We address this concern by focusing on those born in months, implying that they were conceived before the famine. To be specific, we divide famine exposed cohorts into two groups: cohorts who were conceived at least 5 months before the famine (cohorts who were born between January 1958 and June 1959) and cohorts who were conceived and born during the famine (cohorts born from July 1959 to the end of famine).

Cohorts who were conceived and born during the famine could be a selected sample affected by fertility response to the famine. We reestimate the model when excluding those conceived and born during the famine in column 3. Results show that the interaction effect between famine exposure and son preference on cohort sex ratios remains robust. We also report estimates using alternative ways to aggregate the dependent variable. Columns 4-5 report the results when cohort sex ratios are aggregated at the province birth year level. The remaining two columns report the results using the moving average of cohort sex ratios (6 months before and after a particular birth year and month). Overall, our results remain robust to alternative ways to aggregate the dependent variable.

We present the estimates for the fully flexible specification in panel A of Figure 6.[24] The point estimates of coefficients of the triple interaction of famine severity, son preference, and birth year in equation 19 are plotted, and a clear pattern emerges. Before the Great Famine, the coefficients were constant over time and small in magnitude, suggesting no substantial effect of future famine severity on cohorts born before the Great Famine. In addition, the point estimates steadily increase among cohorts born during the famine, reaching a peak for cohorts born in 1961 and conceived in 1960. In summary, our estimates in Table 1 and panel A of Figure 6 verify our predictions on cohort sex ratios (Propositions 1) and suggest that the culture of son preference buffers the negative impact of famine shock on male survival chances.

5.2 The impact on adult height

To test the prediction about the gender gap in survivors' health, we use adult height to proxy for health which is from the 2010 China Family Panel Studies. One caveat is that adult height may not be a perfect proxy for survivors' health parameters. So the interpretations of the following results should be cautious. Due to the small sample size of CFPS data, we aggregate the dependent variable at the province birth year level. Column 1 of Table 2 presents estimates without pre-famine regional controls. In column 2, our preferred estimates further include these pre-famine controls. It shows that the impact of a one standard deviation increase in famine exposure raises the gender gap in height in gender-neutral areas by 1.27 centimeters. Meanwhile, the effect of one

[24] We also divide the sample by the level of son preference and conduct the flexible estimates for both groups in Figure A2. Estimates in Figure A2 demonstrate the same patterns as Figure 6.

standard deviation increase in famine exposure on the gender gap in height in areas with son preference is 0.98 centimeters lower than that in gender-neutral areas (9.80% relative to the mean), consistent with Proposition 2. In the last column, we reestimate the model when excluding those conceived and born during the famine. Results show that the interaction effect between famine exposure and son preference on cohort sex ratios remains robust and statistically significant. In a word, our findings on the gender gap in height imply that the selection effect of intrauterine famine shocks is more evident for girls than boys in areas with son preference, consistent with Gørgens et al. (2012) who find that the selection impact of famine shocks increases the average height of rural female famine survivors.

5.3 The impact on the gender gap in education

Table 3 presents the estimated effect of intrauterine famine exposure and its interacting effect with regional son preference on the gender gap in years of education. The specifications reported in each column of Table 3 are the same as those reported in Table 1. In columns 1-3, the dependent variable is the gender gap in years of education at the province birth year and month level. Column 1 reports the estimates for the full sample without pre-famine regional conditions. We find that intrauterine famine exposure significantly increases the gender gap in years of education. Moreover, we find that the interaction term between famine exposure and son preference is negative. These results remain robust when additionally including pre-famine regional controls in column 2. Column 2 shows that the change in the gender gap in years of education in response to a one standard deviation increase in famine severity in areas with son preference is 0.087 lower than in gender-neutral areas, equal to 6.41% of the mean (or 9.46% of the standard deviation). This means that famine shocks are expected to reduce the gender gap in education in areas with son preference, consistent with Proposition 3. In column 3, we reestimate the model when excluding those conceived and born during the famine. Results show that the interaction effect between famine exposure and son preference on the gender gap in years of education remains robust. We also report estimates using alternative ways to aggregate the dependent variable. Columns 4-5 report the results when the gender gap in years of education is aggregated at the province birth year level. The remaining two columns report the results when using the moving average of the gender gap in years of education (6 months before and after a particular birth year and month). Overall, our results remain robust to alternative ways to aggregate the dependent variable.

Next, we present the results of the fully flexible specification for the determinants of the gender gap in years of education. We plot the coefficients of the triple interaction terms in panel B of Figure 6. The patterns suggest that the relationship between famine shock interacting with regional son preference remains small in magnitude among pre-famine born cohorts and experiences a considerable decrease for cohorts born during the famine. It is worth noting that the largest impact of famine severity interacting with son preference is for those born in 1961 and conceived in 1960, the most severe year of the famine. Taken together, our findings suggest that the culture of son preference generates unintended consequences for the gender gap in years of education under adverse circumstances. Our findings are close to Yi et al. (2015), who find that household resource allocation responses exacerbate the adverse impact of adverse shocks to children on their educational attainment. In our context, household resource allocation (for example, schooling investment) between boys and girls may be responsive to famine experience, specifically to the effect of famine on health.

6 Robustness checks

Our identification strategy relies on the assumption that there are no other contemporaneous shocks affecting outcome variables beyond the famine shock and those factors for which we have controlled. We conduct several exercises to ensure the robustness of the analysis.

Son preference measure. We use the sex ratios of cohorts aged 0-3 in the 1953 Census to construct an alternative measure of son preference, as these younger cohorts were less likely to be influenced by the Chinese Civil War.[25] Columns 1-2 and 5-6 in Table 4 show that our results hold to using sex ratios of younger cohorts as a proxy for regional son preference. We also show that our results hold when using a continuous measure of son preference (columns 3-4 and 7-8 in Table 4).

Famine exposure measure. In panel A of Table 5, we show that our results are insensitive

[25] The Chinese Civil War was a civil war between Chinese Communist Party (CCP) and Kuomingtang (KMT) between 1929 and 1949.

to using the average death rate in 1954-1957 (pre-famine death rate) as the death rate in normal years to construct a new measure of famine severity. We also use province-level actual death rates to construct famine exposure to capture different timing of the famine across provinces. Panel B of Table 5 shows that our results remain robust and statistically significant when using provincial actual death rates to construct famine exposure.

Alternative sample restrictions. We examine the robustness of our estimates by including cohorts born between 1967 and 1970. Table A2 in the Appendix shows that our results are robust to including cohorts who were born after the famine.

Other social changes. We address the concern about the possible impacts of two major social changes: the Chinese Civil War and the Chinese Civil War. The Chinese Civil War is considered to affect individuals' survival chances (or cohort sex ratios). We use the province-level duration of civil war to capture its intensity. We include the interaction term between famine exposure and civil war duration in Table A3 (columns 1-2 and 5-6), and we find our results are still here. The Cultural Revolution may influence individuals' educational attainment. We get the measure of reform intensity from Walder (2014) based on recorded history from county gazetteers. The remaining columns of Table A3 show that our results remain significant and similar in magnitude compared to the main tables when including the controls of the Cultural Revolution.[26]

Famine exposure received in the first year of life. We also show that our results are robust to adding famine exposure during age 0-1. We construct a measure of famine exposure received in the first year of life. We include both famine exposure variables and their interactions with son preference and pre-famine regional conditions in Table A5 in the Appendix. Overall, our results hold when adding this control.

Placebo test. We conduct a placebo test using a younger subsample (those born between 1966 and 1978). Since none of these birth cohorts were ever directly exposed to the famine, we expect that the Great Famine, as well as the interaction between famine severity and regional son preference, would not produce any substantial effects. Table 6 shows that this placebo test works well.

[26] The measure of the Cultural Revolution intensity is the normalized total number of abnormal death of each province during 1966-1976. One SD increase in CR intensity leads to a 0.04 percentage points decrease in sex ratios and a 0.002 in the gender gap in years of education.

Selection during the famine. We examine overall selection effects during the famine using intrauterine famine severity to predict individual characteristics, specifically family background and sibship size, using 2010 CFPS. Overall, Table A7 implies that the gender difference in overall selection effects during the famine is not a serious concern.

7 Conclusions and discussions

In this paper, we study the long-run impact of early life events and their interaction with social norms, in the setting of China's Great Famine. We first put forward a theoretical framework to show how social norms can interact with adverse shocks to affect male and female survival chances, which shape gender inequalities in later life. We empirically test our theoretical predictions using data from the 2000 China Population Census and 2010 China Family Panel Survey.

We find that even though intrauterine famine shocks contribute to a reduction in sex ratios (males to females), the change in male-to-female sex ratios induced by the famine shocks in provinces with son preference is much smaller compared to that in gender-neutral provinces, implying that the culture of son preference buffers the negative impact of famine shock on male survival chances and worsens the survival chance of girls during the famine. In addition, we find that even though intrauterine famine shocks increase the gender gap in height in gender-neutral areas, the interacting effect with son preference leads to a reduction in the gender gap in height. Finally, we show that intrauterine famine shocks increase the gender gap in years of education, but this is less true in areas with a culture of son preference. Taken together, these findings suggest that the culture of son preference buffers the negative impact of famine shock on male-to-female sex ratios (survival chances), but through its impact on the expected health of boys and girls leads to smaller gender gaps in health and education in areas with son preference. These results are proven not to be driven by selective fertility and pass several robustness tests, including alternative sample restriction, alternative measures of treatment variables, controlling for the impact of the Cultural Revolution, and a placebo test.

Our findings link two pieces of literature on the consequences of the cultural preference for sons and the lasting impact of early life events. We show that the culture of son preference can interact with early life adverse shocks to shape gender inequalities in later life. Our theoretical framework and empirical findings, taken together, suggest that context is an important mediator of how local environmental shocks play out with respect to gender. In societies with strong son preference, especially in East Asia, people may discriminate against girls, especially under adverse circumstances. However, the long-term consequences may be different.

References

Almond, D. (2006). Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of Political Economy 114*(4), 672–712.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–72.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature 56*(4), 1360–1446.

Almond, D., L. Edlund, H. Li, and J. Zhang (2007). Long-term effects of the 1959-1961 China famine: Mainland China and Hong Kong. Technical report, National Bureau of Economic Research.

Almond, D. and B. Mazumder (2011). Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy. *American Economic Journal: Applied Economics 3*(4), 56–85.

Andersson, R. and S. Bergstrom (1998). Is maternal malnutrition associated with a low sex ratio¨ at birth? *Human Biology*, 1101–1106.

Arnold, F. and Z. Liu (1986). Sex preference, fertility, and family planning in China. *Population and Development Review*, 221–246.

Ashton, B., K. Hill, A. Piazza, and R. Zeitz (1992). Famine in China, 1958–61. In *The Population of Modern China*, pp. 225–271. Springer.

Barker, D. (1992). Fetal origins of diseases of old age. *European Journal of Clinical Nutrition*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bhalotra, S., C. Valente, and A. Van Soest (2010). The puzzle of Muslim advantage in child survival in India. *Journal of Health Economics 29*(2), 191–204.

Cameron, E. Z. (2004). Facultative adjustment of mammalian sex ratios in support of the Trivers Willard hypothesis: evidence for a mechanism. *Proceedings of the Royal Society B: Biological Sciences 271*(1549), 1723.

Case, A., A. Fertig, and C. Paxson (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics 24*(2), 365–389.

Catalano, R. and T. Bruckner (2006). Secondary sex ratios and male lifespan: damaged or culled cohorts. *Proceedings of the National Academy of Sciences 103*(5), 1639–1643.

Chen, J., Z. Xie, and H. Liu (2007). Son preference, use of maternal health care, and infant mortality in rural china, 1989–2000. *Population Studies 61*(2), 161–183.

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

Chen, Y. and L.-A. Zhou (2007). The long-term health and economic consequences of the 1959–1961 famine in china. *Journal of Health Economics 26*(4), 659–681.

Cui, H., J. P. Smith, and Y. Zhao (2020). Early-life deprivation and health outcomes in adulthood: Evidence from childhood hunger episodes of middle-aged and elderly Chinese. *Journal of Development Economics 143*, 102417.

Currie, J. and T. Vogl (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics 5*(1), 1–36.

Das Gupta, M., J. Zhenghua, L. Bohua, X. Zhenming, W. Chung, and B. Hwa-Ok (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.

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from South African homelands. *The Economic Journal 127*(604), 1906–1939.

Duque, V., M. Rosales-Rueda, and F. Sanchez (2018). How do early-life shocks interact with subsequent human-capital investments? evidence from administrative data. Technical report, Mimeo, University of Michigan.

Ebenstein, A. (2010). The “missing girls” of China and the unintended consequences of the one child policy. *Journal of Human Resources 45*(1), 87–115.

Edlund, L. (1999). Son preference, sex ratios, and marriage patterns. *Journal of Political Economy 107*(6), 1275–1304.

Gørgens, T., X. Meng, and R. Vaithianathan (2012). Stunting and selection effects of famine: A case study of the great Chinese famine. *Journal of Development Economics 97*(1), 99–111.

Guilmoto, C. Z. (2009). The sex ratio transition in Asia. *Population and Development Review 35*(3), 519–549.

Gupta, M. D. (1987). Selective discrimination against female children in rural Punjab, India. *Population and Development Review 13*(1), 77–100.

Jayachandran, S. and I. Kuziemko (2011). Why do mothers breastfeed girls less than boys? evidence and implications for child health in India. *The Quarterly Journal of Economics 126*(3), 1485–1538.

Jayachandran, S. and R. Pande (2017). Why are Indian children so short? the role of birth order and son preference. *American Economic Review 107*(9), 2600–2629.

Kraemer, S. (2000). The fragile male. *BMJ: British Medical Journal 321*(7276), 1609.

Li, W. and D. T. Yang (2005). The great leap forward: Anatomy of a central planning disaster. *Journal of Political Economy 113*(4), 840–877.

Lin, J. Y. and D. T. Yang (2000). Food availability, entitlements and the Chinese famine of 1959–61. *The Economic Journal 110*(460), 136–158.

Luo, Z., R. Mu, and X. Zhang (2006). Famine and overweight in China. *Review of Agricultural Economics 28*(3), 296–304.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–26.

Meng, X. and N. Qian (2009). The long term consequences of famine on survivors: evidence from a unique natural experiment using china's great famine. Technical report, National Bureau of Economic Research.

Meng, X., N. Qian, and P. Yared (2015). The institutional causes of china's great famine, 1959–1961. *The Review of Economic Studies 82*(4), 1568–1611.

Mu, R. and X. Zhang (2011). Why does the great Chinese famine affect the male and female survivors differently? mortality selection versus son preference. *Economics & Human Biology 9*(1), 92–105.

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

Nilsson, J. P. (2017). Alcohol availability, prenatal conditions, and long-term economic outcomes. *Journal of Political Economy 125*(4), 1149–1207.

Park, C. B. and N.-H. Cho (1995). Consequences of son preference in a low-fertility society: imbalance of the sex ratio at birth in Korea. *Population and Development Review*, 59–84.

Pathania, V. (2007). The long run impact of drought at birth on height of women in rural India. *Work. Pap., UC Berkeley*.

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature 465*(7299), 721–727.

Sen, A. (1990). More than 100 million women are missing. *The New York Review of Books 37*(20), 61–66.

Shi, X. (2011). Famine, fertility, and fortune in China. *China Economic Review 22*(2), 244–259.

Steckel, R. H. (1995). Stature and the standard of living. *Journal of Economic Literature 33*(4), 1903–1940.

Tanner, J. M. and J. M. Tanner (1981). *A history of the study of human growth*. Cambridge university press.

Valente, C. (2015). Civil conflict, gender-specific fetal loss, and selection: a new test of the Trivers– Willard hypothesis. *Journal of Health Economics 39*, 31–50.

Wachter, K. W. and J. Trussell (1982). Estimating historical heights. *Journal of the American Statistical Association 77*(378), 279–293.

Walder, A. G. (2014). Rebellion and repression in China, 1966–1971. *Social Science History 38*(34), 513–539.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal 125*(588), F347–F371.

Yi, Z., T. Ping, G. Baochang, X. Yi, L. Bohua, and L. Yongpiing (1993). Causes and implications of the recent increase in the reported sex ratio at birth in China. *Population and Development Review*, 283–302.

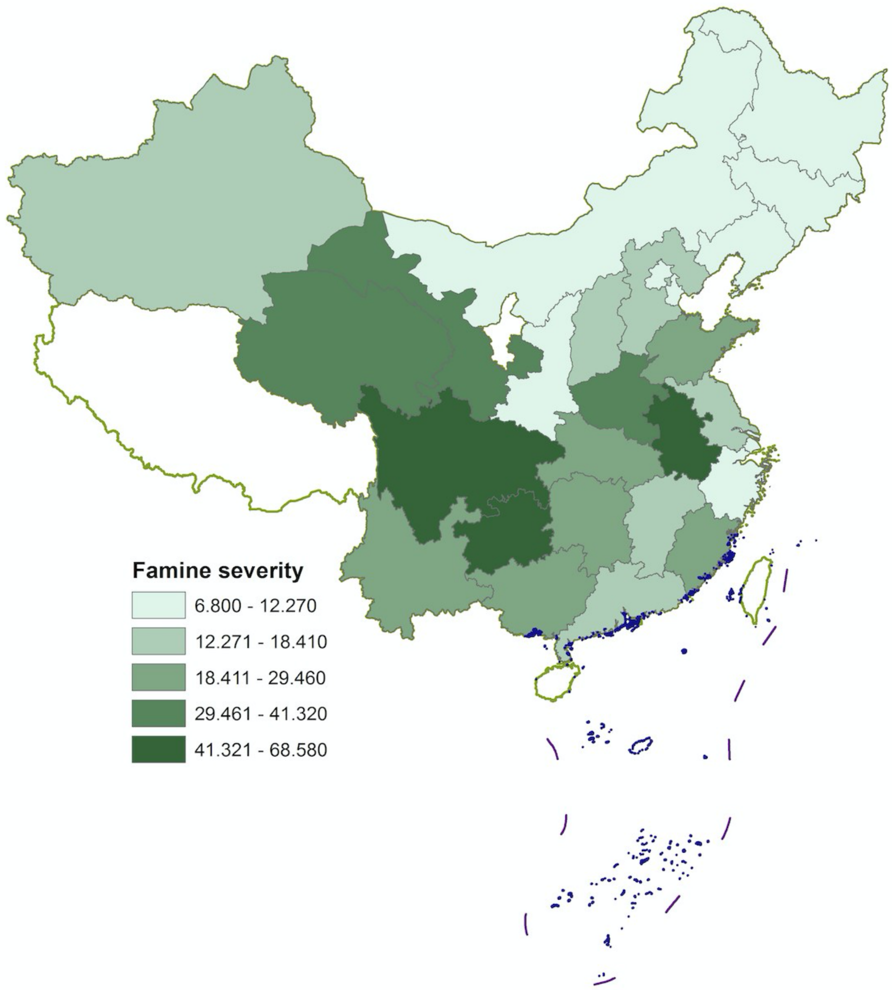


Figure 1: Geographic Distribution of Famine Severity

Notes: This figure shows the geographic distribution of famine severity measured by the provincial number of dead people (per thousand) in 1960. Data source: the Statistical Yearbook released by the National Bureau of Statistics of China.

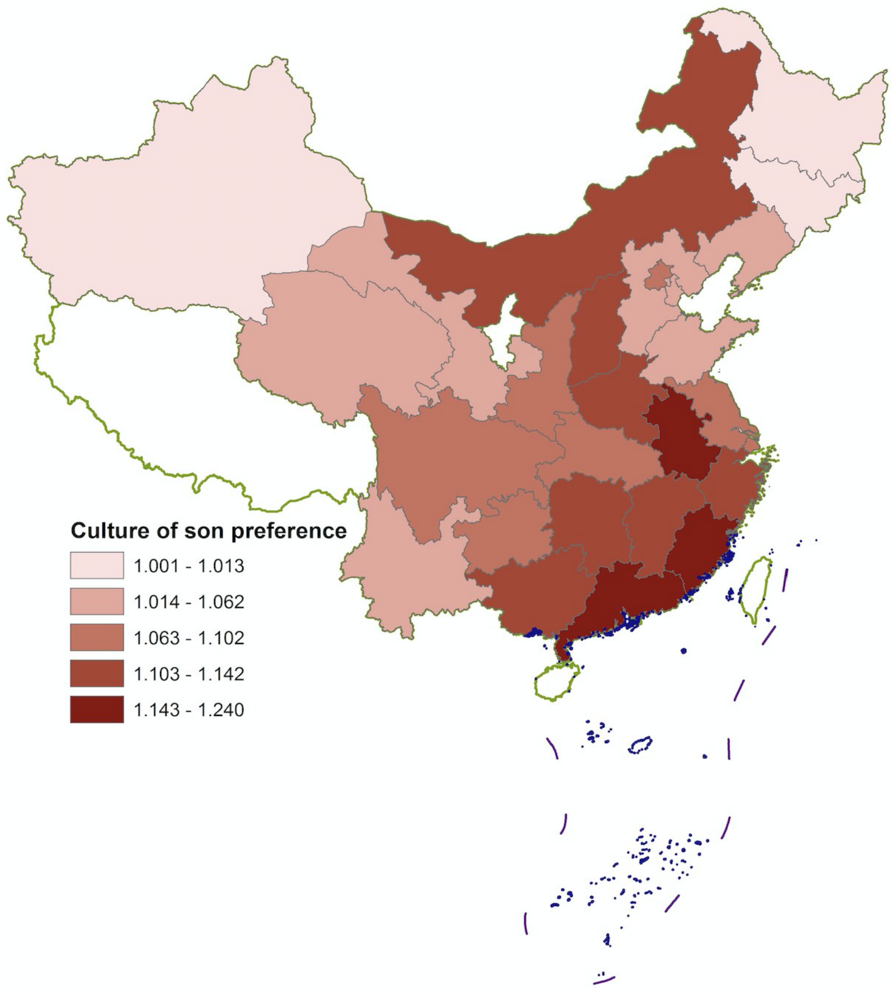
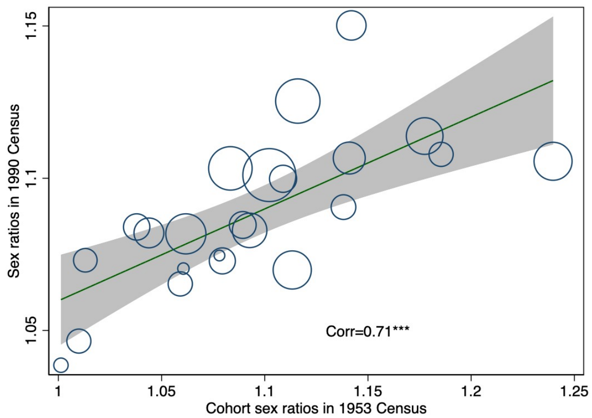
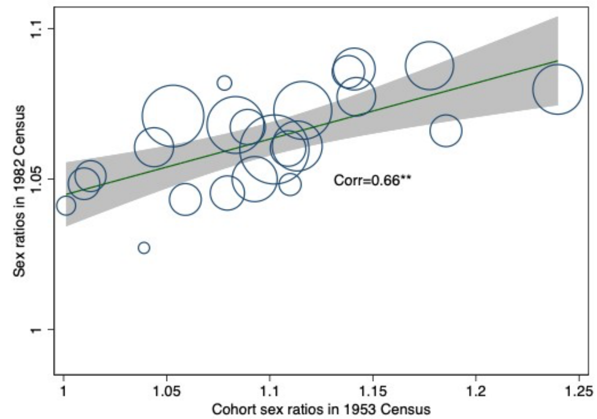


Figure 2: Geographic Distribution of Son Preference

Notes: This figure shows the geographic distribution of the culture of son preference measured by male-to-female sex ratios of cohorts aged from 0-10 in the 1953 census. Data source: the 1953 China Population Census.



(a) 1953 and 1990

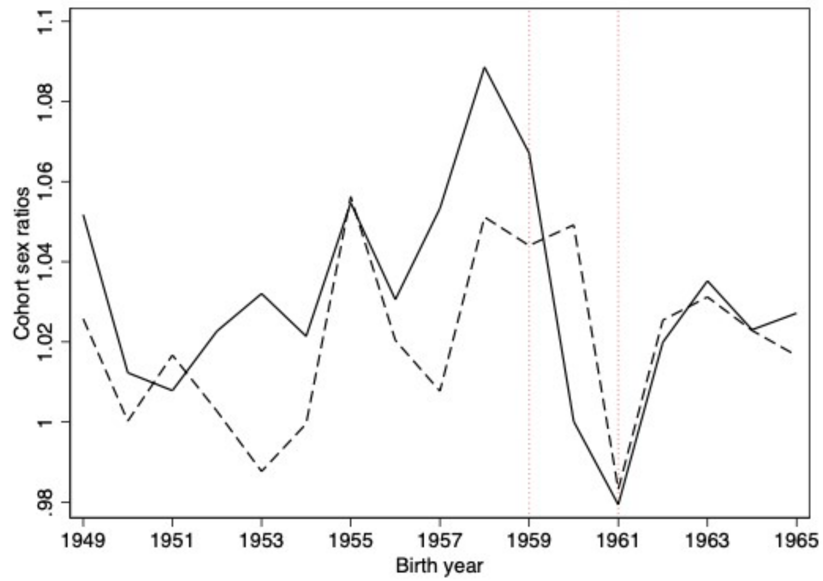


(b) 1953 and 1982

Figure 3: Correlations between Sex Ratios across Different Waves of Censuses

Notes: The figure plots the correlations between cohort sex ratios in 1953 (aged 0-10) and sex ratios in the 1990 and 1982 population censuses. Each circle represents a province. Circle size presents the size of population in the 1953 Census. Data sources: the Comprehensive Statistical Data and Materials on 50 Years of New China, the 1982 and 1990 China Population Census.

(a) Areas with a high degree of son preference



(b) Areas with a low degree of son preference

Figure 4: Sex Ratios by Famine Intensity and Regional Son Preference

Notes: This figure shows cohort male-to-female sex ratios by famine severity and regional culture of son preference. The black solid lines denote cohort sex ratios in high famine intensity areas, and the black dash lines indicate cohort sex ratios in low famine intensity areas. Cohorts born between 1959 and 1961 (within two red vertical lines) are famine-born cohorts. Data source: the 2000 China Population Census and Statistical Yearbook from the National Bureau of Statistics of China.

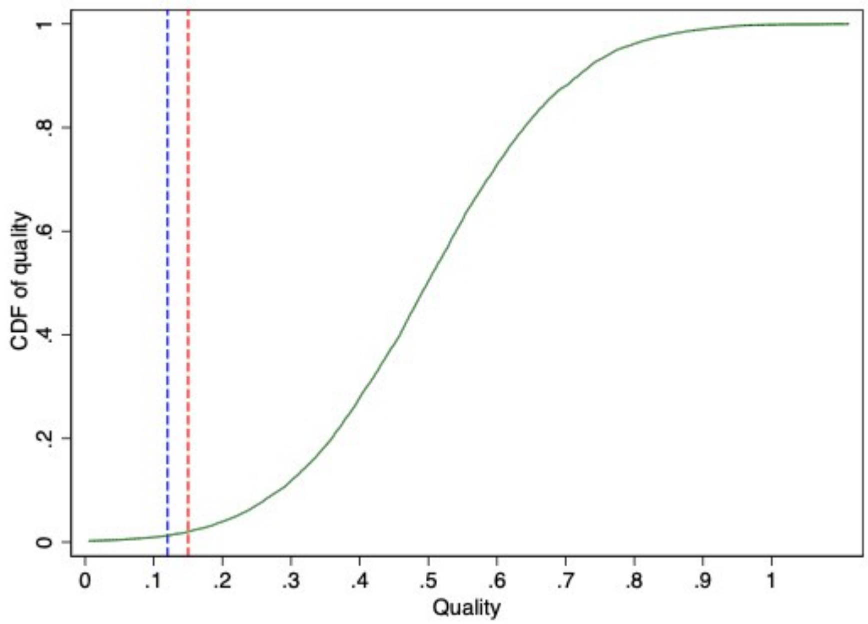
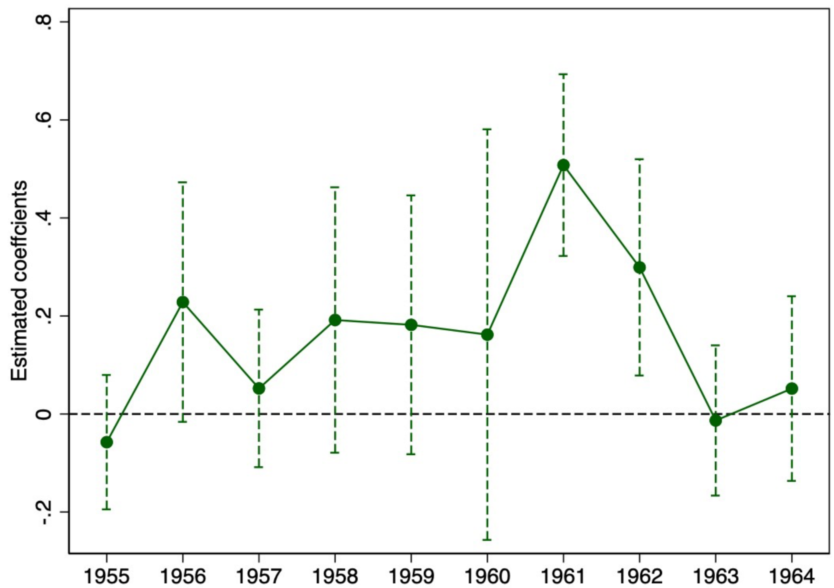
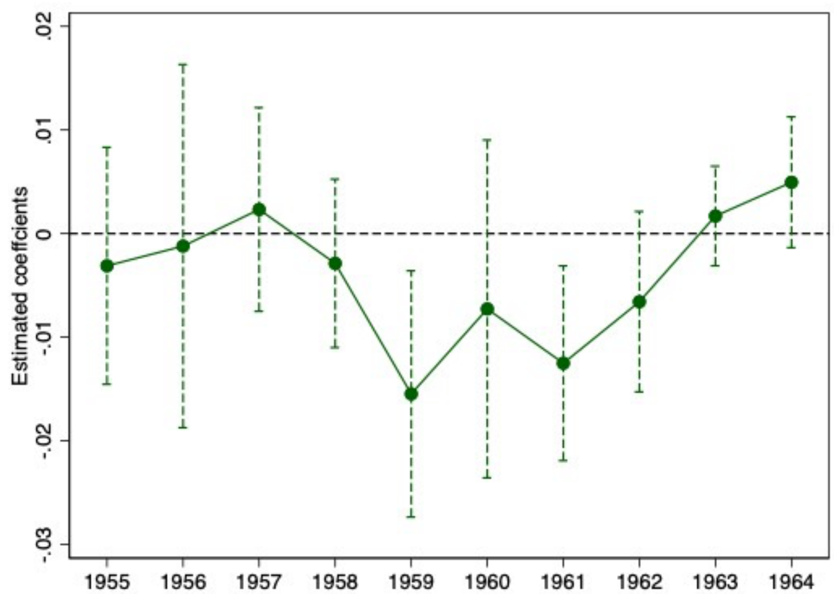


Figure 5: Quality Distributions and Survival Thresholds

Notes: This figure plots the CDF of male quality and the new male and female survival thresholds with respect to male quality distribution. The left blue line denotes the new survival threshold for females, and the right red line represents that for males.



(a) Cohort male-to-female sex ratios



(b) Gender gap in years of education

Figure 6: Flexible Estimates

Notes: The figure presents parameter estimates on interactions between birth year fixed effects, regional excess death rate during the famine and regional son preference on outcome variables, with 95% confidence intervals reported. We multiply sex ratios by 100. We include controls for pre-famine regional characteristics interacted with birth year fixed effects. Standard errors clustered at the province level are used to construct confidence intervals.

Table 1: Estimated Impacts on Cohort Male-to-female Sex Ratios

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	**Male-to-female sex ratios**			**Sex ratios (year)**		**Sex ratios (moving)**	
Famine	-0.2364***	-0.1064	0.1851	-0.1724***	-0.1094**	-0.2273***	-0.0941
	(0.084)	(0.097)	(0.1897)	(0.054)	(0.047)	(0.081)	(0.091)
Famine*son preference	0.2640***	0.3367***	0.4278***	0.1881***	0.2489***	0.2523***	0.3219***
	(0.078)	(0.059)	(0.0845)	(0.055)	(0.052)	(0.073)	(0.057)
Famine*GDP per capita		0.0020**	0.0044***		0.0014**		0.0018**
		(0.001)	(0.0015)		(0.001)		(0.001)
Famine*health institutions		-0.0000	-0.0000		-0.0000		-0.0000
		(0.000)	(0.0001)		(0.000)		(0.000)
Famine* high school students		0.0470	-0.0073		0.0582		0.0531
		(0.073)	(0.0874)		(0.066)		(0.077)
Observations	4212	4212	3402	4212	4212	4212	4212
R-squared	0.165	0.166	0.180	0.466	0.470	0.497	0.502

Notes: This table reports the effect of intrauterine famine exposure and its interaction with son preference on cohort sex ratios. The unit of observation is at the province birth year and month level. In columns 1 through 3, the dependent variable is the male-to-female sex ratios of each cell at the province birth year and month level. In columns 4-5, we use dependent variable aggregated at province and birth year level. In columns 6-7, we use a moving average of the dependent variable (6 months before and after). We multiply sex ratios (dependent variable) by 100. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by the population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table 2: Estimated Impacts on The Gender Gap in Adult Height

	(1)	(2)	(3)
Dependent variables	**Gender gap in height (male-female)**		
Famine	0.0270	0.2157***	0.0895
	(0.0492)	(0.0609)	(0.1322)
Famine*son preference	-0.0747	-0.1663***	-0.5003***
	(0.0506)	(0.0538)	(0.1020)
Famine*GDP per capita		0.2813***	-0.3184
		(0.0898)	(0.2463)
Famine*health institutions		0.2300***	0.1729**
		(0.0382)	(0.0817)
Famine* high school students		-0.2169***	-0.3119***
		(0.0618)	(0.1091)
Observations	291	265	229
R-squared	0.150	0.213	0.214

Notes: This table reports the effect of intrauterine famine exposure and its interaction with son preference on the gender gap in adult height. The unit of observation is at the province birth year level. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2010 China Family Panel Studies (CFPS) and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table 3: Estimated Impacts on The Gender Gap in Years of Education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	**Gender gap in education**			**Gap edu. (year)**		**Gender edu. (moving)**	
Famine	0.0117**	0.0110***	0.0111**	0.0092**	0.0100***	0.0109**	0.0101**
	(0.004)	(0.004)	(0.0052)	(0.003)	(0.003)	(0.004)	(0.004)
Famine*son preference	-0.0127***	-0.0148***	-0.0152***	-0.0111***	-0.0132***	-0.0120***	-0.0143***
	(0.003)	(0.003)	(0.0035)	(0.003)	(0.003)	(0.004)	(0.003)
Famine*GDP per capita		-0.0001***	-0.0001***		-0.0001***		-0.0001***
		(0.000)	(0.0000)		(0.000)		(0.000)
Famine*health institutions		-0.0000***	-0.0000**		-0.0000***		-0.0000***
		(0.000)	(0.0000)		(0.000)		(0.000)
Famine* high school students		-0.0034	-0.0026		-0.0025		-0.0031
		(0.002)	(0.0026)		(0.002)		(0.002)
Observations	4212	4212	3402	4212	4212	4212	4212
R-squared	0.821	0.823	0.832	0.955	0.957	0.955	0.957

Notes: This table reports the effect of intrauterine famine exposure and its interaction with son preference on the gender gap in years of education. The unit of observation is at the province birth year and month level. In columns 1 through 3, the dependent variable is the gender gap in years of education of each cohort at the province birth year and month level. In columns 4-5, we use dependent variable aggregated at province and birth year level. In columns 6-7, we use a moving average of the dependent variable (6 months before and after). Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table 4: Robustness to Different Measures of Son Preference

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables	**Male-to-female sex ratios**				**Gender gap in education**			
Famine	-0.1431	0.0016	-1.8308**	-2.4372***	0.0125***	0.0066*	0.0597	0.1120***
	(0.093)	(0.097)	(0.8828)	(0.8061)	(0.003)	(0.004)	(0.0463)	(0.0380)
Famine*sp03	0.1931**	0.3034***			-0.0149***	-0.0149***		
	(0.088)	(0.063)			(0.002)	(0.002)		
Famine*sp_continuous			1.6750**	2.3171***			-0.0548	-0.1005***
			(0.7876)	(0.7470)			(0.0403)	(0.0350)
Famine*GDP per capita		0.0015		0.0011		-0.0001***		-0.0001**
		(0.001)		(0.0009)		(0.000)		(0.0000)
Famine*health institutions		-0.0001**		-0.0000		-0.0000		-0.0000***
		(0.000)		-0.0000		(0.000)		(0.0000)
Famine* high school students		0.0815		0.1582**		-0.0051**		-0.0083***
		(0.070)		(0.0677)		(0.002)		(0.0025)
Observations	4212	4212	4212	4212	4212	4212	4212	4212
R-squared	0.164	0.166	0.164	0.165	0.822	0.823	0.820	0.822

Notes: We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. Sp_03 uses the sex ratios of cohorts aged 0-3 in the 1953 China Population Census. Sp_continous uses the continuous measure of sex ratios of cohorts aged 0-10 in the 1953 China Population Census. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 5: Robustness to Different Measures of Famine Exposure

	(1)	(2)	(3)	(4)
Dependent variables	**Male-to-female sex ratios**		**Gender gap in education**	
Panel A: alternative excess death rate				
Famine_excess	-0.3566**	-0.3282**	0.0205***	0.0223***
	(0.1507)	(0.1298)	(0.0050)	(0.0039)
Famine_excess*son preference	0.3998**	0.4966***	-0.0227***	-0.0352***
	(0.1488)	(0.1464)	(0.0047)	(0.0039)
Famine_excess*GDP per capita		0.0015		-0.0003***
		(0.0013)		(0.0001)
Famine_excess*health institutions		-0.0000		-0.0000**
		(0.0000)		(0.0000)
Famine_excess* high school students		0.0515		-0.0100*
		(0.1619)		(0.0050)
Observations	4212	4212	4212	4212
R-squared	0.165	0.165	0.821	0.823
Panel B: actual death rate				
Famine_actual	-0.3566*	-0.3576**	0.0127**	0.0127*
	(0.184)	(0.150)	(0.006)	(0.007)
Famine_actual*son preference	0.3600*	0.4666***	-0.0129**	-0.0111
	(0.178)	(0.156)	(0.006)	(0.009)
Famine_actual*GDP per capita		0.0007		-0.0000
		(0.001)		(0.000)
Famine_actual*health institutions		-0.0001**		-0.0000
		(0.000)		(0.000)
Famine_actual* high school students		0.2267**		0.0002
		(0.086)		(0.004)
Observations	4212	4212	4212	4212
R-squared	0.165	0.167	0.820	0.821

Notes: Panel A reports the estimates using an alternative definition of excess death rate. Famine_excess is constructed using the average death rate in 1954 and 1957 (pre-famine death rate) as the death rate in normal years. Panel B reports the estimates using actual provincial death rate to construct famine exposure in utero. Famine_actual denotes the intrauterine famine exposure based on the actual provincial death rate. Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Table 6: Placebo Tests

	(1)	(2)	(3)	(4)
Dependent variables	**Male-to-female sex ratios**		**Gender gap in education**	
Famine	-0.0010	0.0246	-0.0037	-0.0014
	(0.0751)	(0.0921)	(0.0035)	(0.0031)
Famine*son preference	0.0298	0.0486	0.0034	0.0051
	(0.0628)	(0.0709)	(0.0031)	(0.0032)
Famine*GDP per capita		0.0007		0.0001**
		(0.0009)		(0.0000)
Famine*health institutions		0.0000		0.0000
		(0.0000)		(0.0000)
Famine* high school students		0.0208		-0.0019
		(0.0878)		(0.0020)
Sample	1966-1978	1966-1978	1966-1978	1966-1978
Observations	4212	4212	4212	4212
R-squared	0.187	0.187	0.777	0.777
Province FEs	Yes	Yes	Yes	Yes
Birth year and month FEs	Yes	Yes	Yes	Yes

Notes: We construct fake (placebo) famine shock which happened 13 years later (between 1971 and 1974) and conduct our analysis on cohorts who were born between 1966 and 1978. We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Social Norms and The Impact of Early Life Events on Gender Inequality Online Appendices

Wei Luo, University of Hong Kong

Wei Huang, Emory University

Albert Park, Hong Kong University of Science and Technology

A Data Description

2000 China Population Census. The 2000 Chinese census, officially the Fifth National Population Census of the People's Republic of China, was conducted by the government of the People's Republic of China on 1 November 2000. The census contains information on demographics, work, and households. The 2000 China National Population Census sample that we use is from IPUMS, a 1% sample of 11,804,344 persons.

China Family Panel Studies (CFPS). China Family Panel Studies (CFPS) is a nationally representative data. The first national wave was conducted in collaboration with the Institute of Social Science Survey at the Peking University and the Survey Research Center at the University of Michigan from April to August 2010. The five main parts of the questionnaire include data on communities, households, household members, adults, and children. The CFPS sample in 2010 covers 25 provinces/municipalities/autonomous regions, representing 95% of the Chinese population. The 2010 baseline survey interviewed a total of 14,960 households and 42,590 individuals. CFPS implemented Probability-Proportional-to-Size Sampling (PPS) with implicit stratification, taking the regional differences in Chinese society and reducing survey processing costs into consideration. We report the summary statistics for selected variables in Table A1.

B Robustness checks

B.1 Alternative sample restriction

In the main context, we focus on cohorts born between 1954 and 1966 in the 2000 China Population Census. Here, we examine the robustness of our estimates to including more cohorts born after the famine, for example, cohorts born between 1967 and 1970. Table A2 shows that the average treatment effect model estimates on cohort sex ratios and the gender gap in education are similar to what we report in Tables 1 and 3. Estimated impacts of the interaction term between intrauterine famine severity and son preference on cohort sex ratios and the gender gap in years of education remain robust and similar in magnitude.

B.2 Son preference measure

We check the robustness of our analysis using various ways to construct son preference measures.

Sex ratios of younger cohorts. Our main text uses the sex ratios of cohorts aged 0-10 in the 1953 China Population Census to measure son preference. One concern is that cohort before 1950 may be influenced by the Chinese Civil War. We alleviate the concern by utilizing the sex ratios of cohorts born after 1949 to construct an alternative measure. Columns 1-2 and 5-6 in Table 4 show that our results almost remain robust to this alternative measure of son preference, even though the coefficients of the impact of famine exposure on sex ratios lose significance.

Continuous measure. In the main text, we use a dummy variable indicating whether individuals were born in areas with son preference or not. Here, we replicate our main analysis by simply using male-to-female sex ratios of cohorts aged 0-10 in the 1953 China Population Census as the alternative measurement of regional cultural preference for sons. Columns 3-4 and 7-8 in Table 4 show that our results are robust to the continuous measurement of son preference.

B.3 Famine exposure measure

We check the robustness of our analysis to different ways to construct intrauterine famine exposure.

Alternative excess death rate. In the main text, the province-level excess death rate is defined as the difference between the average death rate during the famine period (1958-1961) and the average death rate during normal years (1972-1976). Here, we utilize the average death rate in 1954-1957 (pre-famine death rate) as the death rate in normal years to construct a new provincial excess death rate. Panel A of Table 5 shows that our results remain robust and statistically significant to this new measure, with a slightly larger magnitude in estimated coefficients.

Provincial actual death rate. To address the concern that the timing of the famine may be different across provinces, we use actual province death rates to construct famine exposure. Panel B of Table 5 demonstrates that our results remain robust to using provincial actual death rate to construct famine exposure.

B.4 Other social changes

Here, we address the concern about the possible impacts of two major social changes: the Chinese Civil War and the Cultural Revolution.

Civil war. In our main text, we use the cohort sex ratios of those aged 0-10 in the 1953 China Population Census as a proxy for the culture of son preference. However, these cohorts' survival chances may be influenced by the Chinese Civil War between 1945-and 1949. We utilize the number of months that each province experienced the Civil War to capture the intensity of the war. We introduce the interaction term between famine exposure and civil war duration in Table A3. Columns 1-2 and 5-6 in Table A3 show that our results remain robust to adding the civil war duration.

Cultural Revolution. The Cultural Revolution may also influence the educational attainment of our sample (those born between 1954 and 1966). We measure the intensity of the Cultural Revolution using the data from Walder (2014), based on recorded history from county gazetteers on the number of abnormal deaths due to the revolution between 1966 and 1976. We aggregate the total number of abnormal deaths during the CR period at the province level and introduce the interaction between famine exposure and Cultural Revolution control in Table A3. Our results hold to the inclusion of Cultural Revolution control. Meanwhile, the interaction term between famine severity and Cultural Revolution intensity produces negligible negative impacts on outcome variables of interest.

B.5 Famine exposure received in the first year of life

Even though extensive literature in biology and economics shows that fetuses are vulnerable and adverse fetal conditions would have huge and lasting impacts on adult outcomes (Barker, 1992; Almond and Currie, 2011; Currie and Vogl, 2013) as this intrauterine period may be very crucial for gene programming (Petronis, 2010). Here, we construct famine exposure received in the first year of life to check whether exposure during other periods has some effects. Results in Table A4 show no statistically significant impact on sex ratios. Even though the coefficient of the interaction between famine severity received during age 0-1 and son preference on the gender gap in education is statistically significant, the magnitude and the significance of the coefficient are reduced compared to the estimates using intrauterine famine exposure.

B.6 Placebo test

Changes in provincial excess death rates should not be systematically correlated to other omitted factors that affect cohort sex ratios or educational attainment to ensure that difference-in-differences estimators remain free from bias. We conduct a placebo test using a subsample of individuals born between 1966 and 1978 (containing the same number of birth year cohorts as we use in the main analysis). We define cohorts born between 1971 and 1974 as famine-exposed cohorts to simulate the famine situation. Since none of these birth cohorts were ever directly exposed to the famine, we expect that the Great Famine and the interaction between famine severity and regional son preference would not produce any substantial effects on this subsample. Results in Table 6 indicate that this placebo test works well.

C Additional analysis

C.1 Linking educational and health outcomes

We use the 2010 China Family Panel Studies (CFPS) data to construct the gender gap in height for each birth year cohort at the province level. Then we control for the gender gap in height in our analysis of education in Table A5. We find that the coefficient of famine severity reduces in magnitude and statistical significance after controlling for the gender gap in height. Moreover, the coefficient of the interaction term between famine severity and son preference also slightly loses its magnitude. We confirm the point that stronger girls survived during the famine in areas with son preference than in gender-neutral areas, further shaping the gender gap in education.

At last, we observe that the gender gap in height is positively correlated with the gender gap in education. There are two explanations for the reduction of the coefficient of famine severity in utero after adding health controls. (1) Health (nutrition) accounts for part of the results (around 20%). And there may be other factors influencing education investment decisions. (2) Height may be only one dimension of health that we can control for, and there are many other aspects of health that may influence education decisions.

C.2 Absolute changes

We reproduce the estimates for absolute changes in adult height and years of education of each gender in Table A6. For each outcome variable, we first report the results without pre-famine regional controls and then add these controls in. We find that intrauterine famine exposure slightly increases male height in columns 1-2, even though the coefficient is not statistically significant before adding pre-famine regional controls. However, we do not observe any statistically significant effect of the interaction term between famine exposure and son preference in columns 1-2. In addition, in columns 3-4, we find no impact of intrauterine famine exposure or its interacting effect with son preference on female height.

The remaining four columns report the estimated impacts on male and female years of education, respectively. We find that intrauterine famine exposure and its interaction with son preference have no statistically significant effects on male years of education, regardless of the inclusions of pre-famine regional conditions. Moreover, column 7 shows that intrauterine famine exposure significantly reduces female years of education. Meanwhile, its interacting effect with son preference is positive, implying that the negative impact of famine exposure on female years of education is smaller in areas with son preference. Column 8 shows that these findings hold to the inclusion of pre-famine regional controls.

C.3 Selection during the famine

Some literature suggests that two kinds of selection exist during the famine period. First, fertility decisions could be postponed during the famine. Second, individuals who are observed in later Population Census or survey data should experience the selection of mortality. We have utilized birth month to deal with the selective fertility issue in our main context. Here, we examine the overall selection during the famine using intrauterine famine severity to predict family background and sibship size, using data from the China Family Panel Studies (CFPS) 2010.

Parents' education level and party membership are proxies for family background. Columns 1-4 in Table A7 suggest no substantial impact of intrauterine famine exposure on family background by several yardsticks. Furthermore, we do not observe any gender difference in the effects of famine exposure on family background. The last three columns in Table A7 show that intrauterine famine exposure has no substantial effect on sibship size, regarding the total sibship size and female sibship size. Besides, there is no gender difference in the impact of famine exposure on sibship size. Overall, our results imply that the gender difference in the overall selection during the famine is not so serious regarding respondents' family background and sibship information.

D Proofs of propositions

Proof of Proposition 1

Proof. The impact of famine on survival chances between two groups is determined by the change in CDF of the standardized health distribution. For the standard normal distribution, we know that $\Phi'(x) = \varphi(x)$, $\varphi'(x) = -z\varphi(x)$, where $\Phi(.)$ denotes the standard cumulative distribution function, and $\varphi(.)$ represents the standard normal density function. As $\Phi'(x) > 0$ for any x, we have that $\varphi'(x) > 0$ for any $x < 0$. Here, Lemma 1 stated as follows, according to the property that $\varphi'(x) > 0$ for any $x < 0$.

Lemma 1. *If* $x_1 < x_2 < 0$, *for any small positive* Δv, *we have that* $\frac{\Phi(x_2+\Delta v)-\Phi(x_2)}{\Delta v} > \frac{\Phi(x_1+\Delta v)-\Phi(x_1)}{\Delta v}$.

Our proof of Proposition 1 is constructed as follows.

(1) Preference for boys is greater in areas with son preference than that in gender-neutral areas. Thereby, the survival line for boys is greater in gender-neutral areas than in areas with son preference. By Lemma 1, we know that:

$$\Phi\left(\frac{z-\theta_b^n+\Delta\mu-\mu_b}{\sigma}\right)-\Phi\left(\frac{z-\theta_b^n-\mu_b}{\sigma}\right)>\Phi\left(\frac{z-\theta_b^p+\Delta\mu-\mu_b}{\sigma}\right)-\Phi\left(\frac{z-\theta_b^p-\mu_b}{\sigma}\right)$$

It implies that the impact of famine on boys' survival probability reduction is greater in gender-neutral areas than in areas with son preference.

(2) Similarly, preference for girls is smaller in areas with son preference than in gender-neutral areas. Thus, the survival line for girls is greater in areas with son preference than in gender-neutral areas. By Lemma 1, we know that:

$$\Phi\left(\frac{z-\theta_g^p+\Delta\mu-\mu_g}{\sigma}\right)-\Phi\left(\frac{z-\theta_g^p-\mu_g}{\sigma}\right)>\Phi\left(\frac{z-\theta_g^n+\Delta\mu-\mu_g}{\sigma}\right)-\Phi\left(\frac{z-\theta_g^n-\mu_g}{\sigma}\right)$$

It indicates that the impact of famine on girls' survival probability reduction is greater in areas with son preference than in gender-neutral areas.

(3) According to inequalities in (1) and (2), we have that the gender difference (boys versus girls) in the impact of famine on survival probability reduction in gender-neutral areas is greater than in areas with son preference, and Proposition 1 is proved.

Proof of Proposition 2

Proof. We analyze the impact of famine on the expected health of survivors in gender-neutral areas. The difference in the effect of famine on expected health between the two groups is only determined by the inverse Mills ratios (selection effect). For the standard normal distribution, the inverse mills ratio, we know that $\lambda(x) > 0$, $\lambda'(x) > 0$, and $\lambda''(x) > 0$. As for any x, we have that $\lambda''(x) > 0$. Thereby, we get Lemma 2 as follows.

Lemma 2. If x1 < x2, for any positive Δv, we have $\frac{\lambda(x_2 + \Delta v) - \lambda(x_2)}{\Delta v} > \frac{\lambda(x_1 + \Delta v) - \lambda(x_1)}{\Delta v}$.

The proof of Proposition 2 proceeds as follows.

(1) Preference for boys is greater in areas with son preference than in gender-neutral areas. The survival line for boys is greater in gender-neutral areas than in areas with son preference. By Lemma 2, we know that:

$$\lambda\left(\frac{z - \theta_b^n + \Delta\mu - \mu_b}{\sigma}\right) - \lambda\left(\frac{z - \theta_b^n - \mu_b}{\sigma}\right) > \lambda\left(\frac{z - \theta_b^p + \Delta\mu - \mu_b}{\sigma}\right) - \lambda\left(\frac{z - \theta_b^p - \mu_b}{\sigma}\right)$$

Thus, the impact of famine on boys' expected health is greater in gender-neutral areas than in areas with son preference.

(2) Similarly, preference for girls is greater in gender-neutral areas than that in regions with son preference. Thus, the survival line for girls in areas with son preference is greater than in gender-neutral areas. By Lemma 2, we know that the impact of famine on girls' expected health is greater in regions with son preference than in gender-neutral areas.

(3) According to these two inequalities mentioned above, we have that the gender difference (boys versus girls) in the impact of famine on expected health is greater in gender-neutral areas than that in areas with son preference, and Proposition 2 is proved.

Proof of Proposition 3

Proof. We focus on the impact of famine on years of schooling (logarithm form). The impact of famine on schooling is determined by its effect on the expected health. Our proof of Proposition 3 is constructed as follows.

(1) We begin by comparing the impact of famine on the expected health of boys in areas with son preference and gender-neutral areas. By rearranging the terms, we have that:

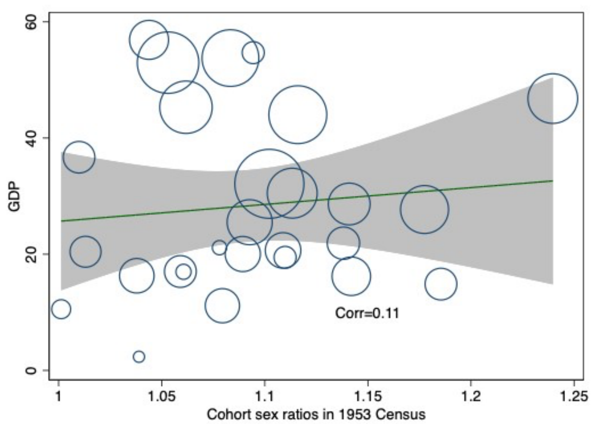
$$\ln\frac{Q\left(E_b > z_b^{\prime p} + \Delta\,\mu\right)}{Q\left(E_b > z_b^{\prime p}\right)} - \ln\frac{Q\left(E_b > z_b^{\prime n} + \Delta\,\mu\right)}{Q\left(E_b > z_b^{\prime n}\right)} = \ln\frac{1 + \dfrac{Q\left(E_b > z_b^{\prime p} + \Delta\,\mu\right) - Q\left(E_b > z_b^{\prime p}\right)}{Q\left(E_b > z_b^{\prime p}\right)}}{1 + \dfrac{Q\left(E_b > z_b^{\prime n} + \Delta\,\mu\right) - Q\left(E_b > z_b^{\prime n}\right)}{Q\left(E_b > z_b^{\prime n}\right)}}$$

It suggests that it is jointly determined by the impact of famine in both areas as well as the expected health of non-exposed cohorts. Proposition 2 shows that the impact of famine expected health is greater for boys in gender-neutral areas than in areas with son preference. Meanwhile, the expected health of non-exposed boys in areas with son preference is higher than that in gender-neutral areas. Taken together, we know that $\ln\frac{Q(E_b > z_b^{\prime p} + \Delta\mu)}{Q(E_b > z_b^{\prime p})} - \ln\frac{Q(E_b > z_b^{\prime n} + \Delta\mu)}{Q(E_b > z_b^{\prime n})} < 0$.

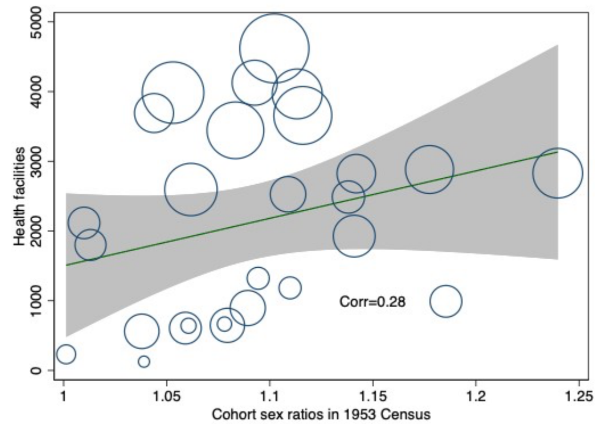
(2) Similarly, we compare these relationships for girls. Proposition 2 shows that the impact of famine on expected health is greater for girls in areas with son preference than in gender-neutral areas. Meanwhile, the expected health of non-exposed girls in gender-neutral areas is greater than that in areas with son preference. Taken together, we know that:

$$\ln\frac{Q\left(E_g > z_b^{\prime p} + \Delta\,\mu\right)}{Q\left(E_g > z_b^{\prime p}\right)} - \ln\frac{Q\left(E_g > z_b^{\prime n} + \Delta\,\mu\right)}{Q\left(E_g > z_b^{\prime n}\right)} > 0$$

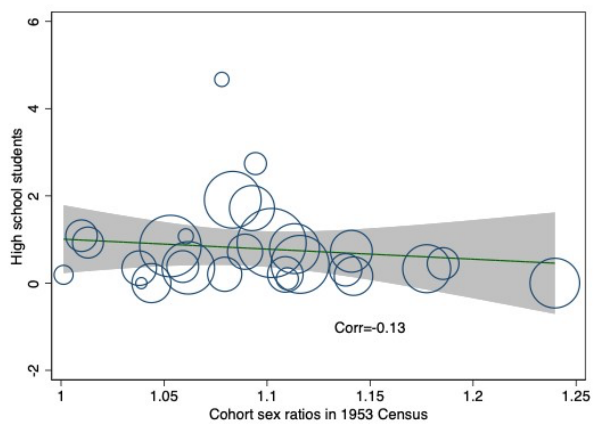
(3) According to two inequalities in (1) and (2), we know that $ds^p - ds^n < 0$. The gender difference in the impact of famine on schooling is smaller in areas with son preference than in gender-neutral areas, and Proposition 3 is proved.



(a) GDP



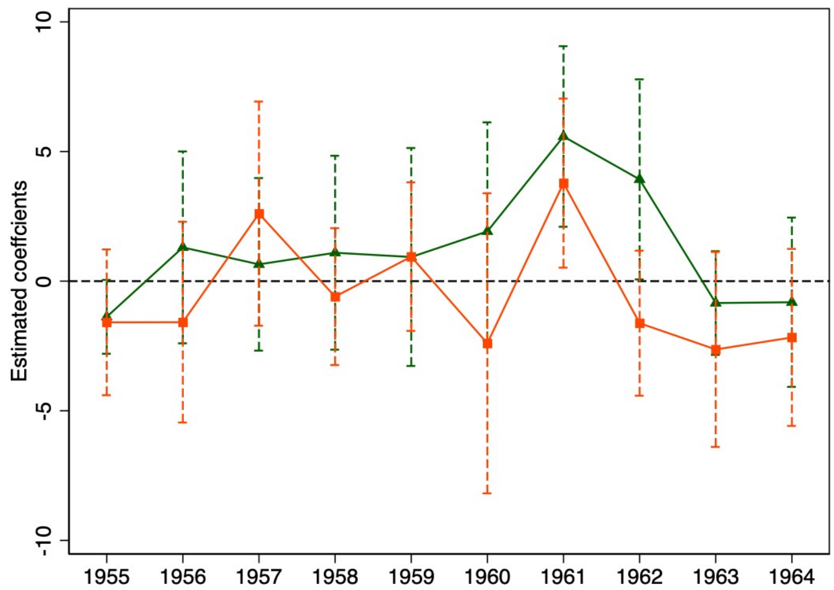
(b) Health facilities



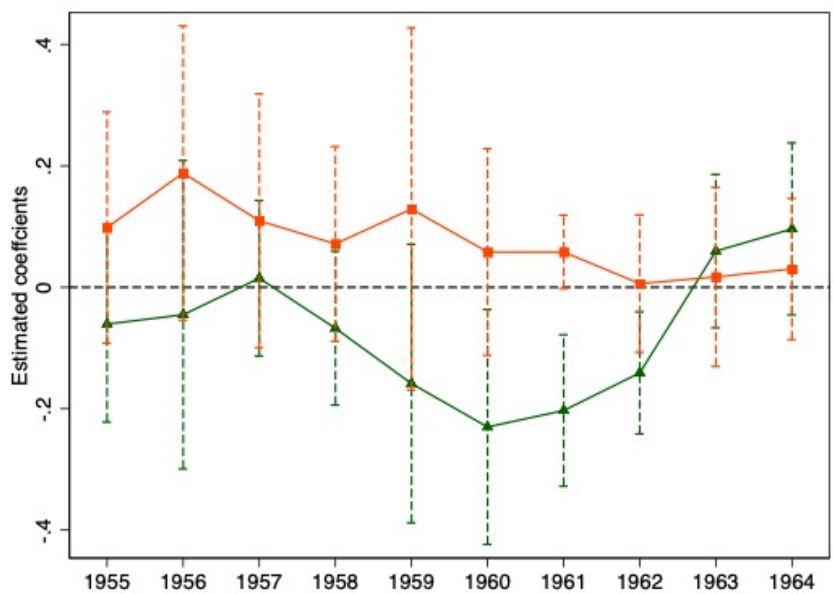
(c) Number of high school students

Figure A1: Correlations between Son Preference and Pre-famine Conditions

Notes: The figure plots the correlations between son preference (measured by cohort sex ratios in 1953) and regional conditions in 1954, including GDP, number of health facilities, and number of high school students. Each circle represents a province. Circle size presents the size of population in the 1953 Census. Data sources: the Comprehensive Statistical Data and Materials on 50 Years of New China.



(a) Cohort sex ratios



(b) Gender gap in education

Figure A2: Flexible Estimates for Areas with Different Levels of Famine Severity

Note: The figure presents parameter estimates on birth year fixed effects and regional excess death rate in areas with a high level of famine severity and regions with a low level of famine severity, with 95% confidence intervals reported. We multiply sex ratios by 100. We include controls for pre-famine regional characteristics interacted with birth year fixed effects. Standard errors clustered at the province level are used to construct confidence intervals.

Table A1: Summary Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
Province birth year and birth month level					
Male-to-female sex ratios	4212	1.05	0.13	0.33	3.00
Gender gap in years of education	4212	1.36	0.92	-1.48	6.84
Famine severity received in utero	4212	3.07	5.89	0.00	34.02
Province level					
Son preference (dummy variable)	27	0.63	0.49	0.00	1.00
GDP per capita in 1954	27	173.60	120.03	66.00	589.00
Number of health institutions in 1954	27	2123.96	1373.06	126.00	4619.00
Number of high school students per 10000 in 1954	27	0.80	1.00	0.00	4.67

Notes: The data on sex ratios and the gender gap in education is from the 2000 China Population Census. The information on province-level year-by-year death rate and pre-famine provincial characteristics are from Comprehensive Statistical Data and Materials on 50 Years of New China which is compiled by the Department of Comprehensive Statistics of the National Bureau of Statistics in China. In addition, information on province-level cohorts sex ratios is calculated from the 1953 China Population Census.

Table A2: Inclusion of Later Cohorts

	(1)	(2)	(3)	(4)
VARIABLES	**Male-to-female sex ratios**		**Gender gap in education**	
Famine	-0.2365***	-0.1029	0.0125*	0.0115**
	(0.0822)	(0.1045)	(0.0067)	(0.0054)
Famine*son preference	0.2983***	0.3663***	-0.0140**	-0.0174***
	(0.0785)	(0.0617)	(0.0054)	(0.0056)
Famine*GDP per capita		0.0017**		-0.0002***
		(0.0008)		(0.0000)
Famine*health institutions		-0.0000		-0.0000***
		(0.0000)		(0.0000)
Famine* high school students		0.0566		-0.0010
		(0.0780)		(0.0034)
Observations	5508	5508	5508	5508
R-squared	0.166	0.167	0.828	0.830

Notes: This table reports the robustness of the estimated impact of intrauterine famine exposure and its interaction with son preference to alternative sample restriction (cohorts born between 1954-and 1970). We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table A3: Robustness to Other Social Changes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables		**Male-to-female sex ratios**				**Gender gap in education**		
Famine	-0.2085***	-0.1135	-0.2554***	-0.1397	0.0122**	0.0109***	0.0113**	0.0097**
	(0.064)	(0.095)	(0.0870)	(0.0973)	(0.005)	(0.004)	(0.0044)	(0.0040)
Famine*son preference	0.2254***	0.2968***	0.2803***	0.3474***	-0.0134***	-0.0155***	-0.0123***	-0.0144***
	(0.063)	(0.072)	(0.0817)	(0.0638)	(0.004)	(0.004)	(0.0035)	(0.0032)
Famine*war duration	-0.0412**	-0.0270			-0.0007	-0.0005		
	(0.019)	(0.023)			(0.001)	(0.001)		
Famine*cultural revolution			-0.0446***	-0.0401**			-0.0010	-0.0016***
			(0.0150)	(0.0189)			(0.0009)	(0.0005)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4212	4212	4212	4212	4212	4212	4212	4212
R-squared	0.166	0.166	0.166	0.166	0.821	0.823	0.821	0.823

Notes: This table reports the robustness to including controls for other social changes. We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. War duration denotes the provincial length of the Chinese Civil War collected from Wikipedia. Cultural revolution denotes the normalized provincial number of abnormal death due to the revolution from Walder (2014) based on county gazetteers. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table A4: Robustness to Inclusion of Famine Exposure during Age 0-1

	(1)	(2)	(3)	(4)
Dependent variables	**Male-to-female sex ratios**		**Gender gap in education**	
Famine	-0.5940***	-0.4120**	0.0214***	0.0159***
	(0.139)	(0.169)	(0.005)	(0.005)
Famine*son preference	0.5138***	0.6326***	-0.0158***	-0.0192***
	(0.122)	(0.097)	(0.004)	(0.004)
Famine (0-1)	0.4743***	0.4109**	-0.0127*	-0.0065
	(0.136)	(0.181)	(0.007)	(0.008)
Famine (0-1)*son preference	-0.3326***	-0.3947***	0.0042	0.0058
	(0.114)	(0.110)	(0.005)	(0.004)
Observations	4212	4212	4212	4212
R-squared	0.168	0.169	0.822	0.824
Controls	No	Yes	No	Yes

Notes: This table reports the robustness of our analysis to including famine exposure during age 0-1. We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. Controls include the interaction terms between pre-famine regional conditions and famine exposures received in utero and during age 0-1. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table A5: Inclusion of Health Controls

	(1)	(2)
Dependent variables	**Gender gap in education**	
Famine	0.0172***	0.0155***
	(0.003)	(0.004)
Famine*son preference	-0.0158***	-0.0149***
	(0.003)	(0.003)
Famine*GDP per capita	-0.0001***	-0.0001***
	(0.000)	(0.000)
Famine*health institutions	-0.0000***	-0.0000***
	(0.000)	(0.000)
Famine* high school students	-0.0040	-0.0028
	(0.003)	(0.004)
Gender gap in height		0.0092**
		(0.003)
Observations	3492	3492
R-squared	0.843	0.844

Notes: This table reports the robustness of the estimated impact of intrauterine famine exposure and its interaction with son preference to the inclusion of controls of health. We multiply sex ratios (dependent variable) by 100. The unit of observation is at the province birth year and month level. The gender gap in height is calculated at the province birth year level from the 2010 China Family Panel Studies (CFPS). Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). Province fixed effects and birth year and month fixed effects are controlled for. All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses.***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2000 China Population Census, 2010 China Family Panel Studies (CFPS), and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table A6: Results of Absolute Changes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variables	**Adult height**				**Years of education**			
Famine	0.0331	0.1740**	0.0134	-0.0515	-0.0045	-0.0049	-0.0164**	-0.0161*
	(0.0772)	(0.0776)	(0.0970)	(0.0908)	(0.0068)	(0.0083)	(0.0059)	(0.0080)
Famine*son preference	-0.0995	0.0068	-0.0399	0.1093	0.0037	0.0011	0.0165***	0.0160***
	(0.0656)	(0.1094)	(0.0812)	(0.0845)	(0.0059)	(0.0045)	(0.0048)	(0.0036)
Famine*GDP per capita		0.3225**		-0.0313		-0.0001		0.0000
		(0.1491)		(0.1002)		(0.0001)		(0.0001)
Famine*health institutions		0.0568		-0.1510***		-0.0000*		-0.0000
		(0.0548)		(0.0441)		(0.0000)		(0.0000)
Famine* high school students		-0.0940		0.0706		0.0096**		0.0132***
		(0.0594)		(0.0424)		(0.0040)		(0.0042)
Sample	Male	Male	Female	Female	Male	Male	Female	Female
Observations	296	270	309	283	4212	4212	4212	4212
R-squared	0.623	0.636	0.626	0.636	0.894	0.895	0.943	0.943

Notes: This table reports the effect of intrauterine famine exposure and its interaction with son preference on the absolute changes of outcome variables of each gender. In columns 1 through 4, the dependent variable is adult height. The unit of observation is at the province birth year level. In columns 5-8, the dependent variable is years of education. The Unit of observation is at the province birth year and month level. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). All models are weighted by population size of each unit. Standard errors are clustered at province level and reported in parentheses. ***p<0.01, **p<0.05, *p<0.1. Data Sources: the 2010 China Family Panel Studies (CFPS), the 2000 China Population Census and the Comprehensive Statistical Data and Materials on 50 Years of New China.

Table A7: Associations between Famine Exposure and Family Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Years of education		***Party Membership***			***Sibship size***	
Dependent variables	Father	Mother	Father	Mother	Total	Male	Female
Famine	0.0488	0.0034	0.0009	0.0007	0.0119	0.0021	0.0098
	(0.0302)	(0.0149)	(0.0014)	(0.0011)	(0.0095)	(0.0085)	(0.0081)
Male	0.0747	0.1087*	-0.0023	0.0021	-0.1470***	-0.1298***	-0.0172
	(0.0803)	(0.0533)	(0.0094)	(0.0032)	(0.0459)	(0.0330)	(0.0375)
Male*famine	-0.0138	-0.0041	-0.0020	-0.0000*	-0.0070	-0.0056	-0.0014
	(0.0166)	(0.0079)	(0.0013)	(0.0000)	(0.0078)	(0.0073)	(0.0070)
Observations	5933	6442	7389	7389	7288	7288	7288
R-squared	0.030	0.063	0.014	0.008	0.047	0.028	0.023

Notes: This table reports the results of using intrauterine famine severity to predict family background and sibship size. The data used in this analysis is from the China Family Panel Survey (CFPS) 2010 and Statistical Yearbook released by the National Bureau of Statistics. Pre-famine regional control variables are provincial GDP per capita, number of health institutions, and number of high school students in 1954 (subtracted from the median values). All models are weighted by population size of each unit. All standard errors are clustered at province level. Standard errors are in parentheses. ***p<0.01, **p<0.05, *p<0.1.