Online Appendix for

Who Is Born Matters: The Demographic Origins of Inequality

Chenxiao Shou,[1] Xuebo Wang,[2]* Junsen Zhang[3]

[1] School of Business, Hunan Normal University. [2] School of Economics, Shanghai University of Finance and Economics. [3] School of Economics, Zhejiang University.

* Corresponding author. Email: wang.xuebo25@gmail.com.

This PDF file includes:

Appendix text
Figures S1 and S2
Tables S1 to S11
SI References

S1. Theoretical Discussions

Conceptually, whether and to what extent differential fertility across income classes affects income inequality largely depends on intergenerational income mobility. If the intergenerational income correlation is one—so that children have the same income levels with their parents—then differential fertility has a direct and strong effect on the income distribution of the next generation through a pure population-composition effect. At the other extreme, if the intergenerational income correlation is zero—so that children's income is uncorrelated with their parents' income—then differential fertility does not matter, because children from all family backgrounds have the same probability of ending up in any income class. In reality, intergenerational income correlation typically lies between zero and one, and children of rich and poor parents face different transition probabilities across income classes, making the relationship between differential fertility and inequality much more complex.

Differential fertility affects income inequality through its interaction with intergenerational mobility. If children's economic outcomes were independent of family background, changes in population composition would have little effect on the distribution of income in the next generation. In reality, however, children from disadvantaged backgrounds face systematically lower probabilities of upward mobility. Consequently, a larger share of births among disadvantaged groups may increase inequality by concentrating a greater fraction of the population in environments with fewer opportunities for human capital accumulation and economic advancement.

Lam (1986) analyzes the dynamics of differential fertility across income classes and income inequality of society. In a model of differential fertility and intergenerational mobility based on a Markov process governing transitions across income classes, there are *1, 2, ……, n* income classes (the income level increases with the number) with different fertility rates, and intergenerational mobility is described by a matrix M, where element M_{ij} is the probability that a child of class *j* becomes a member of class *i*. The first step to examine the effect of a higher fertility of the poor on the income distribution of the society is to investigate how a higher fertility of the poor affects the proportion of the poor. Lam first analyzes the effect of a change in F_1, the fertility of the poorest class, on the proportion in that class in the next period. He concludes that if $M_{1i} \leq M_{11} \forall i,$ implying that parents of other classes are less likely to produce poor offspring than the poor themselves, then an increase in the fertility of the poor will always lead to an increase in the proportion of the poor in subsequent period.

However, the effect of a change in F_1 on any particular inequality index is far more complicated and cannot be predicted without further information. For example, if we focus on the Gini coefficient, an

increase in the proportion of the poor will likely raise the Gini initially, but as the poor come to dominate the distribution and most people become poor, the Gini must eventually fall. Lam also shows that different inequality measures can move in opposite directions: variance of log income and the coefficient of variation respond differently both in steady state and during the transition when fertility differentials are eliminated. These results call into question the robustness of inequality measures and inequality comparisons when fertility differs across income classes.

Subsequent work (Chu, 1987; Dietzenbacher, 1989) shows that the issue is even more subtle and that even Lam's seemingly simple statement—"if $M_{1i} \leq M_{11} \forall i,$ an increase of the fertility of the poor will lead to an increase in the proportion of the poor"—does not necessarily hold without stronger conditions. For instance, suppose children of the poorest class have a very high probability (say, 70%) of moving into the richest class and only a slightly higher probability (say, 30%) of remaining poor than children of the richest class (say, 29%). In that case, higher fertility among the poor could actually increase the proportion of the rich. Chu and Koo (1990) further develop this literature and show that, under several stronger assumptions, higher fertility in the poorest class does indeed increase the proportion of the poor; moreover, the resulting income distribution is conditionally first-degree stochastically dominated by the initial distribution (in a Markov branching process of income dynamics).

If one income distribution conditionally first-degree stochastically dominates (CFSD) another, the former is unambiguously better in welfare terms: it delivers a higher level of social welfare. For example, starting from two identical income distributions, if we raise the income of the poorest class in one distribution (holding everything else constant), that distribution CFSD-dominates the other because the poorest group is strictly better off. CFSD is thus a powerful concept with stronger welfare implications than standard inequality measures. Importantly, a society with lower measured inequality but uniformly low incomes does not necessarily dominate a society with higher inequality and higher overall incomes.

Despite its theoretical appeal, testing CFSD empirically is extremely difficult. In practice, measures of income inequality such as the Gini coefficient are more policy-relevant, and both policymakers and the public focus on them. While a lower level of inequality is not always the most desirable outcome in every conceivable scenario, it is generally preferred to a higher one (aside from the extreme case where everyone is equally poor). Hence, studying income inequality remains practically important and has strong policy implications.

Although pure theory does not yield a definitive prediction for the effect of higher fertility among the poor on standard inequality measures, the existing models suggest some plausible patterns. In Lam's

framework, an increase in the share of the poor is likely to raise the Gini coefficient initially, while the Gini eventually falls only in the extreme case where the poor become the overwhelming majority. Before this extreme is reached, an increase in the proportion of the poor is therefore likely to increase, rather than decrease, inequality.

Our empirical analysis is guided by this theoretical literature. We examine the effect of a plausibly exogenous increase in rural fertility in China on the income distribution of the affected regions several decades later. According to Chu and Koo (1990), under three specific assumptions, higher fertility among the poorest group (rural households in our context) raises the proportion of the poor and yields a less favorable income distribution in the sense of lower social welfare. According to Lam (1986), a higher share of the poor is likely to increase the Gini coefficient, except in the extreme case where almost everyone is poor. Motivated by these results, we discuss and empirically test the three assumptions in Chu and Koo (1990) and estimate the causal effect of an increase in rural fertility on the Gini coefficient of income in 2005.

Macroeconomic studies have also emphasized the reverse direction of causality, where income inequality directly affects differential fertility. Kremer and Chen (2002) find that higher inequality is associated with larger fertility differentials within a country. Croix and Doepke (2003) show that an increase in income inequality can widen the fertility gap between rich and poor, effectively placing more demographic weight on families that invest less in education; this, in turn, lowers average human capital and economic growth. Our empirical findings are not inconsistent with such mechanisms: current inequality may well influence contemporaneous fertility differentials across income classes.

However, our focus is on how post-famine differential fertility affects income inequality as measured in 2005. Because it is implausible that income inequality in 2005 could have influenced fertility decisions in the 1960s, reverse causality from current inequality to past fertility is unlikely to be a major concern. That said, current inequality may still be correlated with past differential fertility through other channels. To address such potential endogeneity, we rely on plausibly exogenous shocks to differential fertility as instruments in our identification strategy.

S2. Differential Fertility and Population Composition: Theoretical Clarification

The central hypothesis of this paper is that differences in fertility across socioeconomic groups can shape the distribution of income in the next generation by altering population composition. Our empirical analysis focuses on the rural population share of a birth cohort rather than the rural–urban fertility

differential itself. This distinction is important because differential fertility and population composition, while closely related, are conceptually different objects.

Let α denote the rural population share of a given birth cohort; let N_1, N_2 denote the number of women of childbearing age (15–49) in rural and urban areas, respectively, in a given year; let n_1, n_2 denote rural and urban fertility in that year, respectively. We then have:

$$\alpha = \frac{N_1 n_1}{N_1 n_1 + N_2 n_2} = \frac{1}{1 + \frac{N_2 n_2}{N_1 n_1}}. \tag{S1}$$

Equation (S1) shows that cohort composition depends on both fertility differentials and the relative size of the childbearing-age population. Holding the latter fixed, a higher rural–urban fertility ratio implies a larger rural share among newborns. In this sense, differential fertility operates through its effect on population composition.

Our empirical objective is not to estimate the effect of fertility rates per se, but rather to identify the consequences of changes in cohort composition. This focus is motivated by the fact that population composition is the economically relevant object linking fertility behavior to long-run inequality. Regardless of whether a larger rural share arises from higher rural fertility, lower urban fertility, or other demographic forces, what ultimately matters for the distribution of income in the next generation is the composition of individuals entering that generation.

The Great Famine provides a useful setting because it generated a large and persistent increase in rural fertility relative to urban fertility, thereby substantially increasing the rural share of post-famine birth cohorts. We exploit this famine-induced variation in cohort composition to examine how changes in "who is born" affect long-run income inequality and intergenerational mobility.

S3. Cross-Sectional Evidence

We provide suggestive evidence that the famine may have significant increased the rural population share and income inequality among post-famine birth cohorts. This section provides cross-sectional evidence of these effects at the prefectural level, examining the famine's long-term impact on demographic composition and income inequality.

Because prefectural-level mortality data for the famine period are unavailable, Meng et al. (2015) use the 1990 census to construct a proxy for famine severity at the county level: the size of surviving birth cohorts during the famine within the agricultural population. They show that cohort size for 1959–1961

is negatively correlated with famine severity, as it reflects both reduced fertility and increased mortality during the famine, and argue that this proxy has several advantages over mortality-based measures.

Following Meng et al. (2015), we first obtain, from the 1990 census, the rural birth cohort size of each year from 1950 to 1970 for every prefecture and then fit a trend line of rural cohort size over this period. We next calculate, for the famine years 1959–1961, the gap between the actual and fitted rural cohort sizes and construct our famine severity measure as the ratio of this gap to the corresponding fitted value at the prefectural level. Figure S1 plots the prefecture-level average rural birth cohort size for each year from 1950 to 1970 (solid line) together with the fitted trend (dotted line). The birth cohort size falls sharply during the famine years, generating a substantial gap between the actual and trend values. The average rural birth cohort size gap over 1959–1961 therefore captures the extent to which rural cohort size shrank during the famine and serves as a proxy for famine severity.

We also compute the analogous rural birth cohort size gap for each province and compare it with the rural excess mortality rate during the famine—an alternative measure of famine severity. The correlation between these two variables is as high as 0.85 at the provincial level, supporting the validity of the rural cohort size gap as an indicator of famine severity. We thus use this rural birth cohort size gap as an instrument for the post-famine rural population share of the 1962–1985 birth cohorts and estimate the effect of the rural population share on the 2005 Gini coefficient of these cohorts at the prefectural level.

We begin by estimating the following equation, which relates famine severity to either the rural population share or the Gini coefficient of post-famine birth cohorts at the prefectural level:

$$Y_p = c + \beta FM_p + \delta_2 X_p + \varepsilon_{p2}, \quad \text{(S2)}$$

where Y_p represents the dependent variable of interest (either the rural population share or the Gini coefficient) for post-famine birth cohorts (1962–1985) in prefecture p, FM_p denotes famine severity, measured by rural birth cohort size gap during the famine in prefecture p, X_p is a vector of prefectural level control, c is a constant term, and ε_{p2} represents the error term. The coefficient β is the parameter of interest.

We include a set of prefectural characteristics in 2005 in the regressions, as these factors may directly affect income inequality. Specifically, we control for indicators of regional economic development, including income per capita, GDP, the shares of agricultural and industrial output in GDP, and the share of migrants in the total population. In addition, because population density and social spending can influence income inequality, we control for population density and fiscal expenditure per capita in 2005.

Table S1 reports the results. Columns (1)–(3) present the coefficients for the rural population share, and Columns (4)–(6) present those for the Gini coefficient. The correlations are statistically significant and remain stable and robust regardless of the inclusion of controls. The results suggest that greater famine severity is associated with a higher rural population share and a higher Gini coefficient among post-famine birth cohorts.

Famine severity is clearly endogenous and correlated with regional characteristics. While a DID framework can absorb time-invariant regional differences through prefecture fixed effects, the cross-sectional OLS estimates above can only be interpreted as correlations rather than causal effects. To address this concern, we use climate shocks during the famine as an instrument for famine severity. As established in the literature, climate disasters played a nonnegligible role in the sharp decline in agricultural productivity during the famine (Kueh, 1995; Yao, 1999; Lin and Yang, 2005). Extreme weather events likely reduced regional grain output and thereby affected famine severity in 1959–1961. Moreover, because extreme weather can be viewed as an exogenous shock to agricultural production, it is unlikely to affect long-run outcomes through channels unrelated to the famine.

In practice, we construct a dummy variable indicating whether a rainfall shock occurred in a prefecture during the famine. Following the methodology widely used in the literature (e.g., Shah and Steinberg, 2017; Addoum et al., 2020; Corno et al., 2020), we define a rainfall shock as annual rainfall below the 5th percentile or above the 95th percentile of the prefecture's long-term rainfall distribution over 1950–1966. The weather data are taken from Meng et al. (2015). The largely random occurrence of such rainfall shocks during the famine makes this variable a valid instrument for famine severity, as weather shocks are unlikely to influence long-term income inequality through channels other than their impact on agricultural production and famine severity.

Previous work shows that the sharp decline in food production and high state procurement from rural areas were the primary causes of the famine (Thaxton, 2008; Meng et al., 2015). To understand how extreme weather affected famine severity, we examine its impact on grain output and government grain procurement during the famine.

Because data on grain output and procurement are available only at the provincial level, we employ a Difference-in-Differences (DID) framework using a provincial panel from 1953 to 1961. We compare per capita grain output and per capita grain procurement between a treatment group (provinces that experienced extreme weather shocks during the famine) and a control group (provinces that did not),

before and during the famine. This design allows us to identify the impact of extreme weather on grain production and procurement during the famine years. Specifically, we estimate:

$$lny_{pt} = a + \alpha_p + \gamma_t + b\left(GF_t \times RainShock_p\right) + \varepsilon_{pt}, \quad \text{(S3)}$$

where lny_{pt} represents the logarithm of outcome variable (either per capita grain output or grain procurement) for province p in year t; GF_t is a dummy variable that equals 1 if year t falls within the famine period (1959–1961) and 0 otherwise; $RainShock_p$ is a dummy indicating whether province p experienced weather shocks during the famine; α_p and γ_t are province and time fixed effects, respectively; and ε_{pt} is an error term. b is the parameter of interest.

Table S2 reports the results. As shown in Columns (1) and (2), the interaction coefficients are significantly negative for both outcomes. These findings indicate that extreme weather during the famine substantially reduced per capita grain output, and the government correspondingly reduced grain procurement from affected regions. As a result, provinces hit by extreme weather retained more grain and appear to have suffered less from the famine, which explains why extreme weather shocks are negatively correlated with our measure of famine severity.

Using extreme weather shocks in 1959–1961 as an instrument for famine severity, we then estimate the impact of the famine on the rural population share and the income inequality of post-famine birth cohorts in 2005 at the prefectural level. Specifically, we estimate the following 2SLS model:

$$M_p = c_1 + bRainShock_p + \delta_1 X_p + \varepsilon_{p1}, \quad \text{(S4)}$$

$$Y_p = c_2 + \beta FM_p + \delta_2 X_p + \varepsilon_{p2}, \quad \text{(S5)}$$

where most variables are the same as in Equation (S2). The new variable, $RainShock_p$, is a dummy that equals 1 if prefecture p experienced extreme rainfall shocks during the famine and 0 otherwise. β is the parameter of interest, which aims to capture the causal effects of the famine on either the rural population share or the Gini coefficient of post-famine birth cohorts.

Table S3 reports the estimates. Columns (1)–(3) show the effects of famine severity on the rural population share of the 1962–1985 birth cohorts. The results indicate that the famine significantly increased the rural population share in the post-famine period. In particular, the IV estimates in Columns (2) and (3) imply that a 10% increase in the average rural birth cohort size gap during the famine is associated with roughly a 0.1 increase in the post-famine rural population share—a substantial effect.

Columns (4)–(6) report the corresponding effects on the Gini coefficient for the 1962–1985 cohorts. The estimates suggest that the famine also had a sizable impact on income inequality within these cohorts.

Specifically, the IV estimate in Column (6) indicates that a 10% increase in the average rural birth cohort size gap during the famine corresponds to an approximately 0.03 increase in the Gini coefficient, signifying a nonnegligible effect on income inequality.

As noted, weather data is available for only 139 prefectures, representing less than half of all prefectures, raising concerns about potential sample selection bias in our analysis. To address this issue, we provide evidence that these 139 prefectures in the IV regression sample do not significantly differ from the full sample in most observable characteristics.

In Table S4, we compare the 139 prefectures and the full sample across various characteristics, including the Gini coefficient, rural birth population share, rural birth cohort size gap, GDP, income per capita, fiscal expenditure per capita, population density, immigrant population share, agricultural and industrial output share in GDP. For most variables, the mean differences between the IV regression sample and the full sample are not statistically significant. The only exceptions are population density and the agricultural output share in GDP, both of which show statistically significant differences at the 5% level. Overall, the evidence suggests that the selection bias in the IV regression sample is minimal.

S4. The Famine's Effects on Population Composition and Income Inequality

This section evaluates whether the post famine rise in the rural population share is a plausible mechanism for the observed increase in cohort inequality, and how this mechanism relates to other direct and indirect channels.

We begin by examining the dynamics of the famine's impact on the population composition of post-famine birth cohorts. We estimate Equation (1) in the main text and present the b_t coefficients in Figure 6. Each dot on the solid line represents the coefficient on the interaction between a birth-cohort dummy and famine severity, with the 95% confidence intervals shown by the dashed lines.

Figure 6 shows that the estimates for the 1951–1956 birth cohorts are close to zero and statistically insignificant, but they drop sharply for the 1957 cohort and become significantly negative. After the famine ends, the estimates rise sharply, become significantly positive, and remain elevated for an extended period. They begin to decline in the mid-1970s and eventually diminish to zero for cohorts born after 1980. These dynamics indicate a persistent, cohort-specific compositional effect of the famine that gradually dissipates over time.

Note that near-zero coefficients for the 1951–1956 cohorts (with 1950 cohort as the reference) do not imply the famine had no demographic impact on those cohorts. Rather, they suggest that the rural shares

of the 1951–1956 cohorts evolved similarly to that of the 1950 cohort. Existing research shows that famine excess deaths were heavily concentrated among very young children—especially infants under age one—and were much lower among those older than five (Peng, 1987; Song, 2013). Most members of the 1950–1956 cohorts were beyond infancy during the famine, so their direct mortality exposure was likely limited. By contrast, the large positive estimates for post-famine cohorts point to a pronounced increase in their rural population share relative to pre-famine cohorts, consistent with a compositional consequence of the famine. Thus, while direct mortality effects on the 1950–1956 cohorts appear limited, larger compositional shifts are concentrated among cohorts born after the famine. Accordingly, the 1950–1956 cohorts constitute a reasonable control group for post-famine cohorts.

Table 2 in the main text presents the DID estimates of the effect of the rural population share on the Gini coefficient for the 1962–1980 birth cohorts. Column (2) reports estimates without controls, and Column (3) reports results with controls. Ideally, the controls would be prefecture-level characteristics measured prior to the famine. Unfortunately, pre-famine prefectural data are generally of low quality or simply unavailable. We therefore control for prefecture-level characteristics measured in 2005. A potential concern is that these characteristics may themselves be affected by the famine, rendering them "bad controls." This concern is mitigated, however, by the fact that our estimates are not sensitive to their inclusion: the results in Columns (2) and (3) are quantitatively similar. Thus, while the controls are imperfect, their imperfection does not threaten the validity of our main findings.

We next examine the famine's impact on cohort inequality in 2005 by estimating Equations (2) and (3) separately for pre- and post-famine cohorts; the results are reported in Table S5. All cohort-specific estimates—1950–1956, 1962–1967, 1968–1973, 1974–1980, and 1981–1985—are positive and statistically significant. Quantitatively, the coefficient for the 1950–1956 cohort is about 0.22: a 10-percentage-point increase in the average rural birth cohort size gap during the famine is associated with roughly a 0.022 increase in the Gini coefficient. The estimates for the 1962–1980 cohorts are larger (around 0.30), while the coefficient for 1981–1985 is smaller but remains significant, indicating a persistent yet attenuated famine effect.

As noted above, the population compositions of the 1950–1956 and 1981–1985 cohorts were minimally affected by the famine. The fact that we still find significant effects on their Gini coefficients indicates that the famine influenced income inequality through channels other than population composition. Importantly, the larger effects for early post-famine cohorts (1962–1980), who did not directly experience the famine, relative to pre-famine cohorts (1950–1956), who did, are not what one

would expect from direct and indirect channels alone. As discussed in the main text, direct and indirect channels (such as health, education, and labor market impacts on those exposed in childhood) should, if anything, generate a larger effect of the famine on the Gini coefficient for pre-famine cohorts than for post-famine cohorts. Absent the population compositional channel, we would therefore expect the famine's effect on inequality to be greater for pre-famine cohorts than for post-famine cohorts, ceteris paribus. The opposite pattern in our estimates—larger coefficients for post-famine cohorts than for pre-famine cohorts—suggests that the population compositional effect dominates, and that our estimated effects for post-famine cohorts represent a lower bound on the contribution of this channel to income inequality.

The results in Table S4 are consistent with this interpretation. The smaller coefficient for the 1950–1956 cohort points to a modest effect of the famine on their income distribution, while the larger coefficients for the 1962–1980 cohorts indicate a stronger effect that is difficult to reconcile without a substantial compositional channel. As the famine's impact on the rural population share fades for later cohorts, the coefficients decline and approach the pre-famine level for the 1981–1985 cohort, for whom the compositional effect is minimal. These patterns support our hypothesis that the population compositional effect is the dominant driver of income inequality among post-famine cohorts. In short, the dominance of this compositional channel explains the larger impact for earlier post-famine cohorts and underpins the plausibility of our IV–DID strategy.

S5. Discussions on Migrations across Regions

As discussed in the main text, we use the 1990 census data to calculate the rural population share of each prefecture, based on respondents' hukou status at the time of the census. To calculate the Gini coefficient, we use the 2005 census data, relying on respondents' place of residence and workplace. All current labor force members working in each prefecture are included to derive the 2005 Gini coefficient for that prefecture. However, since individuals born in one prefecture may migrate to others over time, this approach might have limitations. Although migration was uncommon before 1990, there was a steady inflow of migrants from the western and central regions to the more developed eastern regions by 2005. Given that some migrant workers in each prefecture in 2005 may have originated from other areas, this could pose challenges to our strategy. To address this issue, we use information on respondents' migration status from the 2005 census data.

In the survey of the 2005 census, question "R6" is related to the respondents' hukou registration location. If the respondents' hukous were registered in where they were currently residing, then, we identify them as local residents. By contrast, if the respondents' hukous were registered in other places rather than where they were currently living, we identify them as migrant population and further take the registration location of their hukou as their hometown.

Such an identification of the respondents' migrant statuses may not be completely accurate. For example, if a respondent was born in one prefecture, later successfully admitted to college in another prefecture, obtained a hukou after graduation, and lived there permanently, then, we include her in the sample of the prefecture of her residence rather than her home prefecture in 2005. Theoretically, we should include all the respondents who were born during 1962–1985 in the prefecture in the sample of this prefecture in 2005, and such an inaccurate choice of the sample may induce problems. However, given Chinese government's extremely tight control on the hukou system over a long period, it is very difficult for most people to change their hukou statuses, particularly switch hukou from a region to another. Given that the most typical migration in China was that people left their hometown and worked in other cities without obtaining the hukou there, we can precisely identify the migration statuses of these people with the above method in the census data.

If we include those migrant workers in their home prefectures, then we should also consider the difference in living cost between their residences and their hometowns. For example, a migrant worker from Shaanxi worked in Shanghai and earned a 3000 Yuan monthly salary, while similar workers in his hometown can only earn 1500 Yuan per month. In this case, we cannot say that this worker's salary is twice as high as that of workers in his hometown because the cost of living in Shanghai is also substantially higher than that back home. Therefore, given the difference in the cost of living across prefectures, those migrant workers' income may be incomparable to that in their home prefectures.

We thus further adjust the price differences across prefectures to make the income earned in different regions comparable. Using the spatial price index for all prefectures in 2004, calculated by Brandt and Holz (2006), we adjust the nominal income in all prefectures accordingly. We then re-estimate Equations (2) and (3) in the main text, incorporating these migration and price adjustments, and present the results in Table S6. The results remain consistent with the benchmark findings in Table 2 of the main text, confirming the robustness of our analysis.

S6. More DID Results

As shown in Figure 6, the population compositions of the 1981–1985 birth cohorts were unaffected by the famine, making them a suitable choice for the control groups. Therefore, We perform additional DID regressions using the 1962–1980 and 1981–1985 birth cohorts—whose population composition was and was not affected by the famine, respectively—as the treatment and control groups. The results presented in Table S7 indicate that, compared to the 1981–1985 birth cohorts, the famine induced a much higher rural population share for the 1962–1980 birth cohorts, which, in turn, led to a higher Gini coefficient for them.

To further validate these findings, we conduct a placebo test, using the 1950–1956 and 1981–1985 birth cohorts—both of which had population compositions largely unaffected by the famine—as the treatment and control groups, respectively. Table S8 presents the corresponding estimation results, including both the reduced-form estimates and the IV regression results. The results indicate that both the first- and second-stage coefficients are close to zero and statistically insignificant, and the reduced-form estimate is also not significant. These results further confirm that population composition could be an important mechanism through which the famine has affected long-term income inequality.

S7. Overidentification Test

To further verify the validity of the famine as an IV for the post-famine fertility structure, we present another IV to perform an overidentification test. As shown in the literature, China's population control policy is more strictly implemented in urban areas than the rural ones, and such a two-tier population policy also induces a much higher rural fertility than the urban one (Zhang, 2017; Wang and Zhang, 2018). Therefore, we can use the implementation intensity of the population control policy as another IV for China's rural–urban fertility structure or population composition. The population control policy could be a valid IV because it directly affects fertility and is unlikely to affect the income inequality of affected cohorts by affecting other factors rather than fertility, such as institutions and regional productivity several decades later.

Given that China's one-child policy (OCP) was implemented after 1979 and the famine did not affect the population composition of the post-1980 birth cohorts, the OCP is not a feasible IV for the overidentification test. However, even before the implementation of the OCP, China has begun a voluntary yet strong family planning campaign with the slogan "Later, Longer, and Fewer" (LLF) in the early 1970s.

This policy was very successful, and China's overall fertility rate was halved during 1970 and 1978 (Zhang, 2017). Given that the famine also affects the rural–urban fertility structure of the 1970–1978 birth cohorts, we can use both the famine severity and implementation intensity of the LLF policy to instrument for the population composition of these birth cohorts to perform the overidentification test.

The LLF policy varies only at the provincial level, therefore, its implementation intensity was uniform across all prefectures within the same province. We can conduct the overidentification test at the prefectural level. Specifically, we estimate the following cross-section regression at the prefectural level:

$$Rshare_p = c_1 + \mathrm{b}FM_p + dL_p + \theta_1 X_p + \varepsilon_{p1}, \quad \text{(S6)}$$

$$Gini_p = c_2 + \beta\, Rshare_p + \theta_2 X_p + \varepsilon_{p2}. \quad \text{(S7)}$$

These regressions are similar to the Equations (2) and (3) in the main text; however, they differ from those equations, as we perform a cross-sectional regression here rather than a DID regression with panel data. L_p represents the implementation intensity of the LLF policy in prefecture p, and X_p are control variables.

In practice, we use the birth planning program timing to proxy for the implementation intensity of the LLF policy at the provincial level. As shown in Babiarz et al. (2018), the LLF policy timing shows significant variation across provinces. Table S9 demonstrates the LLF implementation time for different provinces. Some provinces initiated the policy as early as in 1970, and some others launched the campaign after 1975. Given that we focus on the 1970–1978 birth cohorts, we can use the LLF policy timing to construct the variable of the policy implementation intensity for these cohorts. Specifically, if the policy was initiated in one province in 1970, then, all the 1970–1978 birth cohorts in this province were affected. Thus, we assign the number 1 as the policy implementation intensity to this province. Similarly, if a province initiated the policy in 1971, then 8 (1971–1978) out of 9 (1970–1978) birth cohorts were affected. Thus, we assign the number 8/9 as the policy implementation intensity to this province, and so on.

We estimate Equations (S6) and (S7) and report the results in Column (1) and (2) of Table S10. Column (1) shows the IV estimates without controls and the coefficient of rural population share is 0.107 and significant at the 5% level. Column (2) reports the IV estimates with controls and the corresponding coefficient is 0.140 and significant at the 1% level. The bottom Panel presents the results of overidentification test. The *p* values is much larger than 0.1 in Column (1) and nearly 0.10 in Column (2), which indicates that we cannot reject the null hypothesis that the IVs satisfy the exclusion restrictions.

In the above analysis, we assume that the birth planning policy affects the fertility immediately after its implementation. However, given the ten-month child-bearing period, the policy may have a time lag

between its implementation and being effective to reduce fertility. For instance, many women may have already been pregnant when the policy was announced. Thus, the fertility will not decrease immediately in the future months. The policy is likely to reduce fertility effectively after 10 months or one year of its implementation. Therefore, we now assume that the policy affects the fertility after one year of its announcement and calculate the policy implementation intensity for the 1971–1978 birth cohorts and redo regressions (S6) and (S10).

Column (3) and (4) of Table S10 report the estimate results. Compared to Column (1) and (2) of Table 10 , the estimates in Column (3) and (4) show little change, which confirms the robustness of our identification.

S8. Famine Severity and the Implementation Intensity of the One-Child Policy (OCP)

Other factors, such as population control policies, could also influence the post-famine rural–urban fertility structure. However, as long as these factors are not systematically correlated with famine severity, they should not pose a serious concern for our analysis. To verify this, we provide direct evidence on the relationship between famine severity and the implementation intensity of population control policies.

Panel A of Figure S2 plots the implementation intensity of the OCP—measured by fines for excess fertility in 1979—against famine severity—measured by the average rural excess mortality rate in 1959–1961—for all provinces. In 1979, when the OCP was first introduced, fines for excess fertility were largely uniform across provinces and thus uncorrelated with famine severity.

Panel B plots the fines for excess fertility in 1985 against famine severity for all provinces. By 1985, after six years of OCP implementation, fines varied considerably across provinces. However, the points are highly dispersed, and there is no apparent relationship between the two variables.

In sum, we have no reason to expect any systematic correlation between famine severity and the implementation intensity of the population control policy, and the evidence from these figures confirms that the two variables are not systematically related.

S9. The OCP as an IV for Rural Population Share

As noted, China's one-child policy (OCP) was more strictly implemented in urban areas than in rural ones. Such a two-tier population control policy may also induce a much higher rural fertility than the urban one. Therefore, we can use the OCP as an exogeneous shock on China's post-1980 fertility structure and population composition to identify the effect of rural population share on income inequality several

decades later. Specifically, we use the implementation intensity of the OCP to instrument for the rural population share of the 1982–1985 birth cohorts at the provincial level to identify the effect of such plausible exogeneous variations in the rural population share on the rural–urban ratio of the probability of gaining admission to senior high school.

In practice, we use the fines due to the breach of the OCP to measure the implementation intensity of the policy at the provincial level. Under the OCP, households who exceed their fertility limit are forced to pay a fine (usually several years of household income) and are subject to a variety of other monetary punishments, such as the seizure of property and forced dismissal from government employment (Ebenstein, 2010). Such punishments are generally much more effective in urban areas than in rural ones. As discussed in Zhang (2017), rural residents received few benefits from the government; thus, they had nothing or little to lose, and the penalty (e.g., a fine) is typically ineffective because many rural families are too poor to pay them. Consequently, even if rural and urban households face the same fines, they impacts were different.[1] Intuitively, a higher fine may dramatically reduce urban fertility but only mildly cut down rural fertility, leading to a higher rural–urban fertility ratio or rural population share.

We estimate the following fixed-effects model：

$$Rshare_{pt} = c_1 + \alpha_{1p} + \gamma_{1t} + bFine_{pt} + \varphi_1 \mathrm{R}_{pt} + \varepsilon_{pt1}, \quad \text{(S8)}$$

$$Ruratio_{pt} = c_2 + \alpha_{2p} + \gamma_{2t} + \beta Rshare_{pt} + \varphi_2 \mathrm{R}_{pt} + \varepsilon_{pt2}, \quad \text{(S9)}$$

where $Fine_{pt}$ is the fine as measured in years of household income of province p in year t (1982–1985), and other variables are similar to those in Equations $(2')$ and $(3')$ in the main text.

A potential concern is that OCP implementation intensity may be correlated with provincial characteristics that also affect educational outcomes. To mitigate this concern, the specification includes province fixed effects, which absorb time-invariant differences across provinces, as well as cohort fixed effects that capture nationwide changes over time. The resulting estimates should therefore be interpreted primarily as an alternative source of plausibly exogenous variation in cohort composition rather than as the main identification strategy.

The purpose of this exercise is not to establish a new causal effect of the OCP itself, but to verify whether an independent fertility shock generates results similar to those obtained using famine-induced variation in population composition.

[1] Rural and urban families faced the same fines as measured by years of household income within each province. However, given that urban household income was much higher than the urban one, urban families generally paid a higher fine in absolute amount than rural families.

Figures and Tables

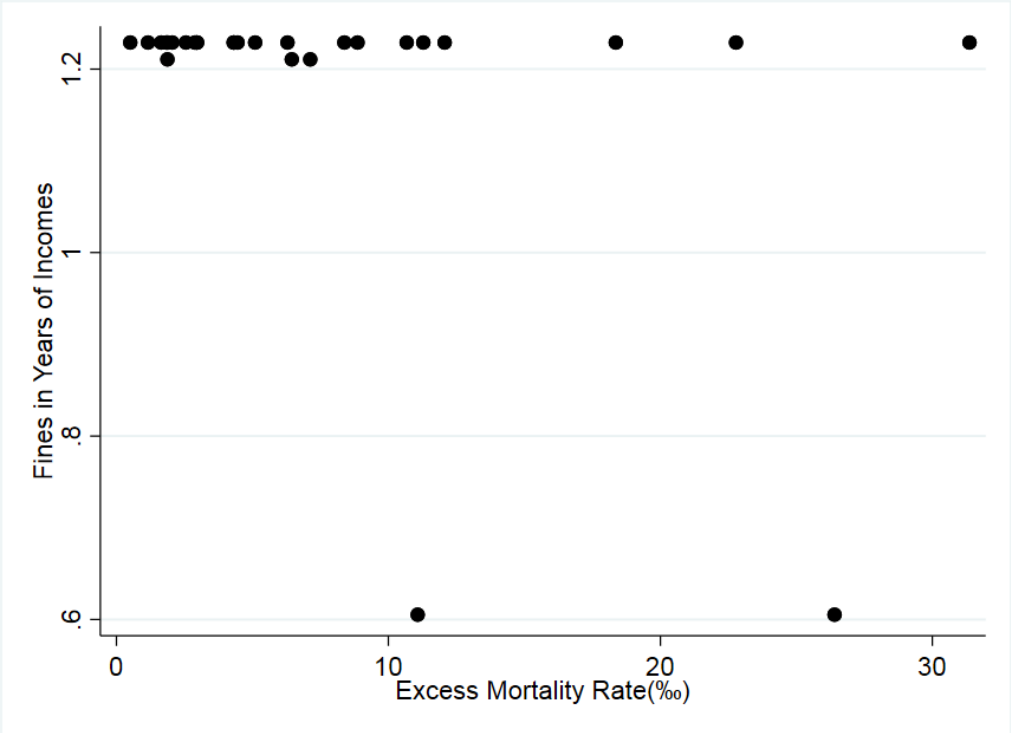
Figure S1: Rural Birth Cohort Size Gap during the Great Famine (1959–1961)
(Prefectural Level)



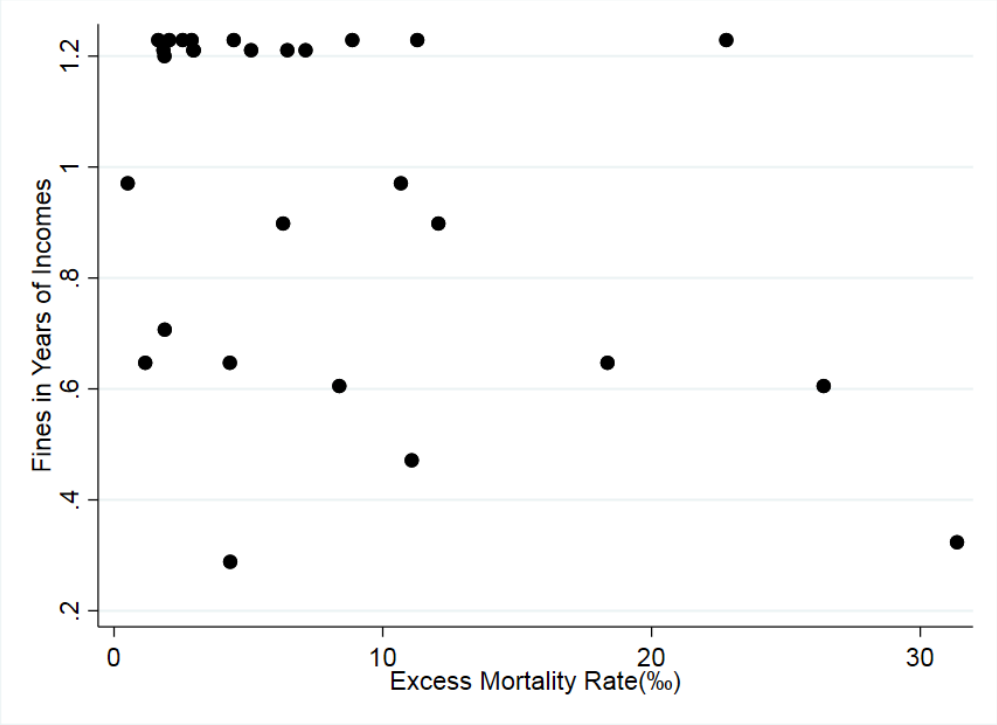
Note: This figure shows rural birth cohort size gap during the famine. Data are from Meng et al. (2015).

Figure S2: Fines for Excess Fertility and the Average Rural Excess Mortality Rate in 1959–1961 (Provincial Level)

Panel A: Fines in 1979



Panel A: Fines in 1985



Note: This figure shows the correlation between famine severity and the implementation intensity of the one-child policy, measured by fines for excess fertility. Panels A and B present results using fines in 1979 and 1985, respectively. Data on fines are from Ebenstein (2010).

Table S1: Correlation between Famine Severity and the Rural Population Share and the Gini Coefficient of Post-Famine Birth Cohorts

	Rural Population Share			Gini Coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Famine Severity	**0.372*****	**0.295*****	**0.319*****	**0.036***	**0.032***	**0.037****
	(0.062)	(0.062)	(0.066)	(0.020)	(0.018)	(0.019)
GDP		0.016	0.017		-0.001	0.000
		(0.019)	(0.019)		(0.005)	(0.005)
Income		-0.032	-0.052		-0.110***	-0.117***
		(0.043)	(0.047)		(0.015)	(0.015)
Expenditure		-0.114***	-0.165***		0.038***	0.027***
		(0.028)	(0.036)		(0.008)	(0.010)
Density		-0.011	-0.018		-0.003	-0.007*
		(0.014)	(0.016)		(0.004)	(0.004)
Immigrant			0.362***			0.005
			(0.106)			(0.045)
Primary			-0.072			-0.210***
			(0.379)			(0.037)
Secondary			0.032			-0.145***
			(0.187)			(0.030)
Observations	291	291	291	291	291	291
R-squared	0.117	0.255	0.276	0.011	0.329	0.388

Notes: This table examines the correlations between famine severity and the rural population share and the Gini coefficient of post-famine birth cohorts (the 1962–1985 birth cohorts). Specifically, it presents estimates of the coefficient β from Equation (S2). The dependent variables are the rural population share and the Gini coefficient for the1962–1985 birth cohorts, respectively. Control variables include a set of prefectural level characteristics in 2005, such as GDP, income per capita (Income), fiscal expenditure per capita (Expenditure), population density (Density), immigrant share of the total population (Immigrant), and the shares of agricultural and industrial output in GDP (Primary and Secondary). Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table S2: DID Estimates of the Effects of Extreme Weather Shocks on Per Capita Grain Output and Per Capita Grain Procurement

	(1) Grain Output	(2) Grain Procurement
GF×Rainfall Shock	-0.161**	-0.176*
	(0.072)	(0.098)
Prefecture FE	Y	Y
Year FE	Y	Y
Observations	261	252
R-squared	0.744	0.772

Notes: This table reports the DID estimates of the effects of extreme weather shocks on per capita grain output and per capita grain procurement during the famine, specifically presenting estimates of the coefficients from Equation (S3). The sample period is 1953-1961. The control variables include the prefecture and year fixed effects. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0$.

Table S3: IV Estimates of the Effect of the Famine on the Rural Population Share and the Gini Coefficient

	Rural Population Share			Gini Coefficient		
	RF	IV	IV	RF	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Famine Severity		**0.963*****	**0.891*****		**0.250*****	**0.261*****
		(0.228)	(0.196)		(0.078)	(0.068)
		First Stage			First Stage	
Rainfall Shock	-0.158***	-0.169***	-0.178***	-0.046***	-0.169***	-0.178***
	(0.037)	(0.028)	(0.026)	(0.011)	(0.028)	(0.026)
GDP	0.002	0.028	0.029	-0.004	0.003	0.004
	(0.021)	(0.031)	(0.026)	(0.007)	(0.010)	(0.009)
Income	-0.057	0.084	0.033	-0.120***	-0.071***	-0.094***
	(0.066)	(0.067)	(0.062)	(0.019)	(0.021)	(0.020)
Expenditure	-0.078	-0.117**	-0.195***	0.028*	0.031**	-0.006
	(0.052)	(0.049)	(0.048)	(0.015)	(0.014)	(0.019)
Density	-0.014	-0.060**	-0.050*	-0.015	-0.022**	-0.026**
	(0.030)	(0.030)	(0.030)	(0.009)	(0.010)	(0.010)
Immigrant	0.317		1.062***	0.020		0.238**
	(0.195)		(0.269)	(0.054)		(0.115)
Primary	0.071		0.234	-0.323***		-0.275**
	(0.300)		(0.312)	(0.084)		(0.108)
Secondary	-0.062		-0.022	-0.192***		-0.180**
	(0.184)		(0.216)	(0.054)		(0.071)
K-P F statistic		35.396	46.144		35.396	46.144
Observations	139	139	139	139	139	139

Notes: This table reports the IV estimates of the famine's effect on the rural population share and the Gini coefficient of post-famine birth cohorts (the 1962–1985 birth cohorts). Specifically, it presents estimates of the coefficients from Equations (S4) and (S5). The dependent variables are the rural population share and the Gini coefficient for the1962–1985 birth cohorts, respectively. The control variables are the same as those in Table S2. The table includes both reduced form (RF) and IV estimation results. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table S4: Balance Test between the Full Sample and IV Sample of Prefectures

Variables	Full Sample		IV Sample		Difference	P Value
	Obs.	Mean	Obs.	Mean		
Gini	291	0.396	139	0.395	0.001	0.886
Rshare	291	0.772	139	0.771	0.001	0.954
Famine severity	291	0.405	139	0.397	0.008	0.632
GDP	291	3.538	139	3.723	-0.185	0.103
Income	291	6.165	139	6.183	-0.018	0.623
Expenditure	291	7.225	139	7.234	-0.009	0.874
Density	291	5.320	139	5.593	-0.273	0.030
Immigrant	291	-0.007	139	-0.009	0.003	0.778
Primary	291	0.186	139	0.163	0.023	0.046
Secondary	291	0.444	139	0.460	-0.015	0.226

Table S5: IV Estimates of the Effect of the Famine on the Gini Coefficient, by Birth Cohort

	1950–1956		1962–1967		1968–1973		1974–1980		1981–1985	
	RF	IV	RF	IV	RF	IV	RF	IV	RF	IV
		(3)								
Famine Severity		0.222***		0.303***		0.294***		0.256***		0.207***
		(0.079)		(0.079)		(0.086)		(0.073)		(0.074)
		1st Stage		1st Stage		1st Stage		1st Stage		1st Stage
Rainfall Shock	-0.038***	-0.173***	-0.052***	-0.173***	-0.051***	-0.173***	-0.044***	-0.173***	-0.036***	-0.173***
	(0.011)	(0.028)	(0.012)	(0.028)	(0.013)	(0.028)	(0.011)	(0.028)	(0.011)	(0.028)
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
K-P F statistic		39.093		39.093		39.093		39.093		39.093
Observations	139	139	139	139	139	139	139	139	139	139

Notes: This table presents the IV estimates of the famine's effect on the Gini coefficient for various birth cohort groups: 1950–1956, 1962–1967, 1968–1973, 1974–1980, and 1981–1985. . Specifically, it presents estimates of the coefficients from Equations (S4) and (S5). The control variables are the same as those in Table S2. The table includes both reduced form (RF) and IV estimation results. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table S6: DID Estimates of the Effect of the Rural Population Share on the Gini Coefficient with Migration and Price Adjustments

	Dependent variable: Gini Coefficient		
	RF	IV	IV
	(1)	(2)	(4)
Rshare		**1.037*****	**0.881*****
		(0.289)	(0.281)
		First Stage	
Famine Severity×*Post*	**0.038*****	**0.042*****	**0.044*****
	(0.010)	(0.008)	(0.009)
Controls	Y	N	Y
Prefecture FE	Y	Y	Y
Year FE	Y	Y	Y
F statistic		28.841	25.175
Observations	7306	7306	7306

Notes: This table reports DID estimates of the effect of the rural population share on the Gini coefficient for the 1962–1980 birth cohorts with migration and price adjustments, using the 1950–1956 birth cohorts as the control group. It presents the estimated coefficients (b) (first stage) and (β) (second stage) from Equations (2) and (3) in the main text. The control variables are the same as in Table 1 in the main text. Robust standard errors are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table S8: Alternative DID Estimates of the Effect of the Rural Population Share on the Gini Coefficient for the 1962–1980 Birth Cohorts

	Dependent variable: Gini Coefficient		
	RF	IV	IV
	(1)	(2)	(4)
Rshare		**0.485****	**0.862*****
		(0.191)	(0.227)
		First Stage	
Famine Severity×*Post*	**0.049*****	**0.055*****	**0.057*****
	(0.010)	(0.008)	(0.009)
Controls	Y	N	Y
Prefecture FE	Y	Y	Y
Year FE	Y	Y	Y
F statistic		43.536	37.104
Observations	6648	6648	6648

Notes: This table reports the DID estimates of the effect of the rural population share on the Gini coefficient for the 1962–1980 birth cohorts, using the 1981–1985 birth cohorts as the control group and the 1962–1980 birth cohorts as the treatment group. The control variables are the same as those in Table 1. Robust standard errors are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table S9: Placebo Test Results for the Effect of the Rural Population Share on the Gini Coefficient

	Dependent variable: Gini Coefficient	
	RF	IV
		(3)
Rshare		-0.120
		(0.958)
		1st Stage
Famine Severity× $Post$	0.002	-0.014
	(0.013)	(0.011)
Controls	Y	Y
Prefecture FE	Y	Y
Year FE	Y	Y
F statistic		1.743
Observations	3,312	3,312
R-squared	0.589	0.112

Notes: This table reports the placebo test results for the effect of the rural population share on the Gini coefficient, using the 1980–1956 birth cohorts as the control group and the 1981–1985 birth cohorts as the treatment group. The control variables are the same as those in Table 1. Robust standard errors are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table S10: LLF Birth Planning Program Timing

Year	Num. of Provinces	Remarks
1970	3	Jiangsu, Guangdong, Hainan
1971	7	Liaoning, Jilin, Guangxi, Chongqing, Sichuan, Guizhou, Gansu
1972	8	Tianjin, Hebei, Heilongjiang, Jiangxi, Shandong, Hubei, Yunnan, Qinghai
1973	5	Shanxi, Shanghai, Fujian, Shaanxi, Ningxia,
1974	3	Anhui, Henan, Hunan
1975	1	Xinjiang
1979	1	Inner Mongolia

Table S11: Overidentification Test Results

	Current Period: 1970-1978		One Period Lag: 1971-1978	
	Without Controls	With Controls	Without Controls	With Controls
	(1)	(2)	(3)	(4)
Rshare	0.107**	0.140***	0.108**	0.143***
	(0.046)	(0.048)	(0.047)	(0.048)
	First Stage		First Stage	
Famine Severity	0.383***	0.360***	0.385***	0.361***
	(0.060)	(0.069)	(0.060)	(0.069)
Birth Control Intensity	0.053	0.089	0.050	0.089*
	(0.055)	(0.056)	(0.051)	(0.053)
Controls		Y		Y
Prefecture FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Kleibergen-Paap F statistic	20.411	15.466	20.480	15.518
Observations	288	223	288	223
Hansen J statistic	0.130	2.834	0.085	2.479
P Value	0.719	0.092	0.770	0.115

Notes: This table reports the results of the overidentification tests, based on estimates from Equations (S6) and (S7). The sample period is 1970-1978 for Columns (1) and (2) and 1971-1978 for Columns (3) and (4). The control variables are the same as those in Table 1 in the main text. Robust standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

References

Addoum, Jawad M., David Ng, and Ariel Ortiz-Bobea. 2020. "Temperature Shocks and Establishment Sales." Review of Financial Studies 33 (3): 1331–66.

Ashton, Basil, Kenneth Hill, Alan Piazza, and Robin Zeitz. 1984. "Famine in China, 1958–1961." Population and Development Review 10 (4): 613–45.

Babiarz, Kimberly S., Paul Ma, Grant Miller, and Shige Song. 2018. "The Limits (and Human Costs) of Population Policy: Fertility Decline and Sex Selection in China under Mao." NBER Working Paper 25130.

Brandt, Loren, and Carsten A. Holz. 2006. "Spatial Price Differences in China: Estimates and Implications." Economic Development and Cultural Change 55 (1): 43–86.

Chu, Cyrus. 1987. "The Dynamics of Population Growth, Differential Fertility, and Inequality: Note." American Economic Review 77 (5): 1054–56.

Chu, Cyrus, and Hui-wen Koo. 1990. "Intergenerational Income-Group Mobility and Differential Fertility." American Economic Review 80 (5): 1125–38.

Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena. 2020. "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments." Econometrica 88 (3): 879–915.

de la Croix, David, and Matthias Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." American Economic Review 93 (4): 1091–1113.

Dietzenbacher, Erik. 1989. "The Dynamics of Population Growth, Differential Fertility, and Inequality: Comment." American Economic Review 79 (3): 584–87.

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

Kremer, Michael, and Daniel L. Chen. 2002. "Income Distribution Dynamics with Endogenous Fertility." Journal of Economic Growth 7 (3): 227–58.

Kueh, Yak-yeow. 1995. Agricultural Instability in China, 1931–1990: Weather, Technology, and Institutions. Oxford: Oxford University Press.

Lam, David. 1986. "The Dynamics of Population Growth, Differential Fertility, and Inequality." American Economic Review 76 (5): 1103–16.

Li, Shi, Hiroshi Sato, and Terry Sicular, eds. 2013. Rising Inequality in China: Challenges to a Harmonious Society. Cambridge: Cambridge University Press.

Lin, Justin Yifu, and Dennis Tao Yang. 2000. "Food Availability, Entitlements and the Chinese Famine of 1959–61." Economic Journal 110 (460): 136–58.

Meng, Xin, Nancy Qian, and Pierre Yared. 2015. "The Institutional Causes of China's Great Famine, 1959–1961." Review of Economic Studies 82 (4): 1568–1611.

Peng, Xizhe. 1987. "Demographic Consequences of the Great Leap Forward in China's Provinces." Population and Development Review 13 (4): 639–70.

Shah, Manisha, and B. Kelsey Jack Steinberg. 2017. "Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital." Journal of Political Economy 125 (2): 527–61.

Song, Shige. 2013. "Identifying the Intergenerational Effects of the 1959–1961 Chinese Great Leap Forward Famine on Infant Mortality." Economics and Human Biology 11 (4): 474–87.

Spence, Jonathan D. 1991. The Search for Modern China. New York: W. W. Norton & Company.

Thaxton, Ralph A. 2008. Catastrophe and Contention in Rural China: Mao's Great Leap Forward Famine and the Origins of Righteous Resistance in Da Fo Village. New York: Cambridge University Press.

Wang, Xuebo, and Junsen Zhang. 2018. "Beyond the Quantity–Quality Tradeoff: Population Control Policy and Human Capital Investment." Journal of Development Economics 135: 222–34.

Xie, Yu, and Xiang Zhou. 2014. "Income Inequality in Today's China." Proceedings of the National Academy of Sciences 111 (19): 6928–33.

Yang, Jisheng. 2012. Tombstone: The Great Chinese Famine, 1958–1962. New York: Farrar, Straus and Giroux.

Yao, Shujie. 1999. "A Note on the Causal Factors of China's Famine in 1959–1961." Journal of Political Economy 107 (6): 1365–72.

Yao, Xinwu, and Hua Yin. 1994. Basic Data of China's Population. Beijing: China Population Publishing House.

Zhang, Junsen. 2017. "The Evolution of China's One-Child Policy and Its Effects on Family Outcomes." Journal of Economic Perspectives 31 (1): 141–60.