NBER WORKING PAPER SERIES

THE LONG TERM CONSEQUENCES OF FAMINE ON SURVIVORS: EVIDENCE FROM A UNIQUE NATURAL EXPERIMENT USING CHINA'S GREAT FAMINE

Xin Meng
Nancy Qian

Working Paper 14917
http://www.nber.org/papers/w14917

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2009

This paper is a companion to "The Institutional Causes of Famine in China, 1959-61" by Xin Meng, Nancy Qian and Pierre Yared. This paper supercedes the previous versions "The Long Run Health and Economic Consequences of Famine on Survivors: Evidence from China's Great Famine," (2005); and "The Causes of Cross-Sectional Inequity in Exposure to China's Great Famine and Its Long Run Health and Economic Consequences on Survivors" (2006). We thank Daron Acemoglu, Andrew Foster, Joshua Angrist, Abhijit Banerjee, Angus Deaton, Esther Duflo, Oded Galor, Bob Gregory, Ivan Fernandez-Vál, Vernon Henderson, Larry Katz, Ashley Lester, Ted Miguel, Gerard Padró-i-Miquel, Rohini Pande, Gerard Roland, Yona Rubinstein, David Weil and Noam Yuchtman for their insights; seminar participants at Brown , UC Berkeley, Princeton, Maryland, CEPR, LSE, Oxford, IIES, IGIER, the SF Federal Reserve, the World Bank and the CCER PKU for useful comments; Tal Levy, Sara Lowes, Katherine Wilson and Ang Sun for excellent research assistance; and the Michigan Data Center and the China Health and Nutrition Survey, funded by NIH (R01-HD30880, DK056350, and R01-HD38700), and the Carolina Population Center and the Chinese CDC for providing data. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

The Long Term Consequences of Famine on Survivors: Evidence from a Unique Natural Experiment using China's Great Famine
Xin Meng and Nancy Qian
NBER Working Paper No. 14917
April 2009, Revised March 2010
JEL No. I1,J01,J1,O1

ABSTRACT

This paper estimates the long run impact of famine on survivors in the context of China's Great Famine. To address problems of measurement error of famine exposure and potential endogeneity of famine intensity, we exploit a novel source of variation in regional intensity of famine derived from the unique institutional determinants of the Great Famine. To address attenuation bias caused by selection for survival, we estimate the impact on the upper quantiles of the distribution of outcomes. Our results indicate that in-utero and early childhood exposure to famine had large negative effects on adult height, weight, weight-for-height, educational attainment and labor supply.

Xin Meng
Australian National University
xin.meng@anu.edu.au

Nancy Qian
Department of Economics
Yale University
27 Hillhouse Avenue
New Haven, CT 06520-8269
and NBER
nancy.qian@yale.edu

The Long Term Consequences of Famine on Survivors: Evidence from China's Great Famine, 1959-61

Xin Meng Nancy Qian*

December 5, 2009

Abstract

This paper estimates the long run impact of famine on survivors in the context of China's Great Famine. To address problems of measurement error of famine exposure and potential endogeneity of famine intensity, we exploit a novel source of variation in regional intensity of famine derived from the unique institutional determinants of the Great Famine. To address attenuation bias caused by selection for survival, we estimate the impact on the upper quantiles of the distribution of outcomes. Our results indicate that in-utero and early childhood exposure to famine had large negative effects on adult height, weight, weight-for-height, educational attainment and labor supply. (O1 Development, I0 Health, J1 Demography)

1 Introduction

The impact of famine has been a question of long-standing interest amongst economists. In the twentieth century, over 100 million individuals have perished from exposure to famines.[1] In 1959-61, an estimated 16-30 million people died in China's Great Famine

*This paper is a companion to "The Institutional Causes of Famine in China, 1959-61" by Xin Meng, Nancy Qian and Pierre Yared. This paper supercedes the previous versions "The Long Run Health and Economic Consequences of Famine on Survivors: Evidence from China's Great Famine," (2005); and "The Causes of Cross-Sectional Inequity in Exposure to China's Great Famine and Its Long Run Health and Economic Consequences on Survivors" (2006). We thank Daron Acemoglu, Andrew Foster, Joshua Angrist, Abhijit Banerjee, Angus Deaton, Esther Duflo, Oded Galor, Bob Gregory, Ivan Fernandez-Vál, Vernon Henderson, Larry Katz, Ashley Lester, Ted Miguel, Gerard Padró-i-Miquel, Rohini Pande, Gerard Roland, Yona Rubinstein, David Weil and Noam Yuchtman for their insights; seminar participants at Brown , UC Berkeley, Princeton, Maryland, CEPR, LSE, Oxford, IIES, IGIER, the SF Federal Reserve, the World Bank and the CCER PKU for useful comments; Tal Levy, Sara Lowes, Katherine Wilson and Ang Sun for excellent research assistance; and the Michigan Data Center and the China Health and Nutrition Survey, funded by NIH (R01-HD30880, DK056350, and R01-HD38700), and the Carolina Population Center and the Chinese CDC for providing data. Comments and suggestions are welcomed at xin.meng@anu.edu.au and nancy.qian@yale.edu.

[1]See Sen (1981) and Ravallion (1997).

alone.[2] Both the fear of future famines and the lingering consequences of those that have already passed are ever relevant as we begin the new millennium. Hundreds of millions of people alive today in developing countries have been affected by a famine at some point during their lifetime. And as recently as 2002, the United Nation's *World Food Programme* calculated that up to 38 million Africans were living under the threat of famine. In China's case, survivors who were exposed to the Great Famine during early childhood or in-utero have only just passed the middle of their lives. While much attention has been paid to the effects of famine on mortality, surprisingly little has been paid to the effects on the millions of those who survived. The net effect of famine exposure on survivors is not *a priori* obvious for either health or labor market outcomes. On the one hand, severe food deprivation can only have adverse effects on the health of any given individual. Besides the direct mechanism of starvation, malnutrition also decreases the human body's ability to resist disease. On the other hand, it has been observed that conditional on survival, children who are malnourished exhibit rapid "catching-up" if proper care and nutrition are given.[3] Furthermore, studies have argued that when the famine is severe and mortality rates are high, survivors tend to be comprised of a small number of selected individuals who have naturally stronger constitutions and are better able to resist the negative effects of famine.[4] In the extreme scenario, these few survivors may be resilient to the extent that they do not exhibit symptoms of the adverse effects. For labor market outcomes, one must in addition consider the impact of famine in reducing labor supply, which could increase wages for survivors. This could potentially offset negative effects of the famine on survivors' health and work capacity.

This study addresses these questions by examining the net effect of childhood exposure to China's Great Famine on adult health and labor supply outcomes thirty years afterwards. We find that exposure to famine *in-utero* and during early childhood had significant negative effects on adult health, educational attainment and labor supply. The results also indicate that selection for survival and measurement error in the intensity of exposure cause standard OLS estimates of famine impacts to be severely attenuated.

Existing studies of the impact of famine on survivors provide somewhat ambiguous results. Epidemiological studies on the long run impact of the Dutch Famine (1944-

[2]See Coale (1981), Yao (1999), Peng (1987), Ashton et al. (1984) and Banister (1987).

[3]Outside of the famine context, there is evidence that conditional on survival, the negative impact of adverse childhood nutritional shocks is mitigated by rapid "catch-up" (Krueger, 1969; Hoddinott and Kinsey, 2001).

[4]For example, see Gorgens et al. (2002, revised 2007) and Deaton (2008).

1945) find that famine is positively correlated with psychological disorders in adulthood (Neugebauer et al., 1999; Brown et al., 2000; Hulshoff et al., 2000); obesity (Ravelli et al., 1999); and glucose intolerance (Ravelli et al., 1998). In contrast, Stanner et al. (1997) study a sample of approximately 600 survivors of the Leningrad siege (1941-1944) and find no effects of exposure. And Luo et al. (2006) find little or no difference in outcomes between survivors and a control group of individuals who were never exposed to China's Great Famine.

All studies of the long-run impact of famine face several empirical difficulties. First, it is difficult to measure the intensity of exposure. Researchers are generally not able to collect data at the time of the famine. The migration typically induced by famine makes it difficult to determine an individual's level of exposure after the fact. Even if researchers can learn of a survivor's location during the famine, the constructed measure of exposure intensity is likely to be measured with error. Second, famine typically occurs in concurrence with other events such as conflict, which can confound causal interpretation. Additionally, one may suspect that famine is more likely to strike in less agriculturally productive regions. In that case, the correlation between famine intensity and survivor outcomes may reflect the effect of being born in a particularly impoverished region rather than the causal impact of famine on survivors.

Finally, selection for survival can attenuate the estimated impact of famine. Survivors are typically from the top of the distribution of important characteristics, such as physical resilience, income, and access to nourishments (Gorgens et. al., 2002, revised 2007). Those from the lower parts of the distribution die. Since these characteristics are likely to be correlated with health and labor market outcomes later in life, comparing the mean of the distribution of survivors to the mean of the distribution of the control group could underestimate the true impact of famine. A recent study by Bozzoli, Deaton and Quintana-Domeque (2008) used cross-country data on infant mortality, which they interpreted as a measure of the disease and nutritional burden in childhood, and the mean adult height of surviving children, to develop a model of selection and stunting. They find that poor nutrition and disease during early childhood is not only responsible for mortality in childhood but also increases long-term health risks for survivors, risks that express themselves in adult height, as well as in late-life disease. Interestingly, their model predicts that, at sufficiently high mortality levels, selection can dominate scarring, leaving a taller population of survivors. This is consistent with the finding that survivors born close before and during the Chinese famine are not obviously shorter in stature relative to those never exposed.

Several recent studies have carefully constructed empirical strategies to address some of these issues for studying the long run effects of China's Great Famine. Gorgens et al. (2002, revised 2007) argue that the observed similarity between survivors and unexposed individuals is driven by attenuation bias from selection and that genetic determinants for survival are more likely to be transmitted to children than exposure to famine. Under this assumption, they show that selection bias can be mitigated by controlling for attributes of children. Their study does not address the problem that exposure to famine may be endogenous to other factors. To establish causal identification, Chen and Zhou (2002) use a difference-in-differences (DD) strategy. They compare exposed cohorts to unexposed cohorts between provinces where famine era mortality rates were high to those where they were low, finding that individuals who were one or two years of age during the famine were shorter in stature on average. Shi (2008) uses a similar strategy and finds that famine exposure as young children reduced education attainment and labor supply for women. Also using a DD strategy, Almond et al. (2006) focus on the effect of in-utero exposure. They compare children born to parents in Hong Kong who were never exposed to those born to parents who immigrated from mainland China and were exposed. They find that in-utero famine exposure has adverse effects on the adult labor supply and marriage outcomes for men. Their results are consistent with the findings of Brandt, Siow and Vogel (2008), who use a decomposition approach and find that the famine cohort is less attractive on the marriage market relative to the cohorts that immediately preceded and followed. These latter three studies do not attempt to address problems from selection. All of the studies mentioned could suffer from measurement error, which if random, would attenuate their estimates of the impact of famine exposure.

The principal contribution of this paper is to improve upon existing studies of the impact of famine for survivors by attempting to simultaneously address the problems of measurement error, endogeneity and selection. We use the average county level cohort size of survivors born during the famine to retrospectively measure exposure. This provides us with a continuous measure of famine intensity at a more disaggregated level than previous studies that have relied on province level variation. Next, to address the issue of positive selection for survival, we estimate the effect of exposure on the upper quantiles of the distribution of outcomes. Assuming that survivors are from the top of the distribution of outcomes, these estimates will more accurately reflect the true effect of exposure to famine. To the best of our knowledge, this is the first study to apply quantile analysis or to use survivor cohort size for studying the impact of famines. Finally, to address the problems of measurement error and omitted variable bias (OVB), we instrument for famine intensity with a unique source of variation derived from the institutional causes of what's often called the "Great Leap Forward Famine".

The Great Famine is widely believed to have been caused by over-procurement. In a companion paper, Meng, Qian and Yared (2009) show that the centrally-planned procurement institution transformed the small drop in production in 1959 into the largest famine in history. More importantly for this study, it produced a striking geographic pattern: regions that produce more grain suffered more during the famine. This pattern was a reversal from normal years, when these productive regions had lower mortality rates. In this paper, we use this institution driven cross-sectional variation in famine intensity, measured as two predictors of famine intensity, grain productivity measured as the suitability for cultivation predicted by natural conditions and area sown for grain in 1997, in combination with birth cohort variation to instrument for exposure. The more grain a region normally produced, the more intense the famine in that region; and the younger an individual was at the time of the famine, the more vulnerable she was to exposure. Only the combination of the two sources of variation can be interpreted as plausibly exogenous.

Note that county level data on historic production is not available. At that disaggregated level, production data is only available for 1997. Using geographic suitability and the area sown for grain in 1997 together to instrument for famine intensity is similar in spirit to proxying for historic production with area sown 1997, which suffers the caveat that production patterns may have changed during the fourty interim years, and then instrumenting for area sown with suitability for cultivation based on natural conditions to address this problem. However, this more structured approach would require specific assumptions about the production function of grain based on natural conditions and the changes over time. Lacking a reliable model for making such predictions, we use a more reduced form approach. This has the advantage of transparency while still addressing key endogeneity and measurement error issues. The main challenge in using our strategy is that regions that produced more grain and suffered more during the famine are also regions that typically were better off (e.g. lower mortality rates). Since survivors presumably recover more fully and quickly under better conditions, our 2SLS estimates should be interpreted as the lower bound estimates of the true impact of famine exposure.

The analysis uses data from several existing sources: the *1990 Population Census*, the *1989 China Health and Nutritional Survey*, the *1997 China Agricultural Census*, and the

FAO GAEZ data set on crop suitability. In addition, we supplement the main county-level analysis with data on historical grain production and death rates at the province level. Our study focuses on agricultural households. They suffered the most from the famine and were relatively undisturbed by the Cultural Revolution (1966-1976) which closed secondary and tertiary education institution during the late 1960s and early 1970s. Our results for individuals in the 90th percentile of the distribution of outcomes indicate that in-utero exposure to famine on average caused a 1.7% (2.8 cm) reduction in height, a 2.3% (1.4 kg) reduction in weight and an 8.6% (0.6 years) reduction in educational attainment. For those exposed during early childhood, the famine on average reduced height by approximately 1.6% (2.7 cm), weight by 5% (3 kg), WFH by 1.2% (0.004 kg/cm) and labor supply by 13.9% (12.6 hours per week).[5]

The adverse effects are consistent with results from other studies of the famine. However, by mitigating the attenuation bias caused by measurement error and selection, our estimated impact of early childhood exposure is larger in magnitude than most previous studies. The OLS estimates show that our crude correction for selection bias – estimating the effect of exposure to famine for survivors in the 90th percentile – doubles the magnitude of the estimated effect that one would obtain from estimating for those in the mean of the distribution. These results are important for both interpreting previous and conducting future studies on the long run impact of famine on survivors. They demonstrate that important long-term effects may be missed or underestimated when improper control groups are used. The 2SLS estimates show that correcting for measurement error further doubles the magnitude of the OLS estimates (for both individuals in the 90th percentile and the mean of the distribution).

The main caveat for interpreting the estimated effects on the 90th percentile as the mean effect is that one needs to assume that exposure to famine does not have heterogenous effects for individuals on the different quantiles of the distribution of outcomes that would produce the same patterns as selection. One could argue that this is unlikely because it would mean that the healthiest individuals were most vulnerable to exposure to famine. Casual observation would suggest that the opposite should be true: those who are naturally weak should be more vulnerable to shocks. However, there is little we can do to test this directly and the medical literature gives us little insight on the presence or patterns of heterogenous effects. Hence, one should extrapolate the results on the 90th

[5]For average effects, we report only the statistically significant outcomes for each group. The estimated effects for each group are shown for all outcomes in the section on results.

percentile with caution. In any case, the insight that the effects on the 90th percentile are much larger than the effects on the mean is important for any study of impacts of large shocks (e.g. war) on survivor's outcomes. Subject to the caveats we describe above, our method of addressing selection has the benefit that it is easy to apply to other contexts where researchers have similar priors about the pattern of selection, and the assumptions necessary for interpretation are transparent.

This study makes several contributions in addition to providing an improved estimate of the long run impact of exposure to famine. The results can shed some light on the impact of severe childhood malnutrition.[6] As recently as 2004, World Development Indicators reported that, worldwide, 30% of children under the age of five are estimated to be severely malnourished.[7] Our results show that these deprivations can have large, adverse long run effects even amongst a sample of "strong" individuals selected to survive severe famine. And this could in turn affect long run economic development.[8] This study also makes a methodological contribution. It shows that under certain conditions, estimation of the impact on the upper quantiles of the distribution of outcomes can potentially be one way for addressing attenuation bias caused by selection for survival. This method is straightforward and can be easily employed for impact evaluations of other events where similar patterns of selection for survival is present.

There are several advantages to studying the long run impact of famine in the context of China's Great Famine. The institutional causes of the Great Famine provide an unusual natural experiment. It is unlike any other famine in history with the exception of the Ukrainian famine (1932-33) in that migration was strictly restricted. This allows us to measure regional famine intensity with the cohort size of survivors with reasonable accuracy. Second, it is a relatively good setting for isolating the effect of famine. Compared to other famines, it occured in a time of relative prosperity and political stability. Unlike notable famines such as the Ethiopian famine, the Chinese famine was not a direct

[6] Recent studies on long run effects of health shocks during childhood include studies by Almond and Mazumder (2005), Almond et al. (2005), Berhman and Rosenzweig (2005), Black et al. (2005), Bleakley (2002), Case et al. (2004), Glewwe et al. (2001), and Maccini and Yang (2008).

[7] Prevalence of child malnutrition is the percentage of children under five years of age whose height-for-age is more than two standard deviations below the median for the international reference population for ages 0 to 59 months. The reference population adopted by the WHO in 1983 is based on children from the United States, who are assumed to be well-nourished.

[8] The correlation between improved health status and economic factors has been found in studies by Fogel (1994), Fogel and Costa (1997), and Smith (1999). Bloom et al. (2001) find a correlation between longer life expectancy and higher economic growth rates. Weil (2005) finds that 26% of the cross-country variation in income can be explained by differences in health.

outcome of civil war or other types of conflict. Famine stricken regions are not known to have received special treatment from the government which could confound our estimates. For political reasons, the Chinese government actually denied the extent of the famine domestically and internationally for three decades. Finally, as a study of the long run effects of childhood malnutrition, the interpretation of the effect of the region-cohort level shocks from the Great Famine are not confounded by omitted variables related to parental heterogeneity.[9]

The paper is organized as follows. Section 2 discusses the background, the empirical strategy and the conceptual framework. Section 3 describes the data. Section 4 presents the results. Section 5 interprets the results. Section 6 offers conclusions.

2 Background

Officially, the cause of the famine, called the "three years of natural disasters" (*san nian zi ran zai hai*), was a fall in grain output due to bad weather. However, recent studies have argued that very little can be explained by weather.[10] Today, it is widely accepted that although there was a fall in agricultural output, the extent of the famine was largely driven by a set of misguided policies.[11] Specifically, we argue that over-procurement of grain from rural areas was a major contributor to the famine. Over-procurement in 1959 led to a decrease in nutrition in rural areas, which in turn led to a decrease in rural workers' physical capacity to produce grain. The reduction in work capacity along with the consumption of inputs such as seeds in the winter of 1959 prolonged the famine. In 1960, the central government had decreased procurement. The famine ended in 1961, when grain reserves were distributed to aid stricken regions. Production recovered to pre-famine levels in few subsequent years. The government is not known to have targeted any subsidies or compensation programs toward famine stricken regions after the event. This is not wholly surprising given that the government has typically tried to minimize

[9] If "bad" parents invest less in children's nutrition and education, then the correlation between the outcomes will reflect parental preferences rather than the causal impact of malnutrition on educational attainment.

[10] For example, see Li and Yang (2005), and Meng and Qian (2009).

[11] The suspected causes of the famine include labor and acreage reductions in grain production (e.g., Peng, 1987; Yao, 1999), implementation of radical programs such as communal dining (e.g., Yang, 1996; Chang and Wen, 1997), reduced work incentives due to the formation of the people's communes (Perkins and Yusuf, 1984), and the denial of peasants' rights to exit from the commune (Lin, 1990). Lin (1990) argues that the removal of exit rights destroyed reduced work incentives for shirkers, and hence decreased overall grain production. See Li and Yang (2005) and Meng and Qian (2008) for a detailed discussion.

people's perceptions of the famine's severity (Li and Yang, 2005). Figure 1 plots average county level cohort sizes by birth year from the 1990 Population Census. The most severe reductions in cohort size are for those born close before and during the famine. The reduction in cohort size is several orders of magnitude larger for agricultural populations relative to non-agricultural populations.

Meng, Qian and Yared (2009), henceforth MQY, expanded on previous studies of the causes of the famine. In this companion paper, we show that despite the small fall in production in 1959, had distribution of food been equal across regions, more than enough food was produced in 1959 to prevent starvation or famine. (The most conservative estimates of production in 1959 show that it was 50% more than what was needed to prevent mortality). More importantly, for this paper, we show a striking geographic pattern in famine intensity. The regions which produced more grain in 1959, the same regions which were always more productive, suffered famine more intensely. This is a reversal from the pattern during normal years, where we find that these same productive regions have lower mortality rates. See Columns 3 and 4 in Table 1 which is taken from that paper.[12] This rather surprising pattern of the distribution of famine is consistent with the observation of journalist, Jasper Becker (1996), who noted that unlike previous famines in China, the Great Famine was most severe in grain rich regions rather than the traditional famine belt. In the companion paper, we provide a theory for how the famine and the geographic patterns of famine were an outcome of the centrally-planned grain procurement system, which was constrained by imperfect information and bureaucratic incapacity.[13]

[12] The estimate for famine years in Column (4) of Table 1 is insignificant, most likely to due insufficient sample size and measurement error in mortality rate data. MQY (2009) addresses this by supplementing the provincial analysis with a county level analysis. The county level analysis shows the same patterns. For brevity, we do not discuss it further in this paper.

[13] We consider a stochastic endowment economy in which a utilitarian government is interested in equalizing food consumption across regions. The government is subject to two constraints. First, though the government knows the expected food production in a given region, it cannot verify its realization since farmers can under-report or hide their production in order to avoid procurement (taxation). Second, the report of production and procurement (collection of taxes) in a given region occurs independently of other regions. Thus, the government is constrained in its bureaucratic capacity since it cannot condition procurement for a particular region on production reported by other regions. (This refers to a context where the central government sends out numerous officials to procure grain in far outlying areas. These officials procure grain and return to cities to report information. It is only then, after procurement has taken place, that the administration can cumulate the information from across regions, and then send the information to provincial capitals and then finally to Beijing. Given the low level of development in communication infrastructure in China in the 1950s, this tedious process could take weeks if not months.) We consider an economy in which all regions can be affected by a binary aggregate proportional shock to food production. In this setting, optimal constrained policy assigns a constant procurement tax or

For the purposes of this paper, the most important finding of MQY (2009) is the strong positive correlation between grain production in 1959 and famine intensity, which was a reversal from normal years when production was negatively correlated with mortality. In that paper, production is measured as both historical production data at the province level, and by using a proxy – county level data on suitability as predicted by natural conditions. The latter is a robustness check in case the historical data on production suffers from reporting errors. In this paper, we will similarly use both a measure of real production and a measure of suitability for production which does not suffer from reporting or other endogeneity concerns to instrument for famine intensity.

Figure 2, which we also take from MQY (2009), plots the coefficients of the interaction terms of grain suitability and birth cohort dummies from a regression of county level cohort size in 1990 on these interaction terms, birth county and birth year fixed effects.[14] Using cohort size in 1990 as a measure of survival, this shows that for agricultural populations who were affected by the grain procurement system, grain suitability in the county of birth as predicted by natural conditions is negatively correlated with survival for those born closely before or during the famine. For non-agricultural populations living in the same counties but who were not affected by the grain procurement system, grain suitability is uncorrelated with survival for non-agricultural populations.[15] We will use this peculiar institution-driven pattern in regional famine intensity to create our instrumental variables.

subsidy for each region. The model predicts that under the negative shock, mortality risk is sharpest in regions which produce more food since they experience a heavier tax burden. The model also predicts no rise in mortality risk for non-agricultural populations during a drop in food production since these receive constant subsidies. Moreover, consistent with the evidence, the model predicts a positive relationship between food production and food consumption in the absence of a negative shock (MQY, 2009).

[14]MQY (2009) uses data from the 1990 Population Census to estimate the following equation for the population of agricultural households and non-agricultural households separately. Figure 2 plots the vectors of coefficients, $\hat{\beta}_t$.

$$\ln pop_{ct} = \sum_{t=1943}^{1966} \beta_t (\ln grain\_suitability_c \times biryr_t) + \alpha + \gamma_c + \delta_t + \varepsilon_{ct} \quad (1)$$

$\ln pop_{ct}$, the natural logarithm of the cohort size of individuals born in county c and birth year t is a function of: the interaction terms between the per capita area suitable for grain cultivation in birth county c, $grain\_suitability_c$, and dummy variables for being born in birth year t, $biryr_t$; and birth cohort dummy variables; county fixed effects, γ_c; and birth year fixed effects, δ_t. The reference group is comprised of individuals born in the period 1930-1942. This group and all of its interactions are dropped. β_t is the correlation between grain suitability and cohort size for those born in year t. The coefficients and standard errors are reported in Appendix Table A1.

[15]This finding is robust to controlling for $province \times year$ fixed effects. We do not show them in this paper

2.1 Empirical Strategy

Our strategy exploits cross sectional and cohort variation in famine exposure. Individuals living in regions that had intense famines will suffer more deprivation. Exposure to famine is likely to have different effects depending on the age of the individual. The elderly and the very young are most likely to suffer. This could be because they are biologically more vulnerable or because households decide to allocate more food to individuals who can work and bring in food or income. Because we can only observe survivors thirty years after the fact, our study focuses on those who were in-utero or very young children at the time of the famine.

We treat exposure during early childhood separately from exposure in-utero. Malnutrition is likely to have different effects at different stages of development. For very young children, exposure to famine will cause severe malnutrition and decrease the quality of care received from adults. For those who were in-utero, exposure to famine is transmitted through malnutrition of the mother and during infancy.[16] Moreover, the attributes that determine survival is different for the two groups. Both groups will be selected based on characteristics that determine one's own survival. For the in-utero group, parents' decisions to have children when there is a famine also play a role. Overall fertility rates were extremely low during the famine. Hence children born during the famine are likely to be very different on average than children born prior to the famine.

Our second stage equation is the following:

$$\begin{aligned} Y_{ct} &= \beta_{54-58}(\ln famp op_c \times born54\_58_t) \\ &\quad + \beta_{59-61}(\ln fampop_c \times born59\_61_t) + \alpha + \gamma_c + \rho_t + \varepsilon_{ct} \end{aligned} \tag{2}$$

The outcome for individuals born in county c during year t is a function of: the interaction between the intensity of the famine in that county measured as the natural logarithm of the average cohort size of individuals born during 1959-61, $\ln fampop_c$, and a dummy variable for whether that individual was born during 1954-58, $born54\_58_t$; the interaction between the intensity of famine and a dummy variable for whether she was born during the famine, $born59\_61_t$; county fixed effects, γ_c, and birth year fixed effects, ρ_t. Individuals born before 1954 and after 1961 form the reference groups. They and

[16] Aizer, Stroud and Bubka (2009) provide evidence that stress-induced elevatation of cortisol levels in pregnant mothers have long term adverse effects on children. Almond et al. (2006) provide a detailed discussion of the differences between in-utero exposure to health shocks.

their interaction terms are dropped. Standard errors are clustered at the province×year level. We do this because for some of the outcomes, we only have data from the seventeen CHNS counties. Clustering on such a small number of counties introduces small sample bias.[17] Each observation is a county-birth year cell. The outcomes are the means of each cell. The cell size is used to weight the regressions by population. The coefficients and standard errors from these population weighted cell level regressions are numerically identical to regressions with individual level data.

Like difference-in-differences, changes across cohorts that affect different regions similarly are controlled for by the comparison across regions. Cohort invariant differences between regions are controlled for by the comparison across cohorts. For example, if regions with bad institutions are more prone to famines and institutions do not change over short periods of time, then differences in institutions will be controlled for by region fixed effects.

There are several reasons why a straight-forward OLS estimate as in equation (2) would under-state the true impact of exposure. First, cohort size may measure famine intensity with error. Since measurement error is most likely random, this will attenuate the OLS estimate. Second, selection bias suggests that the mean of the outcomes in the control group are below the mean of the group of individuals who are "naturally equivalent" to the treated group. Alternatively, OLS may also over-state the true effect. If famine was more severe where local politicians, seeking promotion, allowed more grain to be delivered to the upper levels of government, then these same politicians may also have implemented other policies that had adverse effects on the population.

To address the problems of measurement error and endogeneity, we instrument for famine intensity with the interaction between non-famine grain production and birth year. Two facts are exploited: 1) the grain procurement system caused the famine to be more severe in regions that typically produced more grain; and 2) children who were younger at the onset of the famine were more vulnerable to disease and malnutrition,and children born after the famine were not exposed. We proxy for non-famine grain productivity with the per capita area suitable for grain production as predicted by geographic and climatic suitability, and observed per capita area of grain sown in 1997. Historical production data is not available at the county level. In any case, our proxies have the advantage that they

[17]For the first stage estimate of the effect of grain suitability on survival and the second stage estimate of the effect of famine on educational attainment, we can expand our data to contain all counties in the 1990 Population Census. When we do so, we can cluster at the county level or at the province×year level. The standard errors are similar in the two cases. We do not report the estimates for brevity.

do not suffer from systematic over-reporting by the government, which wanted to deny the extent of the famine and show the outcomes of GLF policies in as favorable of a light as possible.

The instruments are the interaction terms between the grain suitability of the county of birth and birth cohort dummy variables; and the interaction terms of the 1997 per capita grain sown area in the county of birth and birth cohort dummy variables. Only the interaction terms can be interpreted as exogenous. Note that using both sets of instruments is similar to proxying for historical county level production with the 1997 data on production, which by itself may suffer from the problem that it measures true historical production with error, and then instrumenting for this proxy with and geographic suitability which addresses the measurement error (and potential endogeneity) issues. However, this method would require us to make many assumptions about the structure of the relationship between suitability and production, and about how this relationship may have altered from the famine era to 1997. After a thorough review of the agricultural science literature, we could not find a model that would allow us to reliably make the necessary predictions. Hence, rather than doing this, we choose to impose the minimal amount of structure and use suitability and area sown in 1997 as two separate instruments. This method has the same advantages of addressing endogeneity and measurement issues as the more structured method. Namely, because we are using suitability as one of the instruments, our estimates are not vulnerable to bias if the 1997 production data measures true historical production with error or if the famine influenced factors which partly determine production levels in 1997. In addition, this method has the advantage of transparency.[18]

We estimate two first stage equations for the two endogenous independent variables: $\ln fampop_c \times born_t$, where $t = \{54 - 58\ ,59 - 61\}$.

[18]An alternative is to only use interaction terms between birth cohort dummies and natural conditions as instruments. Doing this predicts a similar pattern as when we use both sets of instruments but is much less precise. This is most likely because there are factors besides natural conditions which predict grain production. For example, many cities are situated in regions that are very suitable for agricultural production (e.g. Shanghai). But in practice, very little grain is produced in the regions near Shanghai due to the proximity to that large urban metropolis. Adding interaction terms of birth cohorts with 1997 production levels as an additional set of instruments addresses this issues. For brevity, we only report results using both sets of instruments in this paper.

$$\ln fampop_c \times born_t = \sum_{t=54-58}^{59-61} \beta_t \ln(grain\_pcsuit_c) \times born_t + \sum_{t=54-58}^{59-61} \beta_t \ln(grain\_pcprod97_c) \times born_t + \alpha + \gamma_c + \delta_t + \varepsilon_{ct} \quad (3)$$

The key identification assumption is that suitability for grain production and the adult outcomes of famine survivors in 1990 are not jointly determined by some omitted variable. The most likely problem is that areas that produce more grain normally would recover from the famine sooner. This may be because in non-famine years, there is more food. Individuals born in those regions are better nourished prior to the famine and may be better able to recover. And after the famine, when production recovered, the local population would once again receive better nutrition. The first possibility is not very likely since for many years before and after the famine, agricultural households were left with little surplus (Perkins, 1966). And if it were true, it would bias our estimates towards zero.

Similarly, because these regions are typically richer, they may have better institutions or public goods (e.g. provision of health and education infrastructure) that mitigate the negative effects of famine and aids in the subsequent recovery. If this is true, then the most likely direction of the bias is also towards zero. Survivors in places with better schools and health care are likely to recover better and faster.

We are able to partially check whether this assumption is true to the extent that the institutions that can mitigate the effects of famine are provided for at the county level and provided similarly for agricultural and non-agricultural households. We can compare the effect of grain suitability for agricultural households who were under the procurement system to non-agricultural households living in the same county who were not under the grain procurement system. Figure 2 shows that grain suitability had no effect on cohort sizes for the latter. These results should be interpreted loosely since much public goods provision in China are specific to rural or urban households. Nevertheless, the clear distinction between the effect of grain suitability on agricultural and non-agricultural households lends credibility to our strategy.[19]

[19]Note that despite being isolated from the full extent of the famine, non-agricultural populations cannot be used as a comparison group for rural areas in the second stage because they are subject to different policies regarding the access to labor market and schooling opportunities. These differences also differed by cohort and hence can produce confounding results.

Our instrumental variables strategy aims to correct for biases due to measurement error and omitted variables. But it cannot address attenuation bias caused by selection. "Survival of the fittest" is a general concern in estimating the impact of severe deprivations on survivors. For China's Great Famine, findings from Gorgens et al. (2002; revised 2007) indicate that individuals with higher stature were more likely to survive. This means that the estimated impacts on average height or other outcomes that are correlated with the determinants for survival may be attenuated by selection bias. For example, if the underlying determinants of other outcomes such as labor supply are correlated with the latent indicators of health measured by these outcomes, then the estimated impact on those other outcomes will also be attenuated. Assuming that it is the strongest (or tallest) who survive, the attenuation bias will be smaller in magnitude for individuals on the higher quantiles of the distribution of outcomes because those individuals have comparable control groups. For individuals in the lower quantiles, the strategy compares individuals in the control group with those in the treatment group who would be higher on the distribution absent the famine-induced selection.

Therefore, we can address the problem of selection by estimating the impact of famine on the upper quantiles of the distribution of outcomes. We calculate the sample means of the different quantiles for every birth year-county cell and repeat the same estimation. If height, weight and weight-for-height are correlated with the determinants for survival, then estimates on the upper quantiles (e.g. 90th percentile) will be less attenuated by selection relative to estimates for the lower quantiles. If those determinants are also correlated with education and labor supply, a similar logic implies that estimates at higher quantiles will also be more accurate for those outcomes. Since famine intensity, the treatment variable, is only measured at the birth county and birth year level, this aggregation does not change the independent variables. The only difference between the analysis on the quantiles and the analysis on the mean is in the outcome variables. This strategy is similar to Quantile Regressions and Quantile Instrumental Variables in spirit. It has the advantage that we are able to control for fixed effects.

Another difference is the Cultural Revolution (1966-1976) which targeted non-agricultural populations (Unger, 1982; Meng and Gregory, 2005; Giles et al., 2006). The Cultural Revolution caused widespread closings of schools for approximately three years (1966-1969). Children who survived the famine would have been in school during the Cultural Revolution. Hence, comparing the famine cohort between urban and rural areas would compare outcomes for two different treatments rather than a treatment and a control. We therefore restrict the sample to individuals living in rural areas. Our empirical strategy will be robust to the occurrence of school disruptions in rural areas as long as school closings were not correlated with famine intensity.

Using this method to address selection assumes that differences in the estimated effects of famine exposure across quantiles is due to selection bias rather than to heterogenous treatment effects. We have no way of testing this assumption directly. Hence, this is by no means a fool-proof method of addressing selection bias. We choose this method because the assumptions are transparent and not obviously unreasonable, not many are needed, and the method is simple to replicate. A more detailed discussion is provided later in the paper.

2.2 Conceptual Framework

Exposure to famine at young ages affects adult health and labor market outcomes through two main channels. First, it adversely affects childhood health, which is a product of genetic endowment, fetal health (in-utero nutrition), nutrition and other forms of investment (e.g., health care). The famine potentially also reduced the quality and/or quantity of other forms of investment into children by reducing the health status of parents. Childhood health can in turn affect adult outcomes directly and indirectly (Kuh and Wadsworth, 1993), as poor childhood health can affect adult health directly, which consequently can affect work capacity and labor supply. Barker (1995) and Ravelli et al. (1998) have found that nutrition in-utero can affect health status in middle age, through its impact on chronic conditions such as coronary heart disease and diabetes, in a phenomenon widely known as the "Barker Hypothesis".[20] Poor childhood health could also decrease educational attainment by decreasing returns to education or by increasing the costs of school attendance (Curie and Madrian, 1999; Miguel and Kremer, 2004). This may in turn affect labor supply and/or wages later in life.[21] Second, exposure to famine could potentially have a positive effect by reducing the cohort size of exposed individuals,

[20]Experimental work by Ozanne and Hales (2004) using laboratory mice find that lab mice that are underfed in-utero but who are well-fed after birth catch up rapidly. However, they die earlier than mice that are also well-fed in-utero.

[21]Poor health in children has been associated with lower education and/or labor market outcomes in the U.S. (Case et al., 2004; Doblhammer, 2002), Canada (Currie and Stabile, 2004), Great Britain (Case et al., 2002; Kuh and Wadsworth, 1993; Marmot et al., 2001) and many developing countries (Behrman, 1996; Bleakley, 2002; Brinkley, 1994; Glewwe and Jacoby, 1995; Glewwe et al., 2001; Miguel and Kremer, 2004; and Strauss and Thomas, 1998). See Curie and Madrian (1999) and Curie and Hyson (1998) for a review of studies linking health to educational attainment and labor market outcomes. The latter focuses on the effects of low birth weight. Smith (1999) shows a strong correlation between reported health and income of adults in the U.S. Reduced height has been associated with lower education and labor market outcomes in many countries (Maccini and Yang, 2005; Perisco et al., 2002; Strauss and Thomas, 1998; Schultz, 2001; Schultz, 2002; Strauss and Thomas, 1998).

hence reducing labor market competition and competition for family resources.[22] This paper will estimate the net effect of exposure to famine: the sum of the adverse effect of malnutrition and the potentially positive effects from smaller cohort sizes.

In addition to the health channels that we describe above, famine also affects those exposed in utero by affecting parents' decisions to bear children. The extremely low fertility rates during the famine suggests that very few parents were able to or chose to have children. Because we cannot observe parents of survivors, we cannot investigate the attributes driving this selection. Nor will we be able to reliably predict how these attributes would affect the impact of famine exposure. Therefore, without wild speculation, we cannot say more about the differential impacts between exposure in-utero and exposure during early childhood. The most we can do is to allow the effects for in-utero exposure to differ from the effects of early childhood exposure.

3 Data

This paper constructs a panel of birth cohorts by matching several existing data sets: the 1% sample of the 1990 Population Census, the 1989 China Health and Nutritional Survey (CHNS), the 1% sample of the 1997 Agricultural Census, and GIS data on suitability for grain cultivation which we construct using data from the Food and Agricultural Organization's GAEZ database.

The 1990 Population Census contains 32 variables including birth year, region of residence, and how long an individual has lived in a region of residence. We use these data to count the number of "missing people" from the 1959-61 cohort. The average cohort size across these three birth years for each county is our measure of famine intensity. The large sample size of the Census gives us more precise measures of famine intensity than smaller samples such as the CHNS. We restrict the sample to individuals who report as living in the same county for five or more years and interpret their county of residence as their birth county. For this, we assume that there is little migration between when the famine occured and 1990. Our assumption is supported by studies on migration in China, which find that strict migration controls were well enforced until the early to mid 1990s.[23]

[22]Easterlin (1980) discusses how the size of a generation affects the personal welfare of its members through family and market mechanisms. See Becker and Lewis (1973), Becker and Tomes (1976), Galor and Weil (2000), Hazan and Berdugo (2002) and Moav (2005) for theoretical discussions of the quantity-quality tradeoff; and see Angrist et al. (2006), Black et al. (2004), Qian (2006), Rosenzweig and Zhang (2006), and Schultz (2005) for recent empirical evidence on the quantity-quality tradeoff.

[23]Migrants had no access to government-controlled food rations, housing, schools and medical care

Our outcome data come from the 1989 CHNS. It uses a random cluster process to draw a sample of approximately 2,520 rural households with a total of 10,534 individuals across seven provinces that vary substantially in geography, economic development, public resources, and health indicators. The survey includes a physical examination of all individuals as well as information on labor supply, work intensity and wages. For this study, we use height, weight and weight-for-height (WFH). The CHNS also report diastolic and systolic blood pressure. We use these to create a dummy variable for whether an individual has hypertension (>140/90mmHg). These data have the advantage that they are measured by the surveyor and hence avoid measurement problems from self-reported data. The CHNS also reports education level, which we translate to a continuous variable for the years of educational attainment. The wage data are not used in this study because the majority of rural workers are involved in non-wage earning production and because wages in China during this period do not reflect marginal productivity. Instead, we use labor supply to measure the physical capacity to work. It is calculated as the sum of the number of hours per week spent in wage labor, agricultural labor (e.g. farming, gardening, tending livestock, fishing), and home production.

We create two proxies for regional grain productivity. The first measure is the per capita area suitable for cultivating rice and wheat, the two main crops subject to central procurement targets. This is calculated using the model and data provided by the United Nations Food and Agriculture Organization's (FAO) Global Agro-Ecological Zones (GAEZ), 2002 database, which provides data for 50km by 50km grids for the world. The data are the result of over twenty years of research and are the product of a joint collaboration between the FAO and the International Institute for Applied Systems Analysis (IIASA). Their measure is based purely on the biophysical environment of a region and it is not influenced by which crops were actually adopted in an area. Factors that are easily affected by human actions, such as soil pH, are not parameters in their model. Nunn and Qian (2009) provide a detailed description of the construction of this data and how to calculate suitability measures at the regional level. We use the same method.

We are able to choose the level of inputs that the calculation is based on. Great Leap

once they left their registered homes. The first wave of rural migration did not occur until the early to mid 1990s, during the urban construction boom, and most of those migrants were young adult men. Consequently, it is highly unlikely that the results of this paper are confounded by migration. Using data from China's Ministry of Agriculture RCRE's National Fixed Point Survey for 1986-90, we find that the probability of having a household member work away from the home village is very low and similar for regions that suffered very different levels of famine intensity. They are not reported in the paper. Also see West and Zhao (2000) and De Brauw and Giles (2006) for detailed discussions on migration.

Forward policies required collectives to not use chemical fertilizers. To the best of our knowledge, the use of heavy machinery such as tractors was extremely rare. Hence, our chosen level of inputs allows for rain-fed irrigation but no heavy machinery or chemical fertilizers. We assign a cell with a value of one indicating that it is "suitable" if it can produce 40% or more of the maximum output. The measure of suitability is the fraction of cells in a county that is suitable. We multiply this by the total area of the county and then divide by county population in 1990 to obtain per capita area suitable for grain production. This threshold was arbitrarily chosen, but moderately changing it will not affect the 2SLS estimates. A higher threshold (e.g. suitable defined as cells that can produce 60% or more of the maximum output) will mean that the suitable land can produce more grain per grid. However, this will be offset by the fact that in each region, there will now be fewer grids that are classified as suitable. Our second proxy for non-famine production is the per capita area of rice and wheat sown as reported in the 1997 Agricultural Census. This is the only source of county level data that uses a consistent measure across regions. For each county, we divide the area reported as sown for rice or wheat by county population in 1997.

The data are collapsed and matched by county of residence and birth year. All regressions will be weighted by the county-birth year cell size. To mitigate potentially confounding effects from the Cultural Revolution, which was primarily an urban disturbance, and rural-urban migration, we exclude cities. The CHNS data and the 1990 Population Census matched for seventeen counties within seven provinces. The number of individuals in each county-birth year cell is retained so that we can weight our regressions by population. Figure 1 plots the cohort sizes of agricultural and non-agricultural populations by birth year. The vertical line indicates the beginning of the famine. It shows a significant decrease in cohort size for those born closely before and during the famine in both urban and rural areas. For those born before the famine, the decrease most likely reflects increased mortality due to the famine. For those born during the famine, the decrease is likely to reflect a combination of increased mortality and reduced fertility. This only reflects a part of the estimated 16.5-30 million deaths during the famine because we cannot observe the mortality of the elderly in the 1990 data. The elderly, like the very young, are more vulnerable to health shocks and experienced higher mortality rates relative to other age groups. But because this cohort would be approximately 100 years old in 1990, they do not appear in the 1990 Census.

For ease of interpretation, we measure famine intensity as the natural logarithm of the average cohort size of those born during 1959-61.[24] Figure 3A is a histogram of this measure. It shows that there was much variation in famine intensity. For illustrative purposes, we can also measure 1959 cohort sizes as a fraction of the average cohort size of those born in 1952-54. Figure 3B is a histogram of this normalized measure of famine intensity. In some counties, there were no births during 1959, resulting in this ratio being zero. In other counties, there was little difference in the 1959 cohort size relative to previous cohort sizes. The geographic dispersion of famine can also be observed in Figure 4, where we map the main famine measure (broken up into five equal-frequency categories) by county. Lighter shades reflect more intense famines. The counties on this map are all the counties of the CHNS provinces for which we were able to match the Population Census and the GAEZ data. The counties for which we have survey data from the CHNS are a randomly sampled subset of this map. However, we show all the matched counties in these provinces map to better highlight two facts. First, neighboring counties can have very different famine intensities. Second, the famine was more severe in the south, which is better suited for grain production.

Table 2 shows the summary statistics of our outcome variables by birth year. All of these variables are from the 1989 CHNS data, except for years of schooling, which is obtained from the 1990 population census data. Height is a commonly used measure of the stock of nutritional investments during the fetal and childhood stages of life (Fogel et al., 1982; Fogel, 1994; Steckel, 1986; Micklewright and Ismail, 2001). Average height is approximately 160 cm, four centimeters less than the average height of the same cohort in Japan. Weight, weight-for-height and BMI are crude measures of the body's inability to retain body mass after recovering from a severe nutritional shock. This could be due to the inability to absorb nutrition after suffering severe shocks to the gastrointestinal system such as severe repeated diarrhea during early childhood (Cutler, Deaton and Lleras-Muney, 2006). The sample means for both of these outcomes are similar to comparable cohorts from Japan. Individuals have on average 5.3 years of education. The main economic outcome we examine is total hours worked per week. Adults in the sample work over 73 hours per week, on average. Comparing the outcomes of individuals born prior to or during the famine in columns (2), (3) and (4) to those born after shown in column (5) do not show cohorts affected by the famine as systematically worse off than those not affected. This most likely reflects the strong secular trends in improvement in nutrition,

[24] County-birth year cells for which there are zero births are retained by adding calculating cohort size as $\ln fampop = \ln(1959-61\_cohort\_size + 0.1)$.

sanitation and public health during this period.

4 The Long Run Impact of Exposure to Famine

4.1 OLS

The estimates from equation (2) are shown with their standard errors in Table 3. Recall that a positive coefficient reflects adverse effects of the famine since famine reduced cohort size. To address the attenuation bias from positive selection for survival, we estimate this equation on the 90th percentiles, the means and the 10th percentiles of the distribution of outcomes. For almost all outcomes, the estimates are larger in magnitude and more precisely estimated for the 90th percentile shown in Panel A. For example, columns (1)-(3) show that for the early childhood cohort, the estimated correlation between famine exposure and the relevant outcome variable for the 90th percentile in Panel A is over twice the magnitude as the estimated correlation for the sample mean shown in Panel B. The difference is larger when compared to the 10th percentile shown in Panel C.

We focus the discussion of the results on the 90th percentile in Panel A. For the in-utero cohort, the estimates show that exposure to a famine that decreased average famine cohort size by 1% reduced height by 0.024%, weight by 0.053%, and educational attainment by 0.174%. These estimates are statistically significant at the 5% or 1% levels. The coefficients for WFH, BMI and labor supply have similar signs but are not statistically significant. For the early childhood cohort, exposure to a famine that decreased the famine cohort size by 1% is associated with a 0.027% reduction in height, a 0.1% reduction in weight, a 0.026% in WFH and a 0.01% decrease in BMI. These estimates are statistically significant at the 10% or 5% levels.

4.2 First Stage Estimates

The estimates from equation (3) are shown in Table 4. Columns (1) and (4) show the correlation between the first set of instruments and the endogenous variables. Columns (2) and (5) show the correlation between the second set of instrument and the endogenous variables. Columns (3) and (6) shows the first stage estimates using all of the instruments. The instruments are strongly significant and inclusion of both sets of instruments in columns (3) and (6) does not change the signs or greatly affect the first stage coefficients.

4.3 Two Stage Least Squares

Table 5 shows the 2SLS estimates. Like the OLS estimates, the effects of the famine are more adverse and more precisely estimated for the 90th percentile. Once again, we focus our discussion on those results in Panel A. For the in-utero cohort, exposure to a famine that reduces famine cohort size by 1% decreases height by 0.047%, weight by 0.065%, WFH by 0.018%, BMI by 0.008%, and educational attainment by 0.238%. The estimates are statistically significant at the 1% or 5% levels with the exception of those for WFH and BMI, which are significant at the 20% level. For the early childhood cohort, exposure to a famine that decreased the famine cohort by 1% caused a 0.045% reduction in height, a 0.138% reduction in weight, a 0.033% reduction in WFH, a 0.007% reduction in BMI, and a 0.387% reduction in labor supply. These estimates are statistically significant at the 1% or 5% levels with the exception of the estimate for BMI, which is significant only at the 20% level.

4.4 Selection or Heterogeneous Effects?

Based on previous studies, we entered this study with the presumption that there is non-random selection for survival (Gorgens et al., 2002; Bozzoli, Deaton and Quintana-Domeque, 2008). The finding that the estimated effects are more adverse and larger in magnitude for individuals on the 90th percentile of the distribution of outcomes is consistent with this. However, the same pattern could exist if there are heterogenous effects of exposure to famine such that individuals on the upper quantiles of the distribution are more adversely affected. To the best of our knowledge, there is no evidence from the medical or epidemiological literature arguing this case. Without a concrete prediction of the pattern of heterogenous effects across quantiles, we will not be able to directly test for them or to rule them out empirically. However, we can look for suggestive evidence in the data based on a few reasonable, albeit crude assumptions.

First, we can compare the effect of exposure on the two extreme tails of the distribution. If those that survived were on average the "strongest" or "most able" individuals of the population, and these attributes are correlated with the outcomes of interest, then we would expect the estimated effect of exposure to vary for the lower quantiles but to stabilize for the upper quantiles. In contrast, if the results on the 90th percentile are driven solely by heterogenous effects, *and* the occurrence of heterogenous effects is symmetric for both tails of the distribution, then the effects should be very different for individuals on the lower *and* upper quantiles relative to those on the mean. Figures 5A and 5B plot the estimated effects and their 95% confidence intervals for exposure on labor supply for the early childhood cohort. The coefficients and standard errors are reported in Appendix Table A2. The OLS and 2SLS estimates exhibit similar patterns. The estimated impact of exposure is monotonically increasing with quantile. However, the rate of increase is much lower above the 30th percentile. The evidence is consisten with the presence of selection but does not rule out heterogenous effects.

Second, we provide a speculative argument. Height is typically positively correlated with other measures of health. And healthier individuals are widely believed to be more resilient to health shocks. If this is true, then heterogenous effects would cause the effects of famine exposure to be smaller in magnitude at the 90th percentile. In contrast, our results indicate the adverse effects are larger when estimated at the 90th percentile.

4.5 Average Effects

In Table 6, we calculate the average effect for those exposed to famine during early childhood. The 2SLS estimates for the effect on the 90th percentile from Table 5 Panel A are shown in rows (A) and (B) of Table 5. We use only the estimates that were statistically significant at the 5% or 1% levels. The estimates reveal the effect of exposure to a famine which reduced the population of the 1959 cohort size by 1%. The average effect of famine exposure is the product of the 2SLS estimate and the average percentage reduction of the 1959 cohort size. We measure the latter as one minus the ratio of the average 1959-61 cohort size cohort size to the 1952-54 cohort size. Hence, if we assume that 1959-61 cohort sizes would be equal to 1952-54 cohort sizes, then we calculate that the famine reduced the 1959-61 cohort size by 36%. Row (C) shows that on average, exposure to famine reduced height of the in-utero cohort by 1.7%, weight by 2.3% and educational attainment by 8.6%. Row (D) shows that for the early childhood cohort, exposure on average decreased height by 1.6%, weight by 5%, WFH by 1.2% and labor supply by 13.9%. The average effect in levels are shown in Row (F) and (G). They are the products of the sample means for the 90th percentile shown in Row (E) and the average percentage effects. They show that in-utero exposure on average reduced height by 2.8 cm, weight by 1.42 kg, and educational attainment by 0.6 years. For the early childhood cohort, exposure on average decreased height by 2.7 cm, weight by 3.03 kg, WFH by 0.004 kg/cm, and labor supply by 12.7 hours per week.

5 Interpretation

The main finding of this study is that in-utero and childhood exposure to famine significantly reduces adult health outcomes and labor supply. There are several caveats to the interpretation of the results. If grain rich regions are able to recover more quickly from the famine, then our strategy will cause us to underestimate the adverse effects of famine on survivors. As an estimate of the impact of childhood malnutrition, our results should be interpreted as the lower bound of the magnitude of adverse effects for the additional reason that we are not able to disentangle potentially offsetting effects from being part of a small cohort.

The second caveat arises from our using the results on the 90th percentiles to calculate the average effects of exposure. As we discussed earlier, we cannot empirically rule out the presence of heterogenous effects. This does not affect the validity of our estimates as the estimated impact for each quantile. But it means that caution should be born in mind when extrapolating the estimated effects of the higher quantiles for other segments of the population. A casual look at Figures 5A and 5B suggests that one conservative interpretation of our calculated average effects is that they are relevant for individuals above the 30th percentile.

The contribution of this paper does not hinge on the assumption that the larger effects for individuals on the 90th percentile is due to selection rather than the presence of heterogenous effects. Even if we assume that there are heterogenous effects and conservatively only interpret the estimated effects for those on the mean of the distribution, the instrumental variables strategy, in addressing measurement error and OVB, show that exposure to famine had negative long term effects on height, weight, weight-for-height and labor supply.

An interesting finding is that exposure to famine seems much worse for health outcomes than economic outcomes. This is consistent with the hypothesis that a reduced cohort size could benefit survivors in terms of labor market competition or access to public schools; and potentially offset the negative effects of famine through health channels. A direct examination of the effect of cohort size would be an interesting avenue for future research.

One result which we examined and for brevity did not report in the paper is the effect on hypertension, the most commonly used indicator of heart disease. It can be used to investigate the Barker Hypothesis which predicts that individuals who suffer severe malnutrition will recover and appear healthy until the middle of their life, when they will be more likely to suffer from conditions such as coronary heart disease. We did not find any evidence that exposure to famine increased the rates of hypertension for either those who were exposed as children or those who were exposed in-utero. In addition to the 1989 survey, we used a later wave of the CHNS in 1997, when those born during 1954-61 would have been 36-43 years of age. One possible explanation is that the effects of famine are non-linear in age so that we may only observe the effects as the survivors reach their 50s and 60s. We plan to use future waves of the CHNS to investigate this hypothesis. In addition, we plan to use a similar empirical strategy with mortality inferred from comparing the 2000 and 2010 Chinese Population Censuses to examine the effect of famine exposure on life expectancy (conditional on survival into adulthood). The later censuses will also allow us to investigate whether childhood exposure to famine causes survivors to exit the labor force earlier than those who were never exposed.

6 Conclusion

This paper attempts to overcome significant empirical difficulties to estimate the causal impact of childhood exposure to famine. Our estimates show that exposure in-utero and as very young children can have severe adverse consequences on adult health, educational attainment and labor supply. The instrumental variables strategy we use comes from the specific institutional context of China's Great Famine and cannot be easily used to evaluate the impact of other famines. However, examining the impact on the upper quantiles of the distribution of outcomes to address positive selection for survival can be easily applied to any context that experienced similar patterns of selection. It is important to note that China's Great Famine is far from the most severe in terms of mortality as a fraction of total population. Applying this method to famines with higher mortality rates and more severe selection should make a larger difference than in our context.

In addition to providing a sense of the magnitude of the long term damage caused by one of the largest human devastations in history, the findings of this study provide generalizable insights on the consequences of severe malnutrition during the early stages of life. They suggest that reducing nutritional deprivation in utero and during early childhood should be considered as an extremely worthwhile objective for policy makers.

.

References

[1] Aizer, Anna, Laura Stroud and Stephen Buka (2009) "Poverty, Stress and the Intergenerationsal Transmission of Human Capital" Brown University Working Paper.

[2] Almond, Douglas and Mazumder, Bhashkar. 2005. "The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data." *American Economic Review* 95(2): 258-62.

[3] Almond, Douglas, Kenneth Y. Chay and David S. Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics* 120(3): 1031-83.

[4] Almond, Douglas, Lena Edlund, Hongbin Li and Junsen Zhang. 2006. "Long-term effects of the 1959-1961 China Famine: Mainland China and Hong Kong." Columbia University Working Paper.

[5] Angrist, Joshua, Victor Lavy and Annalia Schlosser. 2006. "New Evidence on the Causal Link Between the Quantity and Quality of Children." MIT Working Paper.

[6] Ashton, Basil, Kenneth Hill, Alan Piazza, and Robin Zeitz. 1984. "Famine in China, 1958-1961." *Population and Development Review* 10 (December): 613-45.

[7] Banister, Judith. 1987. *China's Changing Population*. Stanford, CA: Stanford University Press.

[8] Barker, DJP. 1995. "Fetal Origins of Coronary Heart Disease." *British Medical Journal* 311(6998): 171-174.

[9] Becker, Gary S. and Lewis, H. Gregg, 1973. "On the interaction between the quantity and quality of children." *The Journal of Political Economy* 81(2) Part 2: S279-S288.

[10] Becker, Gary S. and Tomes, Nigel. 1976. "Child endowments and the quantity and quality of children." *The Journal of Political Economy* 84(4) Part 2: S143-S162.

[11] Becker, Jasper, 1996, *Hungry Ghosts: China's Secret Famine*, London: John Murray Publishers

[12] Behrman, Jere. 1996. "The Impact of Health and Nutrition on Education." *World Bank Research Observer* 11: 23–37.

[13] Behrman, Jere R., and Mark R. Rosenzweig. 2005. “The Returns to Birth Weight.” *Review of Economics and Statistics* 86: 586 – 601.

[14] Black, Sandra E., Devereux, Paul J. and Kjell G. Salvanes. 2004. “The more the merrier? The effect of family composition on children’s education.” NBER Working Paper No. W10720.

[15] Black, Sandra E., Devereux, Paul J. and Salvanes, Kjell G. 2005. “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes.” NBER Working Paper No. W11796.

[16] Bleakley, Hoyt. 2006. “Disease and Development: Evidence from Hook Worm Eradication in the American South.” *Quarterly Journal of Economics.*

[17] Bloom, David E., David Canning, and Jaypee Sevilla. 2001. “The Effect of Health on Economic Growth: Theory and Evidence.” NBER Working Paper No. 8587.

[18] Bozzoli, Carlos and Angus Deaton and Climent Quintana-Domeque. "Adult Height and Childhood Disease." *Demography* 46.4 (2009): 647-669.

[19] Brandt, Loren, Aloysius Siow and Carl Vogel. 2008. “Large Shocks and Small Changes in the Marriage Market for Famine Born Cohorts in China,” University of Toronto Working Paper.

[20] Brinkley, Garland. 1994. “The Economic Impact of Disease in the American South, 1860–1940,” Ph.D. Dissertation, University of California at Davis.

[21] Brown, A.S., J. van Os, C. Driessens, H.W. Hoek and E.S. Susser. 2000. “Further Evidence of Relation between Prenatal Famine and Major Affective Disorder.” *American Journal of Psychiatry* 157(2): 190-5.

[22] Case, Anne, Darren Lubotsky and Christina Paxson. 2002. “Socioeconomic Status and Health in Childhood: The Origins of the Gradient.” *The American Economic Review* 92(5): 1308-34.

[23] Case, Anne, Angela Fertig and Christina Paxson. 2004. “The Lasting Impact of Childhood Health and Circumstance.” Princeton Working Paper.

[24] Chang, Gene Hsin, and Guanzhong James Wen. 1997. "Communal Dining and the Chinese Famine of 1958–1961." *Economic Development and Cultural Change* 46 (October): 1–34.

[25] Chen, Y. and L-A. Zhou. 2002."The Long Term Health and Economic Consequences of 1959-1961 Famine in China". SSRN Working Paper 841764. Revised 2007. "The long term health and economic consequences of 1959-1961 famine in China." *Journal of Health Economics*, 26, pp.659-681.

[26] Coale, Ansley J. 1981. "Population Trends, Population Policy and Population Studies in China." *Population and Development Review* 7 (March): 85-97.

[27] Curie, Janet and Hyson, Rosemary. 1998. "Is the Impact of Health Shocks Cushioned by Socio-

[28] Economic Status? The Case of Low Birthweight." UCLA Working Paper.

[29] Curie, J. and Madrian, B. 1999. "Health, health insurance, and the labor market". In Ashenfelter, O. and Card, D. (eds.), *Handbook of Labor Economics*, vol. 3. North Holland: Amsterdam, pp. 3309-3415.

[30] Currie, Jane and Mark Stabile. 2004. "Socioeconomic Status and Health: Why is the Relationship Stronger for Older Children?" *American Economic Review* 93(5): 1813-23.

[31] Cutler, David, Angus Deaton and Adriana Lleras-Muney (2006) "The Determinants of Mortality" *Journal of Economic Perspectives*, 20(3), pp: 97-120.

[32] De Brauw, Alan and Gile, John "Migrant Opportunity and the Educational Attainment of Youth in Rural China" Michigan State University Working paper, 2006.

[33] Doblhammer, Gabrielle. 2002. "Differences in Lifespan by Month of Birth for the United States: The Impact of Early Life Events and Conditions on Late Life Mortality." Max Planck Institute for Demographic Research Working Paper WP 2002-019.

[34] Easterlin, Richard A. 1980. *Birth and Fortune: The Impact of Numbers on Personal Welfare*. The University of Chicago Press.

[35] Fogel, R. 1994. "Economic growth, population theory and physiology: The bearing of long-term processes on the making of economic policy." *American Economic Review* 84 (3), 369–395.

[36] Fogel, Robert and Dora Costa. 1997. "A Theory of Technophysio Evolution, with Some Implications for Forecasting Population, Health Care Costs, and Pension Costs." *Demography* 34(1), 49-66.

[37] Fogel, R., S. L. Engerman, and J. Trussell (1982). "Exploring the uses of data on height." *Social Science History* 6 (4), 401–421.

[38] Friedman, G. C. 1982. "The height of slaves in Trinidad." *Social Science History* 6 (4), 482–515.

[39] Galor, Oded and Weil, David. 2000. "Population, Technology, and Growth: From Malthusian stagnation to demographic transition and beyond." *The American Economic Review* 90(4): 806-828.

[40] Gertler, Paul and Simon Boyce. 2001. "An Experiment in Incentive-based Welfare: The Impact of PROGRESA on Health in Mexico." University of California Berkeley Working Paper.

[41] Giles, John, Albert Park and Juwei Zhang. 2006. "The Great Proletarian Cultural Revolution, Disruptions to Education, and Returns to Schooling in Urban China" Michigan State University.

[42] Glewwe, Paul, and Jacoby, Hanan. 1995. "An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition," *Review of Economics and Statistics* 77: 156–169.

[43] Glewwe, Paul, Hanan Jacoby and Elizabeth M. King. 2001. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics* 81: 345–368.

[44] Gorgens, T., Meng, X., and Vaithianathan, R. (2002, revised in 2007) "Selection and stunting effects of famine: Case study of the Great Chinese Famine." ANU Working Paper and IZA Working Paper.

[45] Gregory, Robert and Xin Meng. Mimeo. "The impact of interrupted education on earnings: the educational cost of the Chinese Cultural Revolution". ANU Working Paper.

[46] Hazan, M. and Berdugo, B. 2002. "Child Labor, Fertility and Economic Growth," *The Economic Journal*, 112(October): 467-475.'

[47] Hoddinott, J. and B. Kinsey. 2001. "Child growth in the time of drought." *Oxford Bulletin of Economics and Statistics* 63(4): 409–436.

[48] Hulshoff Pol, H.E., H.W. Hoek, E, Susser, A.S. Brown, A. Dingemans, H.G. Schnack, N.E. van Haren, L.M. Pereira Ramos, C.C. Gispen-de Wied and R.S. Kahn. 2000. "Prenatal Exposure to Famine and Brain Morphology in Schizophrenia." *American Journal of Psychiatry* 157(7): 1170-2.

[49] Johnson, D. Gale. 1998. "China's Great Famine: Introductory Remarks." *China Economic Review* 9(Fall): 103-9.

[50] Krueger, R. H. 1969. "Sole long-term effects of severe malnutrition in early life." *Lancet* 2: 514–517.

[51] Kuh, D.J. and M.E. Wadsworth. 1993. "Physical Health Status at 36 Years in a British National Birth Cohort." *Social Science and Medicine* 37(7): 905-916.

[52] Wei, Li and Dennis Yang. 2005. "The Great Leap Forward: Anatomy of a Central Planning Disaster." *Journal of Political Economy* 113(4): 840-77.

[53] Lin, Justin Yifu. 1990. "Collectivization and China's Agricultural Crisis in 1959-1961." *Journal of Political Economy* 98: 1228-52.

[54] Luo, Zhehui, Ren Mu, and Xiaobo Zhang. 2006. "The Long Term Health Impact of the Chinese Great Famine." Michigan State University Working Paper.

[55] Maccini, Sharon and Dean Yang. 2008. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall" *American Economic Review*, Forthcoming.

[56] Marmot, M, Shipley, Bruner and Heminway. 2001. "Relative Contributions of Early Life and Adult Socioeconomic Factors to Adult Morbidity in the Whitehall II Study." *Journal of Epidemiology and Community Health* 55: 301-307.

[57] Meng, Xin and Qian, Nancy, 2005. "The Long Run Health and Economic Consequences of Famine on Survivors: Evidence from China's Great Famine," Yale University Working Paper.

[58] Meng, Xin Nancy Qian and Pierre Yared (2009) "The Institutional Causes of Famine in China, 1959-61" Yale University Working Paper.

[59] Micklewright, J. and S. Ismail. 2001. "What can child anthropometry reveal about living standards and public policy? An illustration from Central Asia." *Review of Income and Wealth* 37(1): 65–80.

[60] Miguel, Edward and Kremer, Michael. (2004) "Worms: Identifying impacts on education and health in the presence of treatment externalities." *Econometrica* 72(1): 159–217

[61] Moav, Omer. 2005. "Cheap Children and the Persistence of Poverty." *The Economic Journal* 115: 88-110.

[62] Neugebauer, Richard, Hans W. Hoek and Ezra Susser. 1999. "Prenatal Exposure to Wartime Famine and Development of Antisocial Personality Disorder in Early Adulthood." *JAMA* 281(5): 455-462.

[63] Oi, Jean. 1989. *State and Peasant in Contemporary China: The Political Economy of Village Government*. Berkeley: University of California Press, XXI.

[64] Ozanne, Susan E. and C. Nicholas Hales. 2004. "Catch-up Growth and Obesity in Male Mice." *Nature* 427(29): 411-412.

[65] Peng, Xizhe. 1987. "Demographic Consequences of the Great Leap Forward in China's Provinces." *Population and Development Review* 13: 639–70.

[66] Perkins, Dwight H. (1966), *Market Control and Planning in Communist China*, Cambridge, Massachusetts: Harvard University Press.

[67] Perkins, Dwight H., and Shahid Yusuf. 1984. *Rural Development in China*. Baltimore: Johns Hopkins University Press (for The World Bank).

[68] Persico, Nicola, Andrew Postlewaite and Dan Silverman. 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Poliical Economy* 112(5): 1019-53.

[69] Qian, Nancy. 2006. "Quantity-Quality and the One Child Disadvantage: Evidence from China". NBER Working Paper.

[70] Ravallion, Martin. 1997. "Famines and Economics." *Journal of Economic Literature* 35 (September): 1205-42.

[71] Ravelli, ACJ and JHP van der Meulen, RPJ Michels, C. Ostmond, DJP Marker, CN Haels and OP Bleker. 1998. "Glucose Tolerance in Adults After Prenatal Exposure to Famine." *The Lancet* 351: 173-177.

[72] Ravelli, A.C.J., J.H. van de Meulen, C. Osmond, D.J. Barker and O.P. Bleker "Obesity at the Age of 50 in Men and Women Exposed to Famine Prenatally." *American Journal of Clinical Nutriton* 70(5): 811-6.

[73] Rosenzweig, Mark and Zhang, Junsen. 2006. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy." Yale Working Paper.

[74] Salama, P., F. Assefa, L. Talley, P. Spiegel, A. van der Veen, and C. A. Gotway. 2001. "Malnutrition, measles, mortality and the humanitarian response during a famine in Ethiopia." *Journal of the American Medical Association* 286(5): 563–571.

[75] Sen, Amartya. 1981. *Poverty and Famines: An Essay on Entitlement and Deprivation.* Oxford: Clarendon.

[76] Schultz, T. P. 2001. "Productive benefits of improving health: Evidence from low income countries." Paper for the meetings of the Population Association of America, Washington DC, March 29–31., Yale University.

[77] Schultz, T. Paul (2002). "Wage Gains Associated with Height as a Form of Health Human Capital." *AEA Papers and Proceedings* 92 (2): 349-353.

[78] Schultz, T.P. "Effects of Fertility Decline on Family Well-Being: Evaluation of Population Programs." Draft for MacArthur Foundation Consultation Meeting, 2005.

[79] Shi Xinzheng (2008) "Famine, fertility, and fortune in China" University of Michigan Working Paper.

[80] Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *Journal of Economic Perspectives* 13 (2): 145-166.

[81] Stanner, S.A.,K. Bulmer , C. Andres, O.E. Lantseva, V. Borodina, V.V. Poteen and J.S. Yudkin. 1997. "Does Malnutrition in-utero Determine Diabetes and Coronary Heart Disease in Adulthood? Results from the Leningrad Study, a Cross-sectional Study," *British Medical Journal* 315: 1342-8.

[82] Steckel, R. 1986. "A peculiar population: The nutrition, health and mortality of American slaves from childhood to maturity." *Journal of Economic History* 46 (3): 721–741.

[83] Strauss, John and Thomas, Duncan. 1998. "Health, Nutrition, and Economic Development," *Journal of Economic Literature* 36: 766–817.

[84] Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi Jaswadi, Nathan Jones, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, and Siswanto Wilop. 2003. "Iron Deficiency and the Well-Being of Older Adults: Early Results from a Randomized Nutrition Intervention." UCLA Working Paper.

[85] Unger, Jon, 1982, *Education Under Mao*, New York: Columbia University Press. 82(3): 393-408.

[86] Walker, Kenneth R.. 1984, *Food Grain Procurement and Consumption in China*, Cambridge: Cambridge University Press.

[87] Weil, David. 2005. "Accounting for the Effect of Health on Economic Growth." Brown University Working Paper.

[88] West, Lorraine and Zhao, Yaohui. *Rural-urban migration in China: the past and present*. Institute for East Asian Studies, University of California, Berkeley, 2000.

[89] Wen, Guanzhong James. 1993. "Total Factor Productivity Change in China's Farming Sector: 1952–1989." *Econonomic Development and Cultural Change* 42 (October): 1–41.

[90] Yang, Dali L. 1996. *Calamity and Reform in China: State, Rural Society, and Institutional Change since the Great Leap Famine*. Stanford, CA: Stanford University Press.

[91] Yao, Shujie. 1999. "A Note on the Causal Factors of China's Famine in 1959-1961." *Journal of Political Economy* 107(6), pt 1: 1365-69.

Table 1: The Effect of Pre-Famine Grain Production on Famine Mortality and Population

Coefficients of provincial level grain production; coefficients for the interaction term of provincial level grain production in 1959 and a dummy variable for 1960; and coefficients of the interaction term of provincial level grain production during 1954-58 and a dummy for 1960

	Dependent Variables											
	LnTotalPop		LnDeaths		LnTotalPop		LnDeaths		LnTotalPop		LnDeaths	
	Non-Famine	Famine	Non-Famine	Famine	All	All	All	All	All	All	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ln Grain Prod	0.256	0.00375	-0.261	0.373								
	(0.172)	(0.0859)	(0.165)	(0.445)								
Ln Grain Prod 1959 x 1960 Dummy					-0.0804	-0.0685	0.0797	0.0702				
					(0.0337)	(0.0257)	(0.104)	(0.101)				
Ln Grain Prod 1954-58 x 1960 Dummy									-0.0800	-0.0659	0.0995	0.0817
									(0.0305)	(0.0242)	(0.106)	(0.103)
Controls												
Province x Year Time Trends	N	N	N	N	N	Y	N	Y	N	Y	N	Y
Government Spending on Agriculture	N	N	N	N	N	Y	N	Y	N	Y	N	Y
Government Spending on Health, Edu & Science	N	N	N	N	N	Y	N	Y	N	Y	N	Y
Number of Teachers in Primary and Secondary	N	N	N	N	N	Y	N	Y	N	Y	N	Y
Observations	416	84	369	81	499	357	448	333	503	357	452	333
R-squared	0.970	0.999	0.564	0.661	0.969	0.974	0.558	0.764	0.970	0.974	0.561	0.765

All regressions control for year and province fixed effects.

Standard errors are clustered at the province level.

Source: Meng, Qian and Yared (2009) Note: The historical data used in this regression is at the province and year level.

Table 2: Descriptive Statistics

	Birth Cohort									
	All (1)		Before 1955 (2)		1955-1958 (3)		1959-61 (4)		After 1961 (5)	
Variable	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
Height	337	159.546	171	159.192	59	159.367	39	158.802	68	161.020
		(0.326)		(0.456)		(0.765)		(1.085)		(0.671)
Weight	335	55.485	169	55.673	59	55.636	39	54.644	68	55.369
		(0.346)		(0.493)		(0.908)		(1.107)		(0.636)
WFH	334	0.347	169	0.349	58	0.350	39	0.343	68	0.343
		(0.002)		(0.002)		(0.004)		(0.006)		(0.003)
BMI	334	21.768	169	21.920	58	21.935	39	21.634	68	21.324
		(0.101)		(0.140)		(0.245)		(0.377)		(0.182)
Years of Education	333	5.347	178	4.929	50	5.577	34	6.522	71	5.667
		(0.110)		(0.158)		(0.247)		(0.272)		(0.215)
Total Hours Worked Per Week	313	73.919	174	74.505	49	71.974	31	69.220	59	76.272
		(1.883)		(2.760)		(4.389)		(5.775)		(3.390)

Note: Observations are county-birth year cells.

Table 3: The OLS Estimates of the Effect between Famine Exposure

Coefficients of the interaction terms between the logarithm of 1959-61 cohort size and birth year categorical variables

	Dependent Variables					
	LnHeight	LnWeight	WFH	BMI	LnEduYrs	LnTotWk
	(1)	(2)	(3)	(4)	(5)	(6)
A. 90th Percentile						
Sample Mean (Not Logged)	165.291	61.191	0.376	23.500	6.797	90.802
LnFampop x Born 1959-61	0.024	0.053	0.017	0.008	0.174	0.042
	(0.014)	(0.033)	(0.012)	(0.006)	(0.090)	(0.179)
LnFampop x Born 1955-58	0.027	0.101	0.026	0.009	-0.065	0.098
	(0.008)	(0.022)	(0.008)	(0.004)	(0.124)	(0.119)
Observations	337	335	334	334	333	313
B. Mean						
Sample Mean (Not Logged)	159.546	55.485	0.347	21.768	5.347	73.919
LnFampop x Born 1959-61	0.012	0.037	0.007	0.000	-0.048	0.033
	(0.009)	(0.023)	(0.007)	(0.000)	(0.064)	(0.158)
LnFampop x Born 1955-58	0.010	0.051	0.011	0.000	0.004	-0.005
	(0.005)	(0.017)	(0.005)	(0.000)	(0.096)	(0.098)
Observations	337	335	334	334	333	313
C. 10th Percentile						
Sample Mean (Not Logged)	153.809	50.100	0.320	20.023	4.132	57.633
LnFampop x Born 1959-61	-0.004	0.001	0.000	-0.002	-0.339	0.009
	(0.010)	(0.026)	(0.007)	(0.004)	(0.068)	(0.188)
LnFampop x Born 1955-58	-0.007	0.010	0.000	-0.001	0.012	-0.219
	(0.006)	(0.023)	(0.006)	(0.003)	(0.095)	(0.137)
Observations	337	335	334	334	333	313

All regressions control for birth year and birth county fixed effects. Regressions are population weighted. Standard errors are clustered at the province-year level.

Table 4: The First Stage Estimates of the Effect of Grain Suitability and Grain Sown Per Capita on Famine Exposure

Coefficients of the interaction terms between area suitable for grain cultivation and birth year categorical variables, and between the logarithm of per capita grain sown and birth year categorical variables

	Dependent Variables					
	Ln Fampop x Born 1959-61			Ln Fampop x Born 1955-58		
	(1)	(2)	(3)	(4)	(5)	(6)
Suitable Area x Born 1959-61	-2.699		-2.833	0.007		0.003
	(0.034)		(0.033)	(0.007)		(0.007)
Suitable Area x Born 1955-58	0.002		0.001	-2.680		-2.831
	(0.002)		(0.003)	(0.019)		(0.018)
Ln Sown Area x Born 1959-61		-46.007	-47.333		-0.898	-0.998
		(5.848)	(6.002)		(1.292)	(1.327)
Ln Sown Area x Born 1955-58		-0.090	-0.160		-49.910	-51.171
		(0.603)	(0.612)		(4.279)	(4.317)
Observations	11734	11826	11734	11734	11826	11734

All regressions control for birth year and birth county fixed effects. Regressions are population weighted. Standard errors are clustered at the province-year level.

Table 5: The 2SLS Estimates of the Effect of Famine Exposure

Coefficients of the interaction terms between the logarithm of 1959-61 cohort size and birth year categorical variables

	Dependent Variables					
	LnHeight	LnWeight	WFH	BMI	LnEduYrs	LnTotWk
	(1)	(2)	(3)	(4)	(5)	(6)
A. 90th Percentile						
Sample Mean (Not Logged)	165.291	61.191	0.376	23.500	6.797	90.802
LnFampop x Born 1959-61	0.0470	0.0645	0.0178	0.0083	0.2379	0.1119
	(0.0168)	(0.0390)	(0.0130)	(0.0068)	(0.1409)	(0.1394)
LnFampop x Born 1955-58	0.0454	0.1376	0.0325	0.0067	0.0926	0.3868
	(0.0114)	(0.0257)	(0.0083)	(0.0044)	(0.1213)	(0.1596)
Observations	337	335	334	334	333	313
B. Mean						
Sample Mean (Not Logged)	159.546	55.485	0.347	21.768	5.347	73.919
LnFampop x Born 1959-61	0.0242	0.0483	0.0071	0.0000	-0.0718	0.0180
	(0.0118)	(0.0289)	(0.0082)	(0.0001)	(0.0849)	(0.1316)
LnFampop x Born 1955-58	0.0241	0.0661	0.0135	0.0000	0.0984	0.2401
	(0.0078)	(0.0173)	(0.0049)	(0.0000)	(0.0992)	(0.1430)
Observations	337	335	334	334	333	313
C. 10th Percentile						
Sample Mean (Not Logged)	153.809	50.100	0.320	20.023	4.132	57.633
LnFampop x Born 1959-61	-0.0054	0.0063	-0.0051	0.0083	-0.5264	-0.1794
	(0.0136)	(0.0363)	(0.0086)	(0.0068)	(0.1164)	(0.1658)
LnFampop x Born 1955-58	-0.0023	0.0078	0.0000	0.0067	-0.0204	-0.0349
	(0.0089)	(0.0292)	(0.0073)	(0.0044)	(0.1265)	(0.1794)
Observations	337	335	334	334	333	313

All regressions control for birth year and birth county fixed effects. Regressions are population weighted.

Standard errors are clustered at the province-year level.

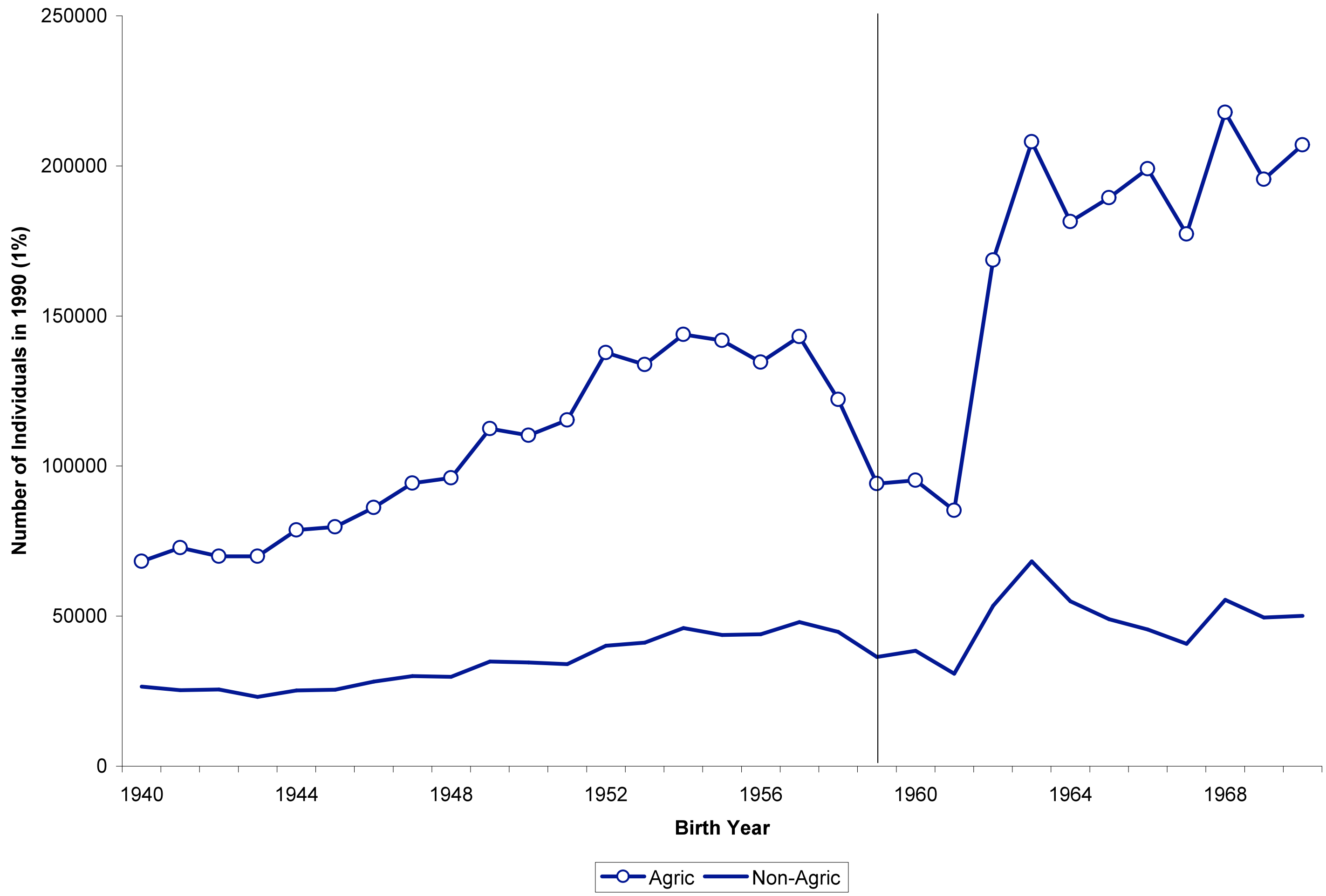
Table 6: The Effect of the Great Famine

Average effects of the famine calculated with the sample mean, 2SLS estimates on the 90th Percentile, and the average estimated intensity of famine

		Variables				
		Height	Weight	WFH	Edu Yrs	Total Work Hrs
		(1)	(2)	(3)	(4)	(5)
A	Coefficient for LnFampop x Born 1959-61	0.0470	0.0645		0.2379	
B	Coefficient for LnFampop x Born 1955-58	0.0454	0.1376	0.0325		0.3868
	% Effect of famine = 2SLS Coefficient x (1 - 1959-61 cohort size/1952-54 Cohort Size)					
C	In Utero Cohort: A x -0.36	-0.017	-0.023		-0.086	
D	Early Childhood Cohort: B x -0.36	-0.016	-0.050	-0.012		-0.139
	Level Effect of famine = % Effect x Sample Mean					
E	Sample Mean for 90th Percentile	165.29	61.19	0.38	6.80	90.80
F	In Utero Cohort: E x C	-2.80	-1.42		-0.58	
G	Early Childhood Cohort: E x D	-2.70	-3.03	-0.0044		-12.64

Notes: Estimates in Rows (A) and (B) show 2SLS estimates on the 90th percentile from Table 4 Panel A. These calculations assume that absent the famine, 1959-61 cohort sizes would be equivalent to the average size of 1952-54 cohorts.

Figure 1: Population by Rural/Urban and by Birth Year



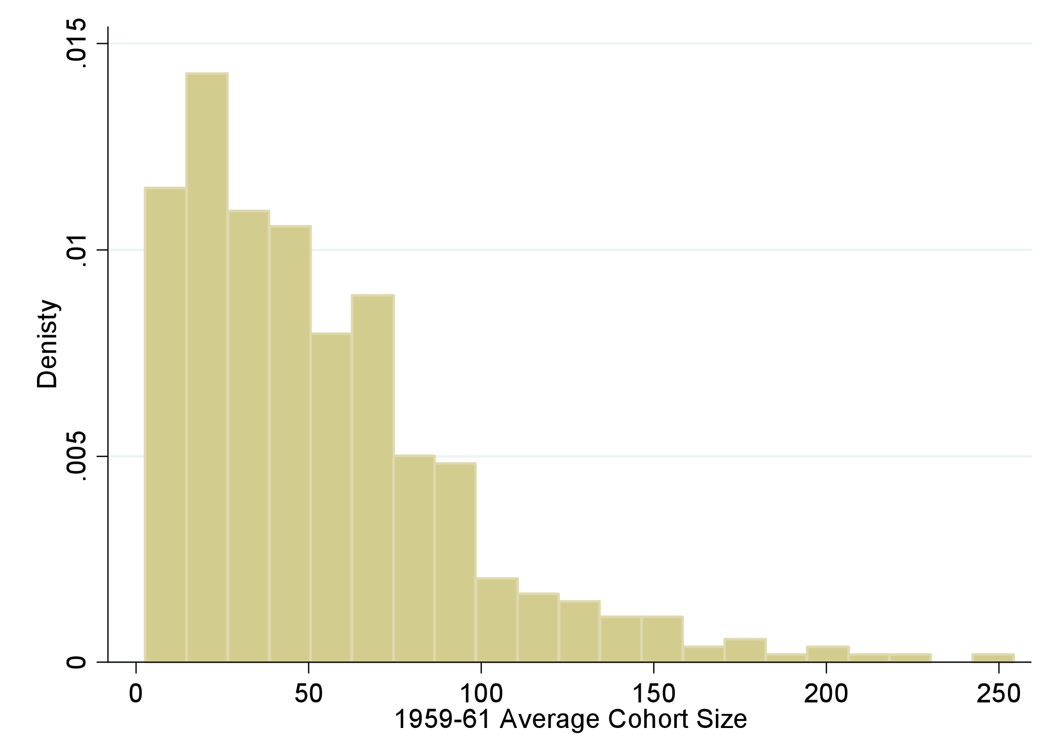
Source: Authors' Calculation

Figure 2: The Correlation between Suitability for Grain Cultivation and Cohort Sizes for Agricultural and Non-Agricultural Populations in All Counties in China

Coefficients of the interactions of the fraction of land suitable for grain and birth year dummy variables

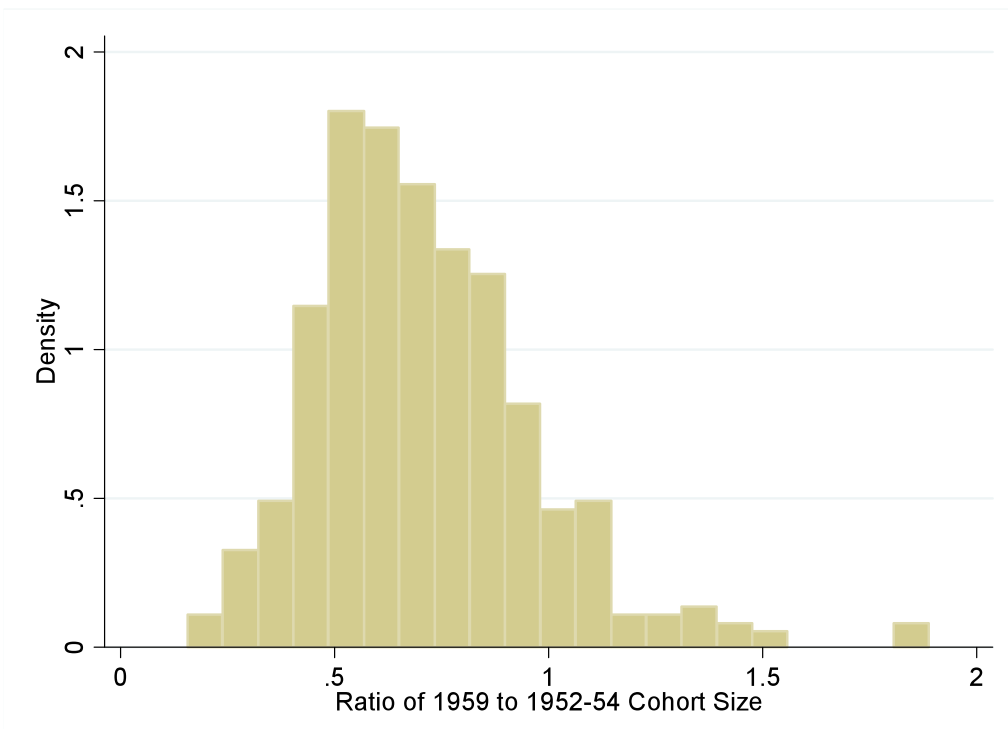
Source: Meng and Qian (2009) Note: This sample excludes cities. Counties in China contain both agricultural and non-agricultural households.

Figure 3A: Famine Intensity across Counties Measured as 1959-61 cohort size



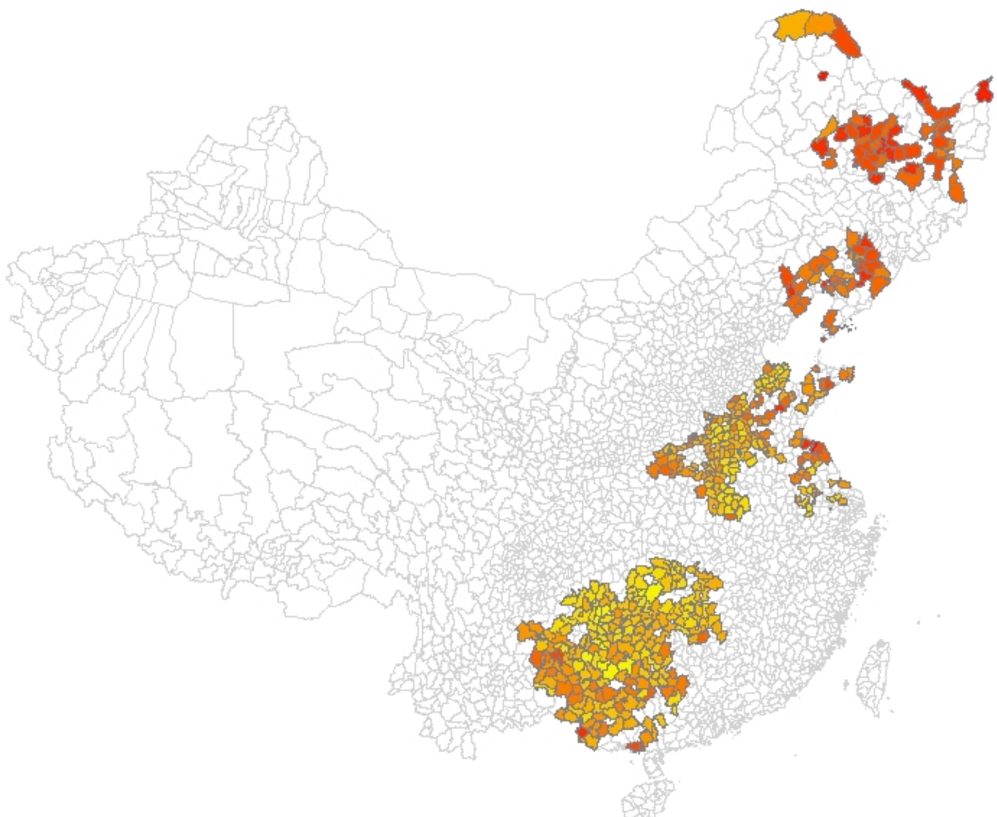
Source: Authors' calculation

Figure 3B: Famine Intensity across Counties
Measured as the Ratio of 1959-61 cohort size to1952-54 Cohort Size



Source: Authors' calculation

Figure 4: Map of County-Level Famine Intensity



Note: *Lighter* shading reflects greater famine intensity; Source: Authors' calculation

Figure 5A: Estimated OLS Effect of Exposure on Labor Supply by Quantile
Coefficients of the interaction terms between the logarithm of 1959-61 cohort size and a dummy variable for being born during 1954-58; and their 95% Confidence Intervals

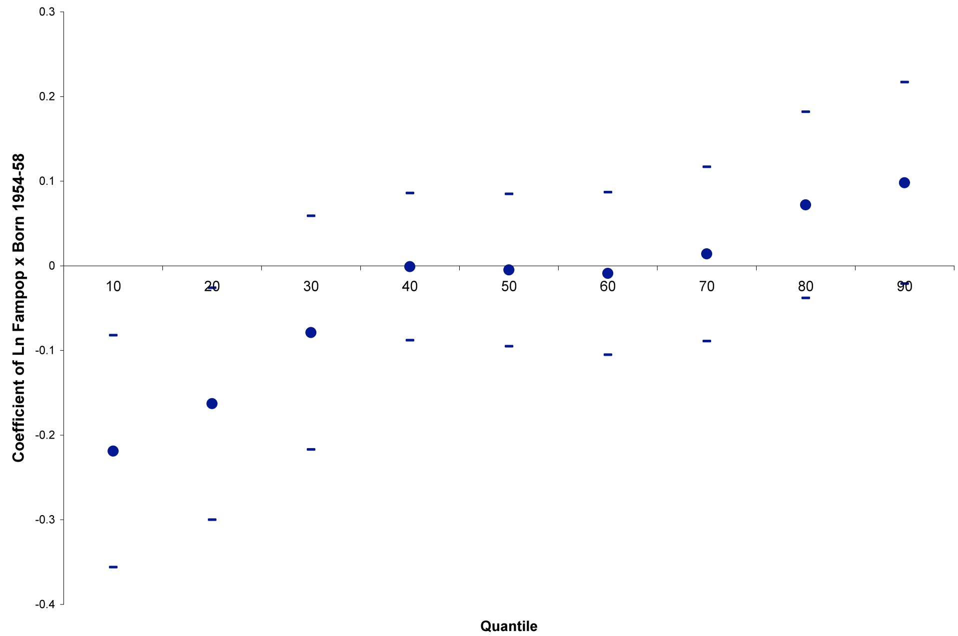


Figure 5B: Estimated 2SLS Effect of Exposure on Labor Supply by Quantile
Coefficients of the interaction terms between the logarithm of 1959-61 cohort size and a dummy variable for being born during 1954-58; and their 95% Confidence Intervals

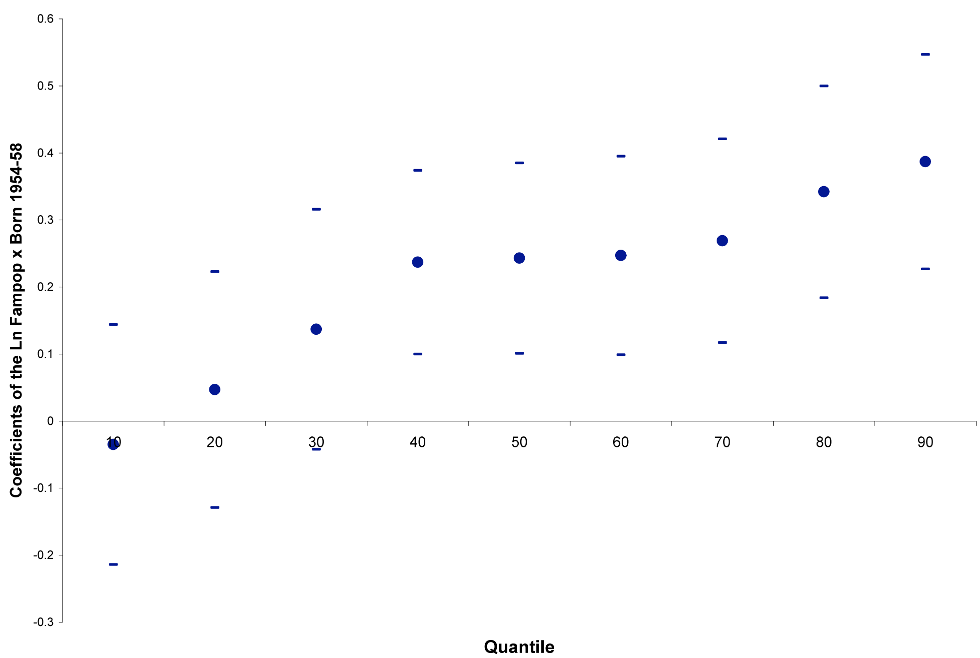


Table A1: The Correlation between Grain Suitability and Per Capita Grain Sown with Cohort Size

Coefficients of the interaction terms between of grain suitability and birth year dummies, and between per capita grain sown and birth year dummies

	Dependent Variable: Ln Population	
	National Sample	
	Agric	Non Agric
	(1)	(2)
Suitability * Born 1943	0.066	0.023
	(0.031)	(0.052)
Suitability * Born 1944	0.096	0.046
	(0.028)	(0.067)
Suitability * Born 1945	0.111	0.065
	(0.036)	(0.075)
Suitability * Born 1946	0.107	0.093
	(0.029)	(0.071)
Suitability * Born 1947	0.122	0.096
	(0.039)	(0.069)
Suitability * Born 1948	0.178	0.132
	(0.032)	(0.066)
Suitability * Born 1949	0.152	0.066
	(0.033)	(0.063)
Suitability * Born 1950	0.036	0.135
	(0.041)	(0.068)
Suitability * Born 1951	0.166	0.193
	(0.040)	(0.067)
Suitability * Born 1952	0.195	0.190
	(0.037)	(0.064)
Suitability * Born 1953	0.148	0.198
	(0.037)	(0.068)
Suitability * Born 1954	0.155	0.206
	(0.031)	(0.069)
Suitability * Born 1955	0.018	0.141
	(0.039)	(0.066)
Suitability * Born 1956	0.093	0.153
	(0.034)	(0.068)
Suitability * Born 1957	0.030	0.211
	(0.038)	(0.072)
Suitability * Born 1958	-0.122	0.102
	(0.047)	(0.074)
Suitability * Born 1959	-0.202	0.144
	(0.067)	(0.069)
Suitability * Born 1960	-0.296	0.012
	(0.074)	(0.063)
Suitability * Born 1961	-0.127	0.101
	(0.062)	(0.062)
Suitability * Born 1962	-0.013	0.043
	(0.037)	(0.057)
Suitability * Born 1963	0.064	0.014
	(0.037)	(0.060)
Suitability * Born 1964	-0.025	-0.083
	(0.027)	(0.074)
Suitability * Born 1965	-0.020	-0.101
	(0.026)	(0.089)
Suitability * Born 1966	0.010	-0.081
	(0.029)	(0.127)
Observations	46212	35175

Regression controls for birth year and birth county fixed effects, and is population weighted. Standard errors are clustered at the province x year level. This Table is taken from Meng and Qian (2009).

Table A2: The Estimates of the Effects of Famine on Labor Supply by Quantile

Coefficients of the interaction terms between the logarithm of 1959-61 cohort size and birth year categorical variables

	Dependent Variable: Ln Total Hours Worked Per Week								
Quantiles	10%	20%	30%	40%	50%	60%	70%	80%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. OLS									
LnFampop x Born 1959-61	0.009	0.040	0.078	0.111	0.061	0.024	0.001	0.036	0.042
	(0.188)	(0.180)	(0.178)	(0.179)	(0.151)	(0.135)	(0.145)	(0.170)	(0.179)
LnFampop x Born 1955-58	-0.219	-0.163	-0.079	-0.001	-0.005	-0.009	0.014	0.072	0.098
	(0.137)	(0.137)	(0.138)	(0.087)	(0.090)	(0.096)	(0.103)	(0.110)	(0.119)
Observations	313	313	313	313	313	313	313	313	313
B. 2SLS									
LnFampop x Born 1959-61	-0.179	-0.134	-0.076	0.048	0.058	0.057	0.065	0.095	0.112
	(0.166)	(0.162)	(0.152)	(0.151)	(0.137)	(0.140)	(0.136)	(0.138)	(0.139)
LnFampop x Born 1955-58	-0.035	0.047	0.137	0.237	0.243	0.247	0.269	0.342	0.387
	(0.179)	(0.176)	(0.179)	(0.137)	(0.142)	(0.148)	(0.152)	(0.158)	(0.160)
Observations	313	313	313	313	313	313	313	313	313

All regressions control for birth year and birth county fixed effects. Regressions are population weighted.

Standard errors are clustered at the province-year level.