Entrepreneurship and the School of Hard Knocks: Evidence from China's Great Famine

Junhong Chu, I.P.L. Png, & Junjian Yi*

May 2016

Abstract

Many entrepreneurs proclaim proudly to have learned from the School of Hard Knocks rather than business schools. These claims challenge the social value of investments in entrepreneurship education and training. Yet, there is little empirical evidence of the effect of earlier life challenges on subsequent success in entrepreneurship. Here, we exploit geographical differences in the intensity of China's Great Famine, 1959-61, as a quasi-natural experiment to identify the effect of hardship. We find robust evidence of more entrepreneurship in counties that experienced greater hardship during the Famine. We investigate whether the difference is due to selective culling or conditioning of personality, and find evidence of conditioning towards greater risk tolerance.

Keywords: Entrepreneurship, Hardship, Nurture, Famine, China JEL codes: L26, O15, O17

*National University of Singapore: junhong.chu@gmail.com, iplpng@gmail.com, junjian.yi@gmail.com. We thank Shangjin Wei, Chien-yi M. Png, Richard Lai, Sng Tuan Hwee and other NUS colleagues, participants at seminars at Cornell University and Syracuse University for helpful advice, and Xiong Xi for outstanding research assistance. This research was funded by Academic Research Fund grant MOE2013-T2-2-045 (R-313-000-109-112).

1 Introduction

> "By definition, an entrepreneur is one who takes risk. It's an attitude and an appetite, one which may be hardwired into one's personality. My grandfather referred to me as having an MBA from the School of Hard Knocks", Stephen Greer (2010).
>
> "From the military school of life – What does not kill me, strengthens me", Nietzsche (1990: 8).

Entrepreneur Stephen Greer transformed a two-person scrap metal business into a multinational enterprise with annual sales of $200 million. Interviewed by the *Harvard Business Review*, he first asserted that risk-taking is "hardwired", but then, referring to the "School of Hard Knocks", suggested that risk-taking and entrepreneurship are cultivated. Likewise, Nietzsche's characterization of the "military school of life" emphasizes both selection and cultivation.

Considering the effort and expense that governments worldwide are investing to promote entrepreneurship, it is essential to understand what drives people to start new high-growth businesses. In particular, does hardship contribute to entrepreneurship, and if so, how? Does hardship contribute by winnowing out weak people or does it contribute by cultivating individual personality and attitudes that support entrepreneurship?

The challenge in any study is causal inference. An experiment that randomly assigns some subjects to sustained hardship and others to normal lives, and then follows their subsequent careers seems inconceivable. Possibly because of this difficulty, there appears to be no research into the effect of prior life experience of hardship on entrepreneurship (Shane and Venkataraman 2000; Rauch and Frese 2007).

By contrast, considerable research has investigated the effect of previous experience of hardship on corporate leaders, particularly CEOs of U.S. listed companies. CEOs who grew up during the Great Depression rely more on internal funding and use less debt (Malmendier, Tate, and Yan 2011). CEOs with military experience use more leverage (Malmendier et al. 2011) and produce better shareholder returns in mergers and acquisitions (Lin et al. 2011). A related line of research investigates the effect of one-off shocks such as natural disasters in early life on CEOs (Bernile, Bhagwat, and Rau 2015). These studies address the possible endogeneity of previous life experience (e.g., voluntary military service) and CEO appointment by timing of birth and exogenous CEO turnover.

In light of the research by Malmendier et al. (2011) and Lin et al. (2011), it is reasonable to expect that previous experience of hardship might also affect engagement in entrepreneurship. Here, we investigate this issue by studying how China's Great Famine of 1959-61 affected individual personality and entrepreneurship. The famine followed the Chinese government's headlong rush to collectivize agriculture and ramp up industrial production. The famine exogenously imposed sustained hardship on the Chinese people to varying degrees by geography. We exploit county-level differences in the intensity of the famine to identify the effect of previous life experience of hardship on individual personality and entrepreneurship.

This study presents three challenges: (i) How to measure the degree of hardship caused by the famine, (ii) How to account for possible endogeneity in the statistical relation between hardship and entrepreneurship, and (iii) How to distinguish whether the relation between hardship and entrepreneurship is due to selective culling or conditioning.

During the period of study, the hukou (household registration) system severely restricted individual mobility within China. The restriction on mobility enables us to use data from the 1990 Population Census to construct an intuitive county-level measure of hardship. Following Yao (1999), Meng and Qian (2009), Meng, Qian, and Yared (2015), and Garnaut (2014), we calculate the population of each county by year of birth, and, fitting a county-specific linear trend to the cohorts born before or after the famine, project the counterfactual number of persons in the cohorts born during the famine. Then we construct the cohort loss as the difference between the projected and actually recorded number of people. Our first key construct is the rate of cohort loss (cohort loss divided by the projected population) in the famine birth cohorts, which represents the intensity of the famine and hardship in the county.

Motivated by prior research into the effect of heat on agricultural productivity (Ritchie and NeSmith 1991; Deschenes and Greenstone 2007; Schlenker and Roberts 2009), our second key construct represents the productivity of agriculture. We define the thermal agricultural productivity as the sum of the degree-days in the growing season.[1] We validate the two

[1]The agronomic concept of "degree-day" is the average daily temperature subject to a floor of 8 degrees Celsius and ceiling of 32 degrees Celsius.

key constructs, showing that the rate of cohort loss is negatively correlated with the thermal agricultural productivity in the previous year during normal periods, but positively correlated in famine periods. This accords with prior findings that the famine was due to excessive procurement of grain from rural areas by officials projecting output from pre-famine years and officials being unwilling or unable to adjust procurement targets or redistribute food during the famine (Kung and Chen 2011; Meng et al. 2015).

The various areas of China differed in the intensity of the Great Famine and subsequent entrepreneurship. We apply an instrumental variables (IV) strategy to account for possible endogeneity in the statistical relation between cohort loss and subsequent entrepreneurship. The instruments for the rate of cohort loss are the thermal agricultural productivity during famine years, and the difference in the thermal agricultural productivity between famine and normal years.

We find an economically and statistically significant relation between the severity of the famine, as represented by the rate of cohort loss among the famine birth cohorts in a county, and the degree of entrepreneurship four to five decades later. In a county with one standard deviation higher rate of cohort loss, an individual is roughly 56.8 percent more likely to engage in entrepreneurship. This effect is significant across all birth cohorts, and is more pronounced the younger the person was at the time of the famine. Further, we also find an economically and statistically significant relation between the severity of the famine and the degree of entrepreneurial success, as measured by entrepreneur's income.

How did greater hardship during the Great Famine give rise to more entrepreneurship in later years? We show that, by county, conditional on the population, the number of entrepreneurs increases with the severity of the famine. Hence, the relation between hardship and entrepreneurship cannot be completely explained by differences in mortality, and so must be due, at least in part, to conditioning of personality. Further, we find an economically and statistically significant relation between the rate of cohort loss and investments in shares. Based on previous research that characterizes risk aversion by investment in shares (Hvide and Panos 2014), we interpret the estimates as showing that hardship cultivated more tolerance of risk.

We rule out the relation between hardship and entrepreneurship as being due to selective culling or conditioning to raise self-confidence, tenacity, resilience, opportunism, or skirting of authority. We also reject several other alternative explanations. One is the famine tilting the gender ratio towards females and inducing people with daughters to engage more in entrepreneurship to counter the imbalance in marriage prospects (Wei and Zhang 2011). Another is the famine leaving more people without sons to support them in old age, and so, causing them to engage more in entrepreneurship and be self-sufficient in later years. Yet, another is the relation between hardship and entrepreneurship being due to local differences in the effectiveness of government and state capacity (Lu, Luan and Sng 2016).

Our findings contribute to a better-grounded appreciation of how hardship cultivates individual personality to foster engagement in entrepreneurship. Many entrepreneurs proclaim proudly to have graduated from the School of Hard Knocks rather than formal business schools. We show that sustained hardship conditions people to be more tolerant of risk, which induces increased entrepreneurship and earnings from entrepreneurship in later life. These results fit with a substantial literature which emphasizes the psychological underpinnings of entrepreneurship (Shane and Venkatraman 2000; Rauch and Frese 2007).

Our work also bears on the issue of whether entrepreneurs are born or nurtured. Using Swedish data, Lindquist, Sol, and Van Praag (2015) compare the effect of birth vis-à-vis adoptive parents on children's entrepreneurship and find that post-birth factors had double the effect of pre-birth factors. Consistent with their findings, we show that early life experience founds entrepreneurship in adulthood, and so, entrepreneurship can be nurtured.

Finally, our research contributes to understanding the longer term effects of famine. Most previous research emphasizes the negative consequences (Ravallion 1997). By contrast, Tan, Tan, and Zhang (2014) find that early childhood victims of the famine managed to shield their own children from reduction of cognitive ability. We identify a rebound on another dimension: hardship engenders more risk tolerance. Apparently, human psychology can adapt and, at least partially, ameliorate the long term effects of sustained deprivation.

2 Institutional Setting

In 1958, the Communist government of China, led by Chairman Mao Zedong, launched the Great Leap Forward, aiming to surpass the industrial output of Britain within 15 years and the United States after a further 15 years. The government collectivized agriculture and forced people to eat in communal canteens, and redirected up to 100 million people from agriculture to collective works and industrial production, while vigorously enforcing increased procurement of food from rural areas for distribution to urban residents and export to foreign countries (Li and Yang 2005).

The result was a famine, starting in the winter of 1959 and lasting until 1961, when Chairman Mao reluctantly reversed his policies.[2] The human toll in deaths and diminished fertility has variously been estimated as 18.5 million (Yao 1999), 32.5 million (Cao 2005), and 45 million (Becker 1996; Dikotter 2010). The government blamed the famine on bad weather: "Over the past two years, regions throughout China have suffered the serious effects of a natural calamity" (*People's Daily*, October 1, 1960, quoted in Yang (2012: 452)).

The severity of the famine during the Great Leap Forward was not uniform across China. Figure 1 depicts the severity by county, as represented by our construct, the relative rate of cohort loss, defined as the difference in the rate of cohort loss between famine and normal periods. Apparently, there was substantial geographical variation in the severity of the famine.

Meng et al. (2015) show that, nationally, China produced sufficient food to meet domestic consumption needs. Shortages of food in particular areas were due to the inflexibility of local government officials. They had set or agreed to targets for procurement, likely based on the previous year's harvest. In the face of shortages due to random fluctuations in weather, they could not or would not relax the procurement or redistribute food geographically. Kung and Chen (2011) attribute the localized famines in part to the career concerns of provincial officials. Officials with lower rank in the Communist Party (hypothesized as aspiring to higher rank) extracted about 3 percent more grain from their provinces than more senior officials.

Although the Chinese government supplied food to urban areas, people in the cities did not escape the famine. According to official Chinese government statistics, the ratio of the death rate during the peak of the famine in 1960 to the pre-famine death rate in 1957 was 1.6 and 2.6 in urban and rural areas respectively (Lin and Yang 2000: Table 2). In 1962, with

[2] China officially dates the famine as between 1959 and 1961, naming it "three years of natural disasters" (*san nian zi ran zai hai*) (Meng and Qian 2009).

an urban population of 2.68 million, Chongqing in the western province of Sichuan, was one of China's most populous cities.[3] A report of the People's Committee for Chongqing City for the period January to September 1962 reported that over 167,000 children suffered from malnutrition and rickets, while 100,000 were in feeble condition (Zhou 2012: Document 23).

In 1958, at the start of the Great Leap Forward, China established the household registration (hukou) system to restrict domestic migration, particularly to prevent migration from rural to urban areas. Under the hukou system, individuals have legal status only in their registered hukou place of residence. The hukou system regulates almost every aspect of an individual's life, including birth, marriage, divorce, public housing, education, and employment. Importantly, local governments issued food rations only to people with the local hukou. Furthermore, the government permits changes of hukou registration only under very stringent conditions. Indeed, the hukou system persists and remains controversial, and the Chinese government is only now taking tentative steps to reform the system.

3 Data

In this study, we compile data from multiple sources. Administratively, China is divided into provinces, which are further divided into prefectures, and which are further divided into counties. (The administrative division of China includes an overlapping category of "city", which might be at provincial, prefectural, or county level.) We carry out the analysis mainly by county, which is the lowest administrative level for which the government and survey researchers publish data.

Our principal sources of information on demography, business, individual employment, investments, and personality are: (i) 1990 Population Census for information on population by county and year of birth, used to estimate cohort losses, (ii) 2004 Economic Census for information on private-sector enterprises, and (iii) 2000 and 2005 Population Censuses and China Family Panel Studies (CFPS) 2010, China Household Finance Survey (CHFS) 2011, and China General Social Survey (CGSS) 2008, 2010, 2011, 2012, and 2013 for information on individual employment, investments, and personality traits.

[3] Chongqing became a municipality in 1997.

We use the 1990 Population Census rather than earlier or later censuses for various reasons. In matching the population, business, and individual data, we drop all counties which changed boundaries. The number of counties with changes in boundaries is over 800 between 1982 and 2004, and about 300 between 1990 and 2004. So, to minimize the loss of data, we prefer the 1990 to the 1982 Population Census. Further, we prefer the 1990 to the 2000 Population Census as migration was more limited at earlier times and the 1990 Census is considered to be more reliable (Lavely 2001; Zhang and Zhao 2006). We use the 2000 Census and the combination of the 1990 and 2000 Censuses in robustness tests.[4]

The 1990 Census reports only current residence and hukou registration. We limit to people whose county of residence is the same as their hukou registration and have lived in the county for five or more years, and assume that their birthplace is in that county. The 2000 Census reports place of birth and county of residence. We limit to people whose birthplace is in the county of residence. After dropping counties whose boundaries differ from those in 2004, the 1990 Census covers 2,265 counties, the 2000 Census covers 2,589 counties, and the combined 1990 and 2000 Censuses cover 2,200 counties.

We collect county-level data on private-sector enterprises from the 2004 Economic Census, which was the first comprehensive national record of business activity. We define a private enterprise as entrepreneurial if it is less than 10 years old and has fewer than 100 employees. The purpose of limiting the employment is to exclude large private businesses formed by spin off from the government or state-owned enterprises. In robustness checks, we define a private enterprise as entrepreneurial if it is less than 10 years old and has fewer than 500 employees. Our results are robust to this alternative measure.

Further, we collect individual data on employment and income from the 2005 Population Census, and define entrepreneurship in two ways – broadly as those who own an enterprise in the private sector or are self-employed, and narrowly as those who own a private-sector enterprise. We collect data on individual employment, investments, self-confidence, tenacity, resilience, and opportunism and skirting of authority from multiple household surveys – CFPS 2010, CHFS 2011, and CGSS 2008, 2010, 2011, 2012, and 2013.

From the China Meteorological Administration, we download daily instrumental weather records for 727 stations over the years 1951-70. We associate each county with the nearest

[4]The 1995 mini-Census is not available at individual level or county level.

weather station by the Euclidean distance from the county seat, and so, on average, match three counties to one station. As explained below, we use the weather to construct instruments, not as the explanatory variable of interest, which is severity of the famine at county level. Nevertheless, to be conservative and account for possible correlation among counties due to the common weather records or similarity in weather, we cluster standard errors by weather station when our measure of entrepreneurship is constructed based on the 2004 Economic Census and the 2005 Population Census. When we investigate the mechanisms for the positive relationship between famine severity and entrepreneurship, we use the various surveys, where we cluster standard errors by county because the counties in each survey are not contiguous to each other.

Table 1 presents summary statistics of the data. Table 1, panel (a), summarizes the 2004 Economic Census data for 2,194 counties. Entrepreneurial firms account for 47.1 percent (s.d. 16.0 percent) of all enterprises and 17.8 percent (s.d. 14.0 percent) of sales, and their median age is 37.7 months (s.d. 12.3 months), as compared with 47.4 months (s.d. 18.1 months) for all private enterprises. Entrepreneurial firms are more profitable. Even though they take up less than one-fifth of sales, they account for more than half of profit. The share of entrepreneurial firms in employment is 20.5 percent, which slightly exceeds their share of sales.

Table 1, panel (b), summarizes the 2005 Population Census data. To focus on the effects of the famine, we limit to people born before 1962. The Census covers over 639,443 such persons in 2,212 counties, of whom 0.7 percent own private enterprises (1.1 percent and 0.3 percent among males and females respectively), and 2.7 percent are self-employed (3.8 percent and 1.6 percent among males and females respectively).[5]

Table 1, panel (c), reports summary statistics for the surveys, limited to people born before 1962. These surveys cover 21,965 such individuals in 248 counties, of whom 2.1 percent own private enterprises, 4.3 percent have financial investments, and 2.7 percent participate in the stock market.

[5]We exclude counties whose boundaries changed between 1990 and 2005. The excluded counties are very similar to those included in our analysis. The percentages of individuals who owned private enterprises or were self-employed were respectively 0.8 and 3.0 in the excluded counties, as compared with 0.7 and 2.7 in the included counties. These differences are not statistically significant.

4 Constructs and Validation

To characterize the severity of the Great Famine in each county, we follow Yao (1999), Meng and Qian (2009), Meng et al. (2015), and use Garnaut (2014) and use the 1990 Population Census to construct the following measure. For each county and year of birth, the rate of cohort loss is the projected, counterfactual population who would have survived until 1990 less the actual population who did survive until 1990 divided by the projected population. Formally, for county, c, and year, y, our first key construct, the rate of cohort loss,

$$\mu_{cy} = \frac{\tilde{P}_{cy} - \hat{P}_{cy}}{\tilde{P}_{cy}} = 1 - \frac{\hat{P}_{cy}}{\tilde{P}_{cy}}, \tag{1}$$

where $\tilde{P}_{cy}$ and $\hat{P}_{cy}$ represent the projected and actual populations in 1990 respectively.

The famine affected individual health, education, employment, marriage, and other outcomes in the long term (Meng and Qian 2009; Mu and Zhang 2011). By taking a retrospective perspective from 1990, our construct for the severity of the famine, the rate of cohort loss, μ_{cy}, accounts for both immediate as well as long-term effects of the famine. We stress that the purpose of calculating the rate of cohort loss is not to measure excess mortality and depressed fertility during the famine as such. Rather, the objective is to represent differences in the geographical intensity of the famine in all its ramifications from a viewpoint in 1990. For instance, the rate of cohort loss encompasses lower wages and worse jobs to the extent that these are correlated with the difference between the projected and actual populations.

Referring to equation (1), we project the counterfactual population as follows. Use ordinary least squares (OLS) to regress the numbers of people born in the normal years, 1949-57 and 1963-70, as recorded in the 1990 Population Census on a linear year trend. Formally, estimate the equation,

$$\hat{P}_{cy} = \alpha_{0c} + \alpha_{1c} y + \epsilon_{cy}, \tag{2}$$

where α_{0c} and α_{1c} are respectively the constant and coefficient of the year trend for county c, and ϵ_{cy} is random error. We start the projection in 1949, the year in which the Communist China was founded, as China was racked by war – first, the Japanese invasion, and then, the civil war between the Nationalists and Communists in the previous one and half decades.

Then, we use equation (2) to project the population for all years from 1949 to 1970, including the pre-famine year, 1958, famine years, 1959-61, and the post-famine year, 1962. These five years are not used in the estimation. Figure 2 depicts the projected and actual populations for China as a whole. Evidently, the population in normal years fluctuated randomly around the time trend, and the difference between the counterfactual and actual populations increased sharply in the famine years. The famine was followed by a sharp baby boom in 1963. Figure 2 is quite similar to Meng et al.'s (2015) Figure 1(b) and Garnaut's (2014) Figure 2.

The major concern with the construct, μ_{cy}, is non-classical measurement error, that is, error that is correlated with the severity of the famine. Classical measurement error is less of a concern because it causes the estimated coefficients to be attenuated, and so, any statistical inference is conservative. Moreover, as explained below, we apply instrumental variables to estimate the effect of the famine on entrepreneurship. The instrumental variables resolve any classical measurement error.

Accordingly, our key concern is with error that is correlated with the severity of the famine. One such source of error is migration between counties. Despite vigorous government enforcement, people did try to move away from famine-stricken areas (Thaxton 2008: 162-170; Yang 2012: 50). However, the 1990 Census enumerates persons by their hukou registration, so, the impact of error due to migration is limited to people who were able to change their hukou registration. Hukou changes and migration in general were severely restricted before the early 1990s (Lin and Yang 2000; Meng and Qian 2009: footnote 23).

The linear projection in equation (2) might give rise to errors, which may or may not be classical. Visually, Figure 2 suggests that the linear projection fits the aggregate pattern well. Indeed, the average R^2 of the estimate is 0.490 (s.e. 0.269), indicating strong predictive power for a very simple regression with just one explanatory variable. To check robustness, we omit the year 1963 from the basis of projection and also apply two non-linear projections, one quadratic and the other exponential. Our results below are robust to these alternative projections.

As an additional check of robustness, we also measure the intensity of the famine by the gender ratio. Mothers under stress produce more adrenal androgens, which increase spontaneous abortion of male fetuses, tilting the gender ratio of births towards females (James 2015). The famine certainly increased stress on mothers. Based on the 1990 Census, the male/female gender ratio was 1.069 in the pre-famine years 1950-57, and fell to 1.054 in the famine years, 1959-61.[6]

We propose to measure the severity of the Great Famine by the rate of cohort loss, as computed from the 1990 Population Census. To validate this construct, we investigate its relation with predictors of agricultural productivity. Agronomic research shows that exposure to heat during the growing season affects crop yields in a nonlinear way. The temperature must exceed some threshold for plants to absorb heat. Yields increase with temperature above the threshold up to some ceiling of around 32 degrees Celsius, beyond which yields decrease (Ritchie and NeSmith 1991; Deschenes and Greenstone 2007).

Accordingly, we measure the potential agricultural productivity in each county and year by the sum of "degree-days", with a floor of 8 degrees Celsius and a ceiling of 32 degrees Celsius, between April 1 and September 30. For each county, c, and day, d, the degree-day is computed according to

$$H_{cd} = \begin{cases} 0 \;\text{ if }\; T_{cd} < 8, \\ T_{cd} - 8 \;\text{ if }\; 8 \leq T_{cd} < 32 \\ 24 \;\text{ if }\; T_{cd} \geq 32, \end{cases} \tag{3}$$

where T_{cd} represents the average temperature in the county on that day. Exposure to extreme temperatures harms plant growth. In robustness checks, we account for the harmful effects through an alternative measure of degree days. Following Richie and NeSmith (1991), the alternative measure is $T_{cd} - 8$ if the daily temperature is between 8 and 33 degrees Celsius, $\frac{25}{8}[41 - T_{cd}]$ if the daily average temperature is between 33 and 41 degrees Celsius, and zero otherwise.

Then, we define our second key construct, the *thermal agricultural productivity* in county,

[6]The gender ratio, as observed in the 1990 Census, measures the intensity of the famine with non-classical error. During famines, Chinese parents discriminated in favor of sons which raised their survival rate relative to daughters, and so, tilted the subsequent observed gender ratio towards males (Li 1991; Coale and Banister 1994; Mu and Zhang 2011).

c, and year, y, as the sum of the degree days during the growing season,

$$\text{Agri}_{cy} = \sum_{d=\text{April 1}}^{\text{Sept 30}} H_{cdy}. \tag{4}$$

This construct is relatively exogenous to manipulation. By contrast, direct measures of agricultural production may have been distorted by government officials during the Great Famine or incompletely adjusted by post-famine statisticians.[7] In the same spirit, Meng and Qian (2009) and Meng et al. (2015) also construct exogenous measures of agricultural productivity, which are based on soil characteristics and monthly weather.

Figure 3 illustrates the evolution of the average thermal agricultural productivity between 1951 and 1970 over all counties. Thermal agricultural productivity fluctuated and was only slightly below average in 1958-60, and actually exceeded the average in 1961. Table A1 in Appendix reports estimates which show that, after controlling for station fixed effects, our construct of thermal agricultural productivity is not serially correlated. This suggests that, at each station, thermal agricultural productivity is a random walk. Incidentally, our estimates refute the official explanation that the famine was caused by bad weather.

Having constructed the measure of thermal agricultural productivity, we estimate the following regression to validate the rate of cohort loss as a measure of the severity of the famine,

$$\mu_{cy} = \beta_0 + \beta_1 \text{Agri}_{c,y-1} + \sum_{y=1953}^{1970} \lambda_y \text{Agri}_{c,y-1} D_y + \sum_{y=1953}^{1970} \eta_y D_y + \nu_c + \varepsilon_{cy}, \tag{5}$$

where D_y is a year indicator, and β_0, β_1, λ_y, and η_y are the constant and coefficients to be estimated, ν_c are county fixed effects, and ε_{cy} is random error. This equation allows the effect of the previous year's thermal agricultural productivity to vary by year through the year-specific coefficients, λ_y. The equation also includes year-specific indicators with coefficients, η_y, to capture any variation in the construct of the severity of the famine not explained by

[7]During the Great Leap Forward, government chief statistician, Xue Muqiao, advised his staff, "Whatever statistical data the party leadership needs, we will provide it, and whatever direction the political campaigns and production campaigns take, statistical work will follow" (Xue (1958) quoted in Yang (2012: 257)).

variation in the previous year's thermal agricultural productivity.[8]

Table 2 reports and Figure 4 illustrates the estimates. The coefficient of the average previous year's thermal agricultural productivity, $\beta_1 = -0.043$ (s.e. 0.024), is negative and marginally significant, suggesting that, on average, a good harvest in the previous year was associated with fewer cohort losses in the subsequent year. This is consistent with part of surplus agricultural production (in excess of government procurement) accruing to people in the county.

Importantly, as Figure 4 clearly illustrates, the coefficient of the previous year's thermal agricultural productivity is positive and significantly larger in the famine years. Indeed, the coefficient is an order of magnitude larger in 1960, at the peak of the famine. The estimates imply that, during the famine years, *higher* agricultural productivity in the previous year was associated with *larger* cohort loss in the subsequent year. These results are consistent with previous research findings (Kung and Chen 2011; Meng et al. 2015) that the famine was due in part to officials over-estimating and over-reporting production, and, during the famine, being unable or unwilling to revise procurement or redistribute food in light of shortfalls in production.

Referring to Table 2, rightmost column, the coefficients of the year indicators, η_y, are positive and significant in the years 1958, 1959, 1961, and 1967, indicating that thermal agricultural productivity in the preceding year did not completely explain the variation in cohort loss. These years were exceptional: 1958 was the year preceding the Great Famine, 1959 and 1961 were famine years, and 1967 was the second year of the Cultural Revolution.

5 Empirical Strategy

Our thrust is to exploit county-level variation in the severity of the Great Famine to study the effect of hardship during the famine on subsequent entrepreneurship. However, statistical estimates of the relation might be confounded by unobserved differences between counties, challenging any causal inference. For instance, in a county with better natural resources, local

[8]The weather data covers 1951-70. Since equation (5) includes a lagged term, $\text{Agri}_{c,y-1}$, the first observation in the regression is for the year 1952.

officials might over-estimate production and increase procurement relatively more, causing more suffering during the famine. During normal times, people in the county with better natural resources might be more engaged in growing cash crops, which lays the foundation for future entrepreneurship. Furthermore, as already discussed, our construct for the severity of the famine, which is the rate of cohort loss, is subject to measurement error.

In light of possible endogeneity and measurement error, we employ two-stage least squares estimation, using two instruments for the severity of the famine. The instruments are the thermal agricultural productivity during famine years and the difference in thermal agricultural productivity between famine and normal years.

To explain and justify the instruments, reformulate equation (5) into two periods, $p = f$ (famine) and $p = n$ (normal),

$$\mu_{cp} = \beta_0 + \beta_1 \text{Agri}_{cp} + \eta D_p + \lambda \text{Agri}_{cp} D_p + \nu_c + \varepsilon_{cp}, \tag{6}$$

where μ_{cp} and Agri_{cp} are the average rate of cohort loss and thermal agricultural productivity in county c in period p.[9] The indicator, $D_p = 1$ for famine years, $y = 1959, 1960$, and 1961, and $D_p = 0$ otherwise.

Referring to equation (6), in normal years, $\mu_{cn} = \beta_0 + \beta_1 \text{Agri}_{cn} + \nu_c + \varepsilon_{cn}$. The coefficient β_1 measures the productivity effect of higher temperatures on the rate of cohort loss during normal years. Owing to localized administration or barriers to trade, greater production of food in a county raises county food consumption in the following year. Accordingly, we expect $\beta_1 < 0$.[10]

The economic interpretation of λ is subtle. Referring to equation (6), in famine years, $\mu_{cf} = \beta_0 + \beta_1 \text{Agri}_{cf} + \eta + \lambda \text{Agri}_{cf} + \nu_c + \varepsilon_{cf}$. Generally, higher temperatures affect cohort losses during the famine in two opposite ways. One way, as in normal years, is through productivity, by which higher temperatures raise food production, and so, reduce cohort

[9]Specifically, $\mu_{cf} = \frac{1}{3}\sum_{y=1959}^{1961} \mu_{cy}$, $\mu_{cn} = \frac{1}{15}(\sum_{y=1953}^{1958} \mu_{cy} + \sum_{y=1962}^{1970} \mu_{cy})$, $\text{Agri}_{cf} = \frac{1}{3}\sum_{y=1959}^{1961} \text{Agri}_{c,y-1}$, and $\text{Agri}_{cn} = \frac{1}{15}(\sum_{y=1953}^{1958} \text{Agri}_{c,y-1} + \sum_{y=1962}^{1970} \text{Agri}_{c,y-1})$.

[10]In principle, the negative coefficient, $\beta_1 < 0$, might also represent an income effect as greater production of food raises overall income. However, prior to the Great Famine, the Chinese government had forced the collectivization of agriculture, and so, rural people did not earn more by producing more.

losses, $\beta_1 < 0$. The other way is through institutions during the famine period. Higher food production was associated with higher grain procurement, which coupled with rigidity in procurement, led to *more* severe famine and greater cohort losses (Meng et al. 2015). Accordingly, we expect the institutional effect, $\lambda > 0$.[11]

We then define the *relative cohort loss rate* as the difference in the rate of cohort loss between famine and normal years. Specifically, $\Delta\mu_c \equiv \mu_{cf} - \mu_{cn} = \beta_1 \Delta \text{Agri}_c + \eta + \lambda \text{Agri}_{cf} + \varepsilon_{cf} - \varepsilon_{cn}$, which simplifies to

$$\Delta\mu_c = \eta + \lambda \text{Agri}_{cf} + \beta_1 \Delta \text{Agri}_c + \Delta\varepsilon_c, \tag{7}$$

where $\Delta \text{Agri}_c \equiv \text{Agri}_{cf} - \text{Agri}_{cn}$ and $\Delta\varepsilon_c \equiv \varepsilon_{cf} - \varepsilon_{cn}$. In our empirical analyses, we use the relative cohort loss rate to measure the county level severity of famine.

Figure 1 shows the geographic variation in the relative cohort loss rate, with darker colors representing more severe famine. Evidently, there is considerable variation in the severity of the famine across counties. The mean of the relative cohort loss rate is 0.363 (s.d. 0.211) with a minimum of −0.716 and a maximum of 1.039. The famine was most severe in the provinces of Sichuan, Anhui, Hunan, Guizhou, and Qinghai, and least severe in Heilongjiang, Inner Mongolia, and Tibet. There was also substantial variation within provinces. For instance, in Sichuan Province, the mean of the relative cohort loss rate is 0.592 (s.d. 0.151) with a minimum of -0.037 and a maximum of 0.789.

Table 3 presents estimates of equation (7) using the relative cohort loss rate as based on the 1990 Census, the 2000 Census, and the combination of both censuses.[12] The estimates are consistent with our prediction. Specifically, the productivity effect, β_1, is negative and the institutional effect, λ, is positive. In all three estimates, the coefficient of the thermal agricultural productivity during the famine years is positive and very precisely estimated, ranging between 0.075 (s.e. 0.013) and 0.079 (s.e. 0.013). These estimates imply that, if thermal agricultural productivity during the famine years had been higher by 0.623 (one standard deviation), then the relative cohort loss rate would have been higher by 0.047 and 0.049, after controlling for the differences in thermal agricultural productivity between

[11]Please refer to Section 2 for institutional details of the Great Famine.

[12]Below, we focus on estimates based on the 1990 Census, and use the 2000 Census and the combined censuses in robustness checks.

famine and normal periods. Relative to the average relative cohort loss rate, these differences represent the famine being 13.4 and 13.7 percent respectively more severe. The coefficient of the thermal agricultural productivity becomes smaller after accounting for spatial correlation in errors, but it is still precisely estimated.

Referring to Table 3, the coefficient of the difference in thermal agricultural productivity between famine and normal years is negative, ranging between -0.191 (s.e. 0.147) and -0.206 (s.e. 0.140). None of the coefficients is statistically significant. This suggests that the difference in thermal agricultural productivity is a weaker instrument for the severity of the famine than the thermal agricultural productivity during the famine years.

The above analysis motivates the cross-sectional empirical model, for county c,

$$Y_c = \omega_1 + \omega_2 \Delta\mu_c + \upsilon_c, \tag{8}$$

where Y_c measures entrepreneurship and υ_c is random error. Using the relative cohort loss rate, $\Delta\mu_c$, to represent the county-level severity of the famine helps in two ways. The relative cohort loss rate removes any county fixed effect, so using the relative cohort loss rate helps to abstract from non-time varying factors that affect entrepreneurship in a county and focus on the net effect of the famine on entrepreneurship. Formally, in equation (7), after controlling for Agri_{cf} and ΔAgri_c, the relative cohort loss rate, $\Delta\mu_c$, excludes the county fixed effect, ν_c.

Further, with two instruments for $\Delta\mu_c$, we can identify the causal effect of hardship on entrepreneurship by estimating equation (8) using two-stage least squares regression with equation (7) as the first stage. The two instruments are thermal agricultural productivity during the famine years and the difference in thermal agricultural productivity between famine and normal years.

Instruments must satisfy relevance and exclusion. The two instruments structurally enter the first stage, equation (7). As these two variables are correlated, the relevant diagnostic is the joint F-statistic, rather than individual t-statistics. The estimates in Table 3 show that the instruments satisfy the relevance condition (the joint F-statistic of the two instruments is about 20 across all specifications).

Although not directly testable, the instruments are likely to satisfy the exclusion condition. First, the instruments are the daily average temperatures in the growing season during the famine period and the difference in temperatures between the famine and normal periods in the years 1951-70. We study the effect of the famine on entrepreneurship and individual personality in the mid- to late 2000s. It seems implausible that temperatures in the 1950s and 1960s would directly affect entrepreneurship or personality over forty years later.

Second, the difference in thermal agricultural productivity between famine and normal periods abstracts from non-time varying differences in entrepreneurship by county. These could be due to differences in climate, as for instance, people living in harsher climates, more mountainous areas, or nearer bodies of water are more or less entrepreneurial. Moreover, Table A1 in the Appendix shows that the difference in temperatures between the famine and normal periods is random and not serially correlated, and so, the instruments should satisfy the exclusion condition. Further, we use the Hansen overidentification test to check the validity of the two instruments in combination. To the extent that these arguments are valid, the instruments satisfy the exclusion condition.

6 Estimates

Figure 5 illustrates our main result using the 2004 Economic Census. Counties which suffered more during the Great Famine were more entrepreneurial in subsequent years. The figure depicts the entrepreneurial share of all enterprises (the ratio of the number of private enterprises less than 10 years old and with fewer than 100 employees to the number of all enterprises) against the severity of the famine.

Following up, Table 4, column (a), presents an OLS estimate of equation (8) with the dependent variable being the entrepreneurial share of all enterprises. The coefficient of the relative cohort loss rate, which represents the severity of the famine, is positive and statistically significant.

Table 4, column (b), reports the first-stage regression of the relative cohort loss rate on the instruments. This estimate is similar to that reported in Table 3, column (a). The instruments are not weak (F-statistic is 20.29). Then, Table 4, column (c), reports the second-stage IV estimate of equation (8) with the dependent variable being the entrepreneurial share of all enterprises. The coefficient of the relative cohort loss rate, 0.658 (s.e. 0.119), is positive and statistically significant. It implies that, if the famine is more severe by 0.210 (one standard deviation), then the proportion of entrepreneurial enterprises would be higher by 0.138, which is 29.3 percent of the average.

Importantly, the IV estimate of the coefficient of the relative cohort loss rate is substantially larger than the OLS estimate. The Hausman test statistic is 30.66 ($p = 0.000$). The significant difference and Hausman test confirm that the severity of the famine is endogenous to entrepreneurship, so, biasing the OLS estimate downward. One potential reason is that counties which suffered more during the Great Famine are those under stricter government control, the control persisted to the 2000s, and so, stifled entrepreneurship. Another potential reason is classical measurement error, which results in the estimated coefficients being attenuated.

Table 4, column (d), reports an IV estimate of equation (8) with the dependent variable being the entrepreneurial share of enterprise sales. The coefficient of the relative cohort loss rate, 0.264 (s.e. 0.086), is positive and statistically significant. It implies that, if the famine is more severe by one standard deviation, then the share of entrepreneurial sales would be 5.54 percentage points higher, which is 31.15 percent of the average share of sales.

Table 4, column (e), reports an IV estimate of equation (8) with the dependent variable being the entrepreneurial share of enterprise profits. The coefficient of the relative cohort loss rate, 0.734 (s.e. 0.147), is positive and statistically significant. It implies that, if the famine is more severe by one standard deviation, then the share of entrepreneurial profits would be 15.41 percentage points higher, which is 29.30 percent of the average share of profits.

Next, Table 4, column (f), reports an IV estimate of equation (8) with the dependent variable being the median age of private enterprises. The coefficient of the relative cohort loss rate, -31.763 (s.e. 12.047), is negative and statistically significant. It implies that, if the famine is more severe by one standard deviation, then the median age of private enterprises would be 6.7 months lower, which is substantial relative to the median age of less than 4 years.

Table 4, column (g), reports IV estimates of equation (8) with the dependent variable being the entrepreneurial share of employment. The coefficient of the relative cohort loss rate, 0.300 (s.e. 0.075), is positive and statistically significant, and the implied effect on entrepreneurship is economically large. If the famine is more severe by one standard deviation, then the entrepreneurial share of employment would be 6.30 percentage points higher, which is 30.73 percent of the average. The Hansen J-statistic is insignificant in four of the five regressions (columns (c) and (e)-(g)), which validates our use of the two instruments.

The preceding analysis shows the effect of hardship on entrepreneurship at the county level. To probe more deeply, we draw on the 2005 Population Census to investigate at the individual level, limiting to persons who were born before 1962. Using the Census record of individual employment, we define an individual to be an entrepreneur in two ways: if he/she reported being the owner of a private enterprise or self-employed (broad definition) or only if an owner (narrow definition). Then, we estimate linear probability models of equation (8) in which the dependent variable is an indicator of being an entrepreneur.

Table 5, column (a), reports the OLS estimate of equation (8). The coefficient of the relative cohort loss rate, which represents the severity of the famine, is not statistically significant. Table 5, column (b), reports the first-stage regression of the relative cohort loss rate on the instruments. This estimate is similar to that reported in Table 3, column (a), with the coefficients being larger (0.094 versus 0.075) due to the difference in sample (the analysis of entrepreneurship at the individual level encompasses only 2,212 counties owing to the omission of counties that changed boundaries between 1990 and 2005). The instruments are not weak (F-statistic is 17.87).

Then, Table 5, column (c), reports the second-stage IV estimate of entrepreneurship. The coefficient of the relative cohort loss rate, 0.088 (s.e. 0.021), is positive and statistically significant, and implies that, if the famine is more severe by 0.219 (one standard deviation), then the probability of the individual being an entrepreneur would be higher by 1.93 percentage points, or 56.76 percent of the average, which is 3.4 percent.

Consistent with the county-level analysis (Table 4), the IV estimated coefficient of the relative cohort loss rate is substantially larger than the OLS estimate. The result reaffirms that the severity of the famine is endogenous to entrepreneurship, and the endogeneity biases the OLS estimate downward.

To check the robustness of the findings, Table 5, column (d), reports an IV estimate with entrepreneurship defined narrowly as owning a private enterprise. The results are similar to those using the broader definition of entrepreneurship, reported in Table 5, column (c). The coefficient of the relative cohort loss rate, 0.013 (s.e. 0.006), is positive and statistically significant, and implies that, if the famine is more severe by one standard deviation, then the probability of the individual being an entrepreneur would be higher by 0.28 percentage points, or 40.67 percent of the average, which is 0.70 percent.

Hardship may affect females differently from males (Elder et al. 1985). Further, men are generally more likely to engage in entrepreneurship than women (Schiller and Crewson 1997). We therefore check whether the famine affected men and women differently. Table 5, columns (e) and (f), report the IV estimates for the two different definitions of entrepreneurship. As expected, the coefficient of the interaction term between the relative cohort loss rate and male is positive and statistically significant. A one standard deviation increase in the severity of the famine is associated with the probability of entrepreneurship (owner or self-employed) being higher by 2.78 and 1.01 percentage points for males and females respectively, which are 56.76 and 53.02 percent of their respective averages.

It is also important to look at the effect of hardship on the intensive as well as the extensive margin of entrepreneurship. Table 5, columns (c)-(f), present estimates of engagement in entrepreneurship, which is the extensive margin. To study the effect on the intensive margin, which is income from entrepreneurship, we use data from the 2005 Population Census on the individual's monthly income. We observe the income of only those who become entrepreneurs. Hence, an estimate of the effect of hardship on income from entrepreneurship must account for selection into entrepreneurship as well as possible endogeneity of the measure of hardship.

Heckman (1979) suggests that observed samples of wages depend on household structure. Here, in the same spirit, we instrument for selection into the entrepreneur sample by the number of generations in the individual's household. Table 5, columns (g) and (h), report IV estimates applying a Heckman correction for selection into entrepreneurship, and, in the second stage, using just one instrument – the thermal agricultural productivity during the famine.[13] The coefficients of the relative cohort loss rate are positive, statistically significant,

[13]The IV estimate with two instruments – the thermal agricultural productivity during the famine and the difference in thermal agricultural productivity between famine and normal periods – is over-identified. Accordingly, we use only the former as an instrument and control for the latter.

and economically large. If the relative cohort loss rate is one standard deviation (0.193) higher, the income of entrepreneurs would be higher by 161.1 percent ($= \exp(4.973 \times 0.193) - 1$) on the broad definition of entrepreneur (owner or self-employed) and by 154.4 percent on the narrow definition (owner).

As we explain below, even more important is possible differences in the effect of the famine by the age of the person when they experienced the hardship. Table 6 presents IV estimates by three-year birth cohort. Figure 6 depicts the estimated coefficients of the relative cohort loss rate and the corresponding 95 percent confidence intervals. Apparently, the marginal effect of famine is larger among the younger people. The difference between the cohorts can be interpreted in two ways. One is that the famine disproportionately killed infants and young children (Li 1991: Tables 8 and 9; Coale and Banister 1994: Table 1), and so, the famine culled relatively more non-entrepreneurial people among the young. The other possible interpretation is that people are more easily conditioned and better able to adapt their personality (equivalently, acquire non-cognitive skills) at early ages (Caspi and Silva 1995; Caspi et al. 2003; Heckman 2007; Nave et al. 2010; Heckman et al. 2013), and so, those who underwent the famine at a younger age were relatively more conditioned towards entrepreneurship.

Apparent differences between the cohorts should be interpreted with caution because they might be confounded by age or cohort effects. We measure entrepreneurship by the 2005 Population Census, in which the various birth cohorts differed in age. For example, in 2005, people born in the 1941-43 cohort were 62-64 years old, whereas those born in the 1959-61 cohort were 44-46 years old. One important cohort difference is the Cultural Revolution, which affected the education of people born between 1948 and 1961. The comparison of the estimates of marginal effect across different birth cohorts is possibly contaminated by age or cohort effects. However, the elasticity of entrepreneurship with respect to the severity of the famine (represented by the relative cohort loss rate) is large, statistically significant, and similar across cohorts, suggesting that the famine affected entrepreneurship across all cohorts.

Overall, the analyses suggest that people who experienced more severe hardship during the Great Famine were subsequently more likely to engage in entrepreneurship. As reported in Table A4 of Appendix, these results are also robust to: (i) excluding 1963 from the years used to project the counterfactual population, (ii) projection of the counterfactual population by a quadratic or exponential growth model, (iii) measuring the intensity of the famine by the gender ratio rather than cohort losses, (iv) measuring thermal agricultural productivity by a four-piece linear formula (Richie and NeSmith 1991), (v) estimating by probit regression rather than the linear probability model, (vi) accounting for spatial correlation in the errors, (vii) clustering standard errors by province, prefecture, or country, and (viii) including controls for ethnicity, gender, age, urbanization, and education.[14]

6.1 Selective Culling or Operant Conditioning

How did greater hardship during the Great Famine induce more entrepreneurship? One possible mechanism is Darwinian – the famine selectively culled the fragile, leaving survivors who were more suited to the challenging environment. Another possibility is that the famine conditioned people to adapt their personality in ways that helped them to cope and survive.

The famine might have selectively culled and/or conditioned people on several dimensions of personality that support entrepreneurship – risk tolerance, self-confidence, tenacity, and resilience. Individuals who are more tolerant of risk (Djankov et al. 2006; Hall and Woodward 2010; Hvide and Panos 2014), self-confident (self-efficacious) (Koellinger et al. 2007; Landier and Thesmar 2009; Hayward et al. 2010), tenacious (Baum and Locke 2004; Markman et al. 2005), or resilient (Bullough et al. 2014) are more likely to engage in entrepreneurship.[15]

It is fairly intuitive that the famine may have selectively eliminated people who were more risk averse, less confident, less tenacious, or less resilient. During the Great Famine, strategies for survival include speculation and profiteering, dealing in the black market, and stealing food (Zhou 2012: Chapter 7). Villagers in Da Fo in northern Henan Province survived by eating raw or unripe crops in the fields, deliberately leaving crops in the fields during the communal harvest for later "gleaning", slacking during communal work, black-market dealing in salt earth and nitrates, and begging (Thaxton 2008: Chapters 5 and 6).

[14]Gender, urbanization, and education are possibly endogenous, and so, the estimates including these variables should be interpreted with caution.

[15]Bernardo and Welch (2001) show theoretically that over-confidence among entrepreneurs may be socially optimal. In an environment with asymmetric information, over-confident people help to break (rational) herding among the uninformed.

Separately, research in psychology suggests that exposure to trauma or adversity may lead to adaptive changes in personality. In particular, exposure to trauma might condition tolerance for risk by reducing the individual's discount rate, decreasing the salience of future losses, regulation of emotions, and impairing higher-order cognitive function (Ben-Zur and Zeidner 2009). Struggle with hardship or cumulative experience of adversity may condition an individual to grow in resilience (Almedom 2005; Seery, Holman, and Silver 2010; Jayawickreme and Blackie 2014).

To investigate the mechanism underlying the differences in entrepreneurship across counties, suppose that the differences are due solely to selective mortality and that the famine did not increase entrepreneurship through any other mechanism. Denote the share of entrepreneurs in the population absent famine by e_c, the actual number of entrepreneurs by $\hat{E}_c$, and the share of entrepreneur deaths in cohort loss due to the famine by λ_c.

Accordingly, the actual number of entrepreneurs in county c,

$$\hat{E}_c = [e_c - \lambda_c \mu_c]\tilde{P}_c = [1 - \mu_c]\tilde{P}_c \cdot \frac{e_c - \lambda_c \mu_c}{1 - \mu_c}, \tag{9}$$

where, as defined around equation (1) above, μ_c is the rate of cohort loss, and $\tilde{P}_c$ the counterfactual projected population in the absence of famine. On the right-hand side of equation (9), $[1 - \mu_c]\tilde{P}_c$ is the actual population as observed during the census, and $[e_c - \lambda_c \mu_c]/[1 - \mu_c]$ is the share of entrepreneurs in the actual population.

Suppose that the famine did not kill all entrepreneurs, $e_c - \lambda_c \mu_c > 0$, so that $\hat{E}_c > 0$. Then, differentiating the logarithm of equation (9) with respect to μ_c, we have, conditional on $\tilde{P}_c$,

$$\frac{\partial \ln \hat{E}_c}{\partial \mu_c} = \frac{-\lambda_c}{e_c - \lambda_c \mu_c} \leq 0, \tag{10}$$

which states that the number of entrepreneurs should not increase with the severity of the famine.

Now, consider estimating the model,

$$\ln \hat{E}_c = \gamma_0 + \gamma_1 \mu_c + \gamma_2 \ln \tilde{P}_c + \epsilon_c, \tag{11}$$

where ϵ_c is an error term. By equation (10), if the famine increased entrepreneurship only through selective culling, $\gamma_1 \leq 0$. However, if, empirically, we find $\gamma_1 > 0$, then some other mechanism must have raised entrepreneurship so much as to outweigh the selective culling.

The challenges in estimating equation (11) are that the rate of cohort loss, μ_c, and the counterfactual population absent the famine, $\tilde{P}_c$, are possibly endogenous. Accordingly, we operationalize μ_c by the relative cohort loss rate, $\Delta\mu_c$, and estimate equation (11) using two instruments – the thermal agricultural productivity during the famine period and the difference in thermal agricultural productivity between famine and normal periods – for the two endogenous variables, μ_c and $\ln \tilde{P}_c$.

As discussed above, famines disproportionately kill the young but people are more easily conditioned at an early age. Hence, we are particularly interested in differences among cohorts in the extent to which the increase in entrepreneurship was not due to selective culling. To broaden the analysis to include people who were teenagers during the famine, we extend the projection of cohort losses (equation (2) above) to include the years 1947 and 1948.

Table 7, columns (a)-(e), report IV estimates of equation (11) by three-year birth cohorts, specifically, 1947-49, 1950-52, 1953-55, 1956-58, and 1959-61, with entrepreneurs defined as owners of private enterprises or self-employed. To ensure comparability across estimates, the sample is limited to counties with a positive number of entrepreneurs in all cohorts.[16]

The coefficient of the relative cohort loss rate is consistently positive across all birth cohorts, and is statistically significant in the three older cohorts (1947-49, 1950-52, and 1953-55), but imprecisely estimated in the two younger cohorts (1956-58 and 1959-61) ($p < 0.4$). The relation between hardship and entrepreneurship is monotone with birth cohort and strongest in the oldest cohort (1947-49), or those who were 10-12 years old at the onset of the famine. In that cohort, if the relative cohort loss rate had been higher by one standard deviation, the number of entrepreneurs would have been 56.4 percent higher.[17]

Referring to equation (11), the reduced-form estimate of γ_1 captures both selective culling and operant conditioning. By equation (10), the selective culling effect is negative. Hence,

[16] If $\hat{E}_c = 0$, the famine did not increase entrepreneurship, so it does not make sense to analyze whether the increase was due to selective culling or conditioning.

[17] The Appendix presents similar results with entrepreneurship defined narrowly as owning a private enterprise.

the positive estimates of γ_1 reported in Table 7 imply that the increase in entrepreneurship goes beyond selective culling. In particular, the estimate of γ_1 provides a lower bound on the positive effect of hardship on entrepreneurship through operant conditioning.

Accordingly, the pattern of the estimates by birth cohort in Table 7 suggest significant operant conditioning among individuals who were seven years and older at the onset of the famine (1950-52 and earlier cohorts), with weaker evidence of conditioning among those aged four to six years (1953-55 cohort). Young children may have observed parents, siblings, other relatives and neighbors engaging in risky actions such as taking raw crops from fields and going out at night to glean food. Older children might have directly participated in gathering food. Children learned behavior and developed personal traits that increased their chances of survival and which stimulated entrepreneurship in later life.

It might seem interesting to compare the patterns of the estimates by birth cohort in Tables 6 and 7. However, the two sets of estimates are not directly comparable (Table 6 reports individual level analyses on a larger set of counties, while Table 7 reports county level analyses, controlling for cohort size, on a subset of the counties). Subject to that proviso, comparing the patterns by birth cohort in Tables 6 and 7, we infer that operant conditioning was stronger among the older children, while selective mortality was stronger among the very young and particularly those born during the famine.

In interpreting the differences between the cohorts, we caution that the differences across birth cohorts are possibly contaminated by age or cohort effects. Further, Table 7 presents the net outcome of two possibly conflicting effects. On the one hand, we argue that famine conditions personality traits that increase entrepreneurship. On the other hand, famine diminishes cognitive skills (Meng and Qian 2009; Ampaabeng and Tan 2013), which would reduce entrepreneurship. Accordingly, the estimates in Table 7 are biased downward in gauging the effect of operant conditioning of personality traits.

Development of both cognitive and non-cognitive skills (personality traits) is strongest in early childhood (Heckman et al. 2013). The empirical pattern of weaker effects on entrepreneurship in later cohorts is consistent with the famine having affected cognitive development relatively more than non-cognitive development among young children.

6.2 Conditioning and Adaptation

As reviewed above, previous research points to several dimensions of personality as being important ingredients in entrepreneurship – risk tolerance, confidence (self-efficacy), tenacity, and resilience. Through which of these personality traits did the Great Famine stimulate entrepreneurship? To investigate, we use individual data from various surveys to estimate the relation between severity of the famine and each of these traits.

A conventional measure of risk tolerance is investment in shares (equities). Using data on 400,000 Norwegian individuals, Hvide and Panos (2014) validate investment in shares as a measure of risk tolerance and show that it is correlated with starting a new business.[18]

The CFPS 2010, CHFS 2011, and CGSS 2010, 2012, and 2013 include questions on investments.[19] We combine responses to these surveys, limiting to individuals born before 1962. Table 8, column (a), reports the first-stage regression of the relative cohort loss rate on the thermal agricultural productivity during the famine and the difference in thermal agricultural productivity between famine and normal periods. The instruments are not weak ($F = 13.96$).

Table 8, column (b), reports the IV estimate of a linear probability model with the dependent variable being an indicator of investment in the share market. The coefficient of the relative cohort loss rate, 0.091 (s.e. 0.045), is positive and statistically significant. The result indicates that people who suffered more during the famine are more likely to invest in shares, which suggests that they are more tolerant of risk. This finding is buttressed by the IV estimate, reported in Table 8, column (c), of financial investment, in which the coefficient of the relative cohort loss rate is positive and significant, 0.127 (s.e. 0.062).[20]

Further, Table 8, column (d), reports an IV estimate of ownership of a private enterprise based on these survey data. The coefficient of the relative cohort loss rate is positive and statistically significant. People who suffered more during the famine are also more likely to

[18]In a Netherlands sample of 2,288 entrepreneurs and managers, Koudstaal et al. (2015) find that the entrepreneurs are more loss averse but not more risk averse than the managers.

[19]CGSS 2008 and 2011 do not include questions on investments.

[20]It might seem intuitive to control for income and wealth in these estimates, but to the extent that our finding that the famine caused people to adapt towards more risk-taking, that would raise their incomes and wealth. Accordingly, income and wealth would be bad controls in a regression of risk tolerance on the severity of the famine.

own a private enterprise. This estimate provides further evidence in support of our main finding of a positive relation between hardship and entrepreneurship.[21]

The three other personality traits that promote entrepreneurship – self-confidence, tenacity, and resilience – are related. An individual who is more confident in her/his own abilities and more tenacious will be better able to bounce back from setbacks, and so, more resilient (Hayward et al. 2010). Bearing in mind this proviso, we investigate whether the Great Famine selectively culled or conditioned people on these other traits.

Self-efficacy, a psychological concept which we interpret as self-confidence, has been related to entrepreneurship (Koellinger et al. 2007; Landier and Thesmar 2009; Hayward et al. 2010; Bullough et al. 2014). The CGSS 2011 includes five questions that gauge self-confidence. As described in the Appendix, we recode all measures in the positive direction, so that a larger value represents a higher degree of self-confidence.

Table 9 reports IV regressions of self-confidence on the severity of the famine, as represented by the relative cohort loss rate. The coefficients of the relative cohort loss rate for self-assessed power and ability to overcome difficulties are positive. However, perhaps due to small sample size, the coefficients are not statistically significant. The coefficients of the relative cohort loss rate are negative in other three measures of self-confidence, which is not consistent with hardship inducing greater confidence. Accordingly, we infer that the data do not support an inference that the famine induced greater self-confidence.[22]

Another personality trait that promotes entrepreneurship is determination, persistence, or tenacity (Baum and Locke 2004; Markman et al. 2005). The CGSS 2008 and 2011 include two questions that elicit this trait. One asks the respondent the degree to which he/she agrees with the statement, "I will exert effort to finish the task even when not feeling well", and the other asks the degree to which he/she agrees with the statement, "I will exert effort to finish the task even it takes a long time to finish".

[21]Although investment is a binary variable, we use linear regression to facilitate estimation and interpretation. As shown in the Appendix, Table A5, the findings are robust to estimation by a probit model.

[22]Although self-confidence, tenacity, resilience are measured on integer Likert scales, we use linear regression to facilitate estimation and interpretation. As shown in the Appendix, Tables A6 and A7, the findings are robust to estimation by ordered probit model.

Table 10 presents IV estimates of tenacity. The coefficients of the relative cohort loss rate are negative, which is not consistent with hardship raising tenacity. We infer that the data do not support an inference that the famine induced greater tenacity.

Yet another personality trait that promotes entrepreneurship is resilience (Bullough et al. 2014). Did hardship during the Great Famine underpin later entrepreneurship by selecting more resilient people or conditioning people to be more resilient? To gauge resilience, we exploit Chairman Mao Zedong's policy during the Cultural Revolution to move graduates of junior and senior high school from urban areas to live in the countryside. In a double randomization, we study the contingent effect of the famine and having been "sent-down" on physical health and tenacity.[23]

The CGSS 2008 includes information on whether the respondent reported being sent down during the Cultural Revolution, and the respondent's health and tenacity. Table 11, columns (a)-(c), report IV estimates of the respondent's self-reported health. In Table 11, column (c), the coefficient of the interaction between the relative cohort loss rate and having been sent down, -11.196 (s.e. 9.772), is negative, which is not consistent with a more severe famine increasing resilience.

Table 11, columns (d)-(f), report IV estimates of the respondent's tenacity, as represented by the degree of agreement with statement, "I will exert effort to finish the task even when not feeling well".[24] In Table 11, column (f), the coefficient of the interaction between the relative cohort loss rate and having been sent down, 3.328 (s.e. 3.689) is positive, which is consistent with people experiencing more severe famine becoming more resilient. However, perhaps owing to the small sample, the coefficient is not precisely estimated. Accordingly, we infer that the data do not support an inference that the famine induced greater resilience.

Risk tolerance, self-confidence, tenacity, and resilience are positive traits. Besides these, particularly in societies with weak institutions, entrepreneurship in China is also associated with opportunism, craftiness, and skirting authority. An old Chinese saying is *wu shang bu jian, wu jian bu shang*, which means "no businessman is not crafty, without craftiness there

[23]Gong et al. (2015) exploit the "send-down" movement to analyze the effect of adversity during adolescent years on non-cognitive skills.

[24]To address the endogeneity of being sent-down, we use the birth cohort and hukou registration type at birth as instrumental variables.

is no business". It is plausible that crafty or opportunistic people were more likely to survive the famine, or the famine conditioned people to become more crafty or opportunistic, which led to more entrepreneurship in later life.

The CGSS 2008 includes six questions to elicit respondents' views on the use of *guanxi* (personal connections) to gain advantage. We use these to gauge the individual's degree of opportunism or craftiness. As described in the Appendix, we apply principal components analysis to the six questions and code the principal component such that a higher score represents stronger agreement with the use of *guanxi*. Table 12, column (a), reports an IV estimate of the respondent's agreement with using *guanxi* to gain advantage. The coefficient of the relative cohort loss rate is negative, which is not consistent with more opportunistic or crafty people being more likely to survive the famine or famine conditioning people to become more opportunistic or crafty.

Another test of opportunism or craftiness uses the CFPS 2010, which asks respondents the degree to which they agree that "To achieve success in today's society, it is impossible to avoid bribery" on a 5-point Likert scale with higher score representing stronger agreement. Table 12, column (b), reports the IV estimate of the respondent's agreement with using bribery. The coefficient of the relative cohort loss rate, 0.223 (s.e. 0.463), is positive, which is consistent with people experiencing more severe famine being more opportunistic or crafty. However, perhaps due to the small sample, the coefficient is imprecisely estimated.

Yet another trait through which the famine might have stimulated entrepreneurship is skirting or outright disobedience to authority. Absent better data, we investigate this aspect through the number of police in the county with control for the population. Counties whose residents are less law-abiding might employ more police. We use information on occupation in the 2000 Population Census to enumerate the police. Table 12, column (c), reports a county-level IV estimate of the logarithm of the number of police officers. The coefficient of the relative cohort loss rate, -5.494 (s.e. 1.684), is negative and statistically significant. Apparently, counties that experienced more severe famine had fewer police, which suggests that their residents are more law-abiding. This empirical finding is inconsistent with the famine inducing people to skirt or disobey authority.

Overall, we conclude that the empirical evidence favors the explanation that the Great

Famine increased entrepreneurship by selecting on or conditioning tolerance for risk, rather than self-confidence, tenacity, resilience, or craftiness or opportunism. However, we should qualify that the inference depends on self-confidence, tenacity, and resilience being measured by self assessment and based on small samples.

6.3 Other Explanations

The Great Famine killed male fetuses at a higher rate than females, and so, raising the ratio of girls to boys (Mu and Zhang 2011). When these children reach the age of marriage, the imbalance in gender would challenge the girls to get attractive marriage partners. This might induce parents of girls to work harder and take more risk, and specifically, engage in entrepreneurship to earn more, and so to better compete to win marriage partners for their daughters (Wei and Zhang 2011).

To check, we investigate whether the famine induces people with daughters to engage in entrepreneurship relatively more. The 2005 Population Census asks only women for the gender of their children, and so, our analysis is limited to women. Table 13, column (a), reports an IV estimate of entrepreneurship (defined as owner or being self-employed) among women. The coefficient of the relative cohort loss rate, 0.071 (s.e. 0.019), is slightly smaller than the corresponding coefficient for men and women, 0.088 (s.e. 0.021) (reported in Table 5, column (c)), which is consistent with women being less likely to engage in entrepreneurship.

Table 13, column (b), reports an IV estimate of entrepreneurship among women, distinguishing those with daughters. The coefficient of having a daughter is negative and significant, suggesting that women with daughters were less likely to engage in entrepreneurship. This is consistent with the Chinese practice of the bridegroom's family bearing the expenses of marriage. The coefficient of the relative cohort loss rate interacted with having a daughter is positive, 0.022 (s.e. 0.018), which is consistent with women with daughters and who suffered relatively more during the famine being more likely to engage in entrepreneurship. However, the estimate is not statistically significant.

The 2005 Population Census does not report the children of men, and so, we cannot carry out the corresponding analysis for men. Intuitively, given Chinese men's strong preference for sons (Tan et al. 2014), we expect any relation among men with daughters to be weaker than the relation among women with daughters. The estimate of the effect of famine on women with daughters is imprecise. Accordingly, we are inclined to rule out the hypothesis that the famine induced people to engage in entrepreneurship so as to earn more to attract spouses for their daughters.

Another alternative explanation emphasizes the effect of the famine on old-age support. The Great Famine killed many children and reduced fertility, even leaving some women permanently unable to bear children. Traditionally, Chinese people look to their sons for support in old age. Without any sons, a couple must plan to provide for themselves in later years. One way is to work harder and take more risk and engage in entrepreneurship to build up savings and even to continue economic activity after the normal retirement age.

Table 13, column (c), reports an IV estimate of entrepreneurship among women, distinguishing those with no sons. The coefficient of having no son is positive and significant, consistent with the traditional notion that people without sons must worker harder and take more risks to support themselves in old age. The coefficient of the relative cohort loss rate interacted with not having any son is negative, -0.013 (s.e. 0.025), which is not consistent with women with no sons and who suffered relatively more during the famine being more likely to engage in entrepreneurship.

Another possible explanation of the relation between the severity of the famine and entrepreneurship is selection by health. The famine may have culled weaker people, leaving a relatively healthier population, which is better able to cope with the vagaries of entrepreneurship. Table 11, column (a), presents direct evidence against the culling hypothesis. In a regression of self-reported health, the coefficient of the relative cohort loss rate is not significant.

Yet another possible explanation of the relation between the severity of the famine and entrepreneurship is the role of government, which might be positive or negative. One possibility is that the government gave more assistance to the areas that suffered most during the famine, and, this support fostered more entrepreneurship in subsequent years. Another possibility is that the differences in the intensity of the famine reflected differences in state capacity (Lu et al. 2016). In areas with greater state capacity, officials were better able to enforce procurement, and so, caused more suffering. Alternatively, in areas with weaker state capacity, officials were less able to mobilize production, and so, caused more suffering. Differences in state capacity would affect the opportunity for private-sector entrepreneurship, depending on whether state action is complement or substitute for entrepreneurship. For instance, Jia and Lan (2014) show the value of political connections to entrepreneurs in China.

To investigate, we use the density of roads (kilometers of roads per square kilometer area) in the year 2000 to represent government assistance or state capacity in the county. Since this might be endogeneous to entrepreneurship, we instrument for road density by the average gradient of the terrain. The gradient would affect the cost of building roads but not directly affect entrepreneurship.

Accordingly, we extract the length of roads in each county in the year 2000 from the China Data Center at the University of Michigan, and compute the density as the ratio of the length of roads in kilometers to the area of the county in square kilometers. We obtain information on terrain from the China Historical Geographic Information System (CHGIS), Harvard Yenching Institute, and calculate the average gradient as the difference in altitude between the highest and lowest points in the county divided by the area of the county.

We first regress, at county level, road density on the severity of the famine, as represented by the relative cohort loss rate. As Table 13, column (d) reports, the coefficient of the relative cohort loss rate, -1.777 (s.e. 0.425), is negative and statistically significant. This suggests that, in counties where the famine was more severe, the government in recent times has been less effective and state capacity is lower.

Next, we regress entrepreneurship on the relative cohort loss rate and road density. Referring to Table 13, column (e), the coefficient of road density is not statistically or economically significant. The coefficient of the relative cohort loss rate, 0.080 (s.e. 0.019) is quite similar to the estimate without controlling for road density (Table 5, column (c)).

Finally, we add the interaction between the relative cohort loss rate and road density, with road density specified relative to the sample average. As Table 13, column (f), reports, the coefficient of the relative cohort loss rate, 0.056 (s.e. 0.018), is positive and statistically significant, and the coefficient of road density,-0.006 (s.e. 0.022), is negative and insignificant.

Both coefficients are much smaller than the estimates without controlling for the interaction between the relative cohort loss rate and road density (Table 13, column (e)). The coefficient of the relative cohort loss rate interacted with road density, -0.004 (s.e. 0.062), is negative but imprecisely estimated. Nearly zero coefficient of the relative cohort loss rate and road density seems to suggest that hardship and state capacity are relatively independent in stimulating entrepreneurship.

7 Discussion

Exploiting geographical variation in the severity of China's Great Famine, we find robust evidence that sustained hardship is associated with more entrepreneurship, particularly among those who experienced the famine at early ages. The relation is partly due to the famine selectively culling non-entrepreneurial people. However, among those aged seven or older at the time, the relation between hardship and entrepreneurship is at least partly the result of conditioning towards becoming more entrepreneurial. We find evidence consistent with conditioning of risk tolerance, but no evidence of conditioning of confidence, tenacity, resilience, opportunism, or craftiness.

We should qualify that our analysis is cross-sectional, not a life cycle study. Even though we analyze the population by birth cohorts separately, the data do not allow us to follow the same persons over time to investigate how the famine affected their development at various stages of life. Our findings only address the long term effect of the famine on conditioning of personality and entrepreneurship.

Another limitation is that the famine imposed hardship across the board rather than on particular individuals. Everyone in the county suffered the same deprivation, except to the extent that they could dodge the authorities and get more food. Our research context does not allow us to distinguish the direct personal effects of hardship from peer effects due to the hardship borne by family members and friends.

Our research does contribute to research on whether entrepreneurs are born or nurtured (Lindquist et al. 2015), and more particularly, whether entrepreneurs can be trained. We show that people can be conditioned to become more entrepreneurial. In particular, we investigate the effect of hardship, a factor much emphasized by entrepreneurs, and show that hardship cultivates both more entrepreneurship and greater risk tolerance. This finding suggests that entrepreneurship education and training should aim to cultivate higher risk tolerance. It supports experiential learning – where students learn by facing and overcoming challenges in unstructured situations.

While we find that hardship conditioned individuals aged between seven and twelve to become more entrepreneurial, our research design is limited in drawing inferences about the conditioning at earlier ages (six years old or younger). The effect for the younger children is also positive, albeit imprecise. The challenge with regard to younger children is that we only observe the net effect of the famine on entrepreneurship through cognitive and non-cognitive development. Research on child development (Heckman et al. 2013) shows that development of both cognitive and non-cognitive skills is fastest at early ages. Our inconclusive findings on the effect of hardship on entrepreneurship at early ages may be due to the conflict between negative effects on cognitive skills and positive effects on non-cognitive skills. An important direction for future research is to distinguish these two effects of hardship on entrepreneurship.

Another direction for future research is the oft-cited view that society should encourage risk-taking and entrepreneurship by reducing the stigma of failure (see, for instance, European Commission 2011). During China's Great Famine, people who failed paid the ultimate price with their own lives. Even parents and children, brothers and sisters turned against each other (Yang 2012). There was certainly no support for failure. Yet, under such dire conditions, people developed personality traits that fostered entrepreneurship in later years. This means that policies and programs to reduce the stigma associated with failure can facilitate the development of entrepreneurship.

References

Almedom, Astier M, "Resilience, Hardiness, Sense of Coherence, and Posttraumatic Growth: All Paths Leading to "Light at the End of the Tunnel"?," *Journal of Loss and Trauma*, 2005, *10* (3), 253–265.

Ampaabeng, Samuel K and Chih Ming Tan, "The Long-term Cognitive Consequences of Early Childhood Malnutrition: The Case of Famine in Ghana," *Journal of Health Economics*, 2013, *32* (6), 1013–1027.

Baum, Robert J. and Edwin A. Locke, "The Relationship of Entrepreneurial Traits, Skill, and Motivation to Subsequent Venture Growth," *Journal of Applied Psychology*, 2004, *89* (4), 587–598.

Becker, Jasper, *Hungry Ghosts: Mao's Secret Famine*, New York: The Free Press, 1996.

Ben-Zur, Hasida and Moshe Zeidner, "Threat to Life and Risk-Taking Behaviors: A Review of Empirical Findings and Explanatory Models," *Personality and Social Psychology Review*, 2009, *13* (2), 109–128.

Bernardo, Antonio E. and Ivo Welch, "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics & Management Strategy*, 2001, *10* (3), 301–330.

Bernile, Gennaro, Vineet Bhagwat, and P Raghavendra Rau, "What Doesn't Kill You Will Only Make You More Risk-Loving: Early-Life Disasters and CEO Behavior," *Journal of Finance, Forthcoming*, 2015.

Bullough, Amanda, Maija Renko, and Tamara Myatt, "Danger Zone Entrepreneurs: The Importance of Resilience and Self-Efficacy for Entrepreneurial Intentions," *Entrepreneurship Theory and Practice*, 2014, *38* (3).

Cao, Shuji, *The Great Famine: China's Population in 1959-61* (in Chinese), Hong Kong: Time International Publishing, 2005.

Caspi, Avshalom, HonaLee Harrington, Barry Milne, James W. Amell, Reremoana F. Theodore, and Terrie E. Moffitt, "Children's Behavioral Styles at Age 3 Are Linked to Their Adult Personality Traits at Age 26," *Journal of Personality*, 2003, *71* (4), 495–514.

Caspi, Avshalom and Phil A. Silva, "Temperamental Qualities at Age Three Predict Personality Traits in Young Adulthood: Longitudinal Evidence from a Birth Cohort," *Child Development*, 1995, *66* (2), 486–498.

Coale, Ansley J and Judith Banister, "Five Decades of Missing Females in China," *Demography*, August 1994, *31* (3), 459–479.

Commission, European, "A Second Chance for Entrepreneurs: Prevention of Bankruptcy, Simplification of Bankruptcy Procedures and Support for a Fresh Start," 1 2011.

Deschenes, Olivier and Michael Greenstone, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *American Economic Review*, 2007, *97* (1), 354–385.

Dikotter, Frank, *Mao's Great Famine: The History of China's Most Devastating Catastrophe, 1958-1962*, London: Bloomsbury Publishing, 2010.

Djankov, Simeon, Yingyi Qian, Gérard Roland, and Ekaterina Zhuravskaya, "Entrepreneurship in China and Russia Compared," *Journal of the European Economic Association*, 2006, *4* (2-3), 352–365.

Elder, Glen H, Tri Van Nguyen, and Avshalom Caspi, "Linking Family Hardship to Children's Lives," *Child Development*, 1985, *56* (2), 361–375.

Garnaut, Anthony, "The Geography of the Great Leap Famine," *Modern China*, May 2014, *40* (3), 315–348.

Gong, Jie, Yi Lu, and Huihua Xie, "Rusticated Youth: The Send-Down Movement and Beliefs," 2014.

Greer, Stephen, "Does an Entrepreneur Need an MBA?," Harvard Business Review Blog 11 2010.

Hall, Robert E. and Susan E. Woodward, "The Burden of the Nondiversifiable Risk of Entrepreneurship," *American Economic Review*, 2010, *100* (3), 1163–1194.

Hayward, Mathew, William Forster, Saras Sarasvathy, and Barbara Fredrickson, "Beyond Hubris: How Highly Confident Entrepreneurs Rebound to Venture again," *Journal of Business Venturing*, 2010, *25* (6), 569–578.

Heckman, James J, "Sample Selection Bias as a Specification Error," *Econometrica: Journal of the Econometric Society*, 1979, pp. 153–161.

Heckman, James J, "The Economics, Technology, and Neuroscience of Human Capability Formation," *Proceedings of the National Academy of Sciences of the United States of America*, 2007, *104* (33), 13250–13255.

Heckman, James, Rodrigo Pinto, and Peter Savelyev, "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, October 2013, *103* (6), 2052–86.

Hvide, Hans K. and Georgios A. Panos, "Risk Tolerance and Entrepreneurship," *Journal of Financial Economics*, 2014, *111* (1), 200–223.

James, William H, "Proximate Causes of the Variation of the Human Sex Ratio at Birth," *Early Human Development*, 2015, *91* (12), 795–799.

Jayawickreme, Eranda and Laura ER Blackie, "Post-traumatic Growth as Positive Personality Change: Evidence, Controversies and Future Directions," *European Journal of Personality*, 2014, *28* (4), 312–331.

Jia, Ruixue and Xiaohuan Lan, "Capitalism for the Children: Entrepreneurs with Cadre Parents Under Big Government," University of California San Diego December 2014.

Koellinger, Philipp, Maria Minniti, and Christian Schade, ""I think I can, I think I can": Overconfidence and Entrepreneurial Behavior," *Journal of Economic Psychology*, 2007, *28* (4), 502–527.

Koudstaal, Martin, Randolph Sloof, and Mirjam Van Praag, "Risk, Uncertainty, and Entrepreneurship: Evidence from a Lab-in-the-Field Experiment," *Management Science*, 2015.

Kung, James Kai-sing and Shuo Chen, "The Tragedy of the Nomenklatura: Career Incentives and Political Radicalism during China's Great Leap Famine," *American Political Science Review*, 2011, *105* (1), 27–45.

Landier, Augustin and David Thesmar, "Financial Contracting with Optimistic Entrepreneurs," *Review of Financial Studies*, 2009, *22* (1), 117–150.

Lavely, William, "First Impressions from the 2000 Census of China," *Population and Development Review*, 2001, *27* (4), 755–769.

Li, Lillian M, "Life and death in a Chinese famine: infanticide as a demographic consequence of the 1935 Yellow River flood," *Comparative Studies in Society and History*, 1991, *33* (03), 466–510.

Li, Wei and Dennis Tao Yang, "The Great Leap Forward: Anatomy of a Central Planning Disaster," *Journal of Political Economy*, 2005, *113* (4), 840–877.

Lin, Chen, Yue Ma, Micah S. Officer, and Hong Zou, "CEOs' Military Experience, Agency Costs and Acquisition Decisions," SSRN 2345151 2011.

Lin, Justin Yifu and Dennis Tao Yang, "Food Availability, Entitlements and the Chinese Famine of 1959-61," *Economic Journal*, January 2000, *110* (460), 136–158.

Lindquist, Matthew J, Joeri Sol, and Mirjam Van Praag, "Why do Entrepreneurial Parents Have Entrepreneurial Children?," *Journal of Labor Economics*, 2015, *33* (2), 269–296.

Lu, Yi, Mona Luan, and Tuan-Hwee Sng, "The Effect of State Capacity Under Different Economic Systems," SSRN 2728605 January 2016.

Malmendier, Ulrike, Geoffrey Tate, and Jon Yan, "Overconfidence and Early-life Experiences: The Effect of Managerial Traits on Corporate Financial Policies," *Journal of Finance*, 2011, *66* (5), 1687–1733.

Markman, Gideon D., Robert A. Baron, and David B. Balkin, "Are Perseverance and Self-Efficacy Costless? Assessing Entrepreneurs' Regretful Thinking," *Journal of Organizational Behavior*, 2005, *26* (1), 1–19.

Meng, Xin and Nancy Qian, "The Long Term Consequences of Famine on Survivors: Evidence from a Unique Natural Experiment Using China's Great Famine," NBER 14917 April 2009.

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

Mu, Ren and Xiaobo Zhang, "Why does the Great Chinese Famine Affect the Male and Female Survivors Differently? Mortality Selection versus Son Preference," *Economics & Human Biology*, 2011, *9* (1), 92–105.

Nave, Christopher S., Ryne A. Sherman, David C. Funder, Sarah E. Hampson, and Lewis R. Goldberg, "On the Contextual Independence of Personality: Teachers' Assessments Predict Directly Observed Behavior after Four Decades," *Social Psychological and Personality Science*, 2010, *1* (4), 327–334.

Nietzsche, Friedrich W., *Twilight of the Idols or How to Philosophize with a Hammer*, Harmondsworth: Penguin Books, 1990. Duncan Large, translator.

Rauch, Andreas and Michael Frese, "Let's Put the Person Back into Entrepreneurship Research: A Meta-analysis on the Relationship between Business Owners' Personality Traits, Business Creation, and Success," *European Journal of Work and Organizational Psychology*, 2007, *16* (4), 353–385.

Ravallion, Martin, "Famines and Economics," *Journal of Economic Literature*, 1997, *35* (3), 1205–1242.

Ritchie, J.T. and D.S. NeSmith, "Temperature and Crop Development," in John Hanks and J.T. Ritchie, eds., *Modeling Plant and Soil Systems*, Crop Science Society of America, Inc., 1991, chapter 8, pp. 5–29.

Schiller, Bradley R. and Philip E. Crewson, "Entrepreneurial Origins: A Longitudinal Inquiry," *Economic Inquiry*, 1997, *35* (3), 523–531.

Schlenker, Wolfram and Michael J. Roberts, "Nonlinear temperature effects indicate severe damages to US crop yields under climate change," *Proceedings of the National Academy of sciences*, 2009, *106* (37), 15594–15598.

Seery, Mark D., E. Alison Holman, and Roxane Cohen Silver, "Whatever Does Not Kill Us: Cumulative Lifetime Adversity, Vulnerability, and Resilience," *Journal of Personality and Social Psychology*, 2010, *99* (6), 1025–1041.

Shane, Scott and Sankaran Venkataraman, "The promise of entrepreneurship as a field of research," *Academy of Management Review*, January 2000, *25* (1), 217–226.

Tan, Chih Ming, Zhibo Tan, and Xiaobo Zhang, "Sins of the Fathers: The Intergenerational Legacy of the 1959-61 Great Chinese Famine on Children's Cognitive Development," IFPRI 01351 2014.

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

Wei, Shangjin and Xiaobo Zhang, "Sex Ratios, Entrepreneurship, and Economic Growth in the People's Republic of China," NBER 16800 2011.

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

Yao, Shujie, "A Note on the Causal Factors of China's Famine in 1959–1961," *Journal of Political Economy*, 1999, *107* (6), 1365–1369.

Zhang, Guangyu and Zhongwei Zhao, "Reexamining China's Fertility Puzzle: Data Collection and Quality over the Last Two Decades," *Population and Development Review*, 2006, *32* (2), 293–321.

Zhou, Xun, *The Great Famine in China, 1958-1962: A Documentary History*, New Haven, CT: Yale University Press, 2012.

Zhou, Xun, *Forgotten Voices of Mao's Great Famine, 1958-1962*, New Haven, CT: Yale University Press, 2013.

Table 1. Summary statistics

Panel (a): Economic Census 2004

VARIABLES	Unit	Obs.	Mean	Std. dev.
Share of entrepreneurial firms	Ratio	2,194	0.471	0.16
Share of sales in entrepreneurial firms	Ratio	2,194	0.178	0.14
Share of profits in entrepreneurial firms	Ratio	2,194	0.526	0.19
Median age of entrepreneurial firms	Month	2,194	37.708	12.289
Median age of private enterprises	Month	2,194	47.448	18.126
Share of employees in entrepreneurial firms	Ratio	2,194	0.205	0.125
Share of high school educated employees	Ratio	2,194	0.17	0.11
Relative cohort loss rate		2194	0.366	0.21
Thermal agricultural productivity during famine	1,000 degrees	2,194	2.41	0.623
Difference in thermal agricultural productivity between famine and normal years	1,000 degrees	2,194	-0.009	0.057

Panel (b): Population Census 2005

VARIABLES	Unit	Obs.	Mean	Std. dev.
2,212 counties included in the analysis				
Owner or self-employed	Indicator	639,443	0.034	0.18
Owner of private enterprise	Indicator	639,443	0.007	0.081
Owner of private enterprise: Monthly income	Yuan	4,221	1,669	2,586
Self-employed	Indicator	639443	0.027	0.162
Self-employed: Monthly income	Yuan	17,105	795.9	690.5
Males: Owner of private enterprise	Indicator	322,933	0.011	0.103
Males: Self-employed	Indicator	322,933	0.038	0.19
Females: Owner of private enterprise	Indicator	316,510	0.003	0.051
Females: Self-employed	Indicator	316,510	0.016	0.125
Relative cohort loss rate		639,443	0.349	0.219
Thermal agricultural productivity during famine years	1,000 degrees	639,443	2.527	0.593
Difference in thermal agricultural productivity between famine and normal years	1,000 degrees	639,443	-0.008	0.054
612 counties excluded owing to boundary change				
Owner or self employed	Indicator	145,062	0.037	0.19
Owner of private enterprise	Indicator	145,062	0.008	0.087
Self-employed	Indicator	145,062	0.030	0.17

Panel (c): Combined survey sample

VARIABLES	Unit	Obs.	Mean	Std. dev.
Owner of private enterprise	Indicator	21,965	0.021	0.143
Any financial investment	Indicator	21,965	0.043	0.203
Stock market participation	Indicator	21,965	0.027	0.161
Relative cohort loss rate		21,965	0.343	0.212
Counties		248		

Notes: Sample: Respondents to CFPS2010, CHFS2011, CGSS2010, CGSS2012, or CGSS2013 born before 1962.

Table 2. Thermal agricultural productivity and cohort loss rate

VARIABLES	β_1	λ_y	η_y
Thermal agricultural productivity in previous year	-0.043*		
	(0.024)		
Year 1953		0.015	-0.024
		(0.018)	(0.042)
Year 1954		-0.007	-0.026
		(0.020)	(0.047)
Year 1955		0.029	-0.073
		(0.020)	(0.049)
Year 1956		-0.002	0.088*
		(0.021)	(0.050)
Year 1957		-0.022	0.083
		(0.021)	(0.050)
Year 1958		0.014	0.127**
		(0.021)	(0.050)
Year 1959		0.044**	0.250***
		(0.021)	(0.051)
Year 1960		0.127***	0.037
		(0.022)	(0.054)
Year 1961		0.040**	0.374***
		(0.020)	(0.047)
Year 1962		-0.014	0.044
		(0.020)	(0.048)
Year 1963		-0.039*	-0.095*
		(0.020)	(0.049)
Year 1964		0.016	-0.032
		(0.018)	(0.044)
Year 1965		-0.006	0.041
		(0.018)	(0.043)
Year 1966		-0.005	0.049
		(0.018)	(0.042)
Year 1967		0.000	0.162***
		(0.017)	(0.041)
Year 1968		0.003	-0.027
		(0.017)	(0.041)
Year 1969		0.018	0.050
		(0.018)	(0.042)
Year 1970		0.014	0.025
		(0.017)	(0.041)

Notes: This table reports the OLS estimate of equation (5); Dependent variable: Cohort loss rate based on the 1990 census; with county fixed effects; 33,327 observations of 2,326 counties, covering years 1952-1970; R-squared: 0.366; F statistic: 31.79; Column (a): Coefficient of previous year's thermal agricultural productivity, β_1; Column (b): Coefficient of previous year's thermal agricultural productivity interacted with year, λ_y; Column (c): Coefficient of year indicator, η_y; Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 3. Thermal agricultural productivity and relative cohort loss rate

VARIABLES	(a) Based on Census 1990	(b) Based on Census 2000	(c) Based on Censuses 1990 and 2000	(d) Based on Census 1990: Spatial correlation in errors
Thermal agricultural productivity during famine, λ	0.075***	0.077***	0.079***	0.032***
	(0.013)	(0.013)	(0.013)	(0.009)
Difference in thermal agricultural productivity between famine and normal periods, β_1	-0.191	-0.192	-0.206	-0.105
	(0.147)	(0.140)	(0.140)	(0.081)
Observations	2,265	2,589	2,200	2,265
R-squared	0.058	0.062	0.069	
F statistic	21.71	21.89	22.02	16.45
Prob > F	< 0.001	< 0.001	< 0.001	<0.001

Notes: Sample: All counties; Estimated by ordinary least squares; Dependent variable: Relative cohort loss rate; Robust standard errors clustered by station in parentheses (* p<0.1; ** p<0.05; *** p<0.01). Column (a): Relative cohort loss rate estimated using the 1990 Population Census; Column (b): Relative cohort loss rate estimated using the 2000 Population Census; Column (c): Relative cohort loss rate estimated using the 1990 and 2000 Censuses; Column (d): Relative cohort loss rate estimated using the 1990 Population Census, allowing for spatial correlation in errors (estimated by Stata routine, spreg).

Table 4. Famine severity and entrepreneurship: County analysis

VARIABLES	(a) OLS: Entrepreneurial share of enterprises	(b) First stage	(c) IV: Entrepreneurial share of enterprises	(d) IV: Entrepreneurial Share of enterprise sales	(e) IV: Entrepreneurial share of enterprise profits	(f) IV: Median age of private enterprises	(g) IV: Entrepreneurial share of employees
Relative cohort loss rate	0.161***		0.658***	0.264***	0.734***	-31.763***	0.300***
	(0.021)		(0.119)	(0.086)	(0.147)	(12.047)	(0.075)
Thermal agricultural productivity during famine		0.075*** (0.013)					
Difference in thermal agricultural productivity (famine – normal)		-0.191 (0.149)					
Counties	2,194	2,194	2,194	2,194	2,194	2,194	2,194
R-squared	0.04	0.06					
Hansen J statistic	.	.	0.56	10.85	0.39	1.96	1.08
p-value	.	.	0.45	0.00	0.53	0.16	0.30

Notes: Sample: Enterprises covered by the 2004 Economic Census; Across columns (c) – (g), instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Columns (a) and (c): Dependent variable is entrepreneurial firms' share in all enterprises; Column (d): Dependent variable is entrepreneurial firms' share of all enterprise sales; Column (e): Dependent variable is entrepreneurial firms' share of all enterprise profits; Column (f): Dependent variable is median age in month of private enterprises; Column (g): Dependent variable is entrepreneurial firms' share of enterprise employment. Columns (c)-(g): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$), and first-stage F statistic = 20.29. Robust standard errors in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 5. Famine severity and entrepreneurship: Individual analysis

VARIABLES	(a) OLS: Owner or self-employed	(b) First Stage	(c) IV: Owner or self-employed	(d) IV: Owner	(e) IV: Owner or self-employed	(f) IV: Owner	(g) Selection & IV: Income of owner or self-employed	(h) Selection & IV: Income of owner
Relative cohort loss rate	-0.003		0.088***	0.013**	0.046***	-0.001	4.973***	4.839**
	(0.005)		(0.021)	(0.006)	(0.014)	(0.002)	(0.963)	(2.082)
Thermal agricultural productivity		0.094***						
during famine		(0.017)						
Difference in thermal agricultural		-0.180					2.022***	2.850***
productivity (famine – normal)		(0.169)					(0.374)	(0.963)
Male					-0.000	-0.002		
					(0.007)	(0.003)		
Relative cohort loss					0.081***	0.027***		
rate x male					(0.020)	(0.009)		
Observations	639,443	639,443	639,443	639,443	639,443	639,443	21,326	4,221
Counties	2,212	2,212	2,212	2,212	2,212	2,212	2,025	1,298
R-squared	0.000	0.060						
Hansen J statistic			1.300	1.890	6.966	2.148		
p-value			0.254	0.169	0.0307	0.342		

Notes: Sample: Local residents born before 1962 covered by the 2005 Population Census; Columns (c)-(f) present estimates by two-stage least squares weighted by population, with instruments being thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Columns (g)-(h) present estimates by Heckman selection with number of generations in household as instrument and two-stage least squares with thermal agricultural productivity in famine period as instrument and difference in thermal agricultural productivity between famine and normal years as control. Columns (a), (c), (e): Dependent variable is indicator of owning private enterprise or self-employed; Columns (d) and (f): Dependent variable is indicator of owning private enterprise; Column (g): Sample limited to owners of private enterprises or self-employed, dependent variable is logarithm of monthly income; Column (h): Sample limited to owners of private enterprises, dependent variable is logarithm of monthly income; Columns (c)-(d): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$), and first-stage F statistic = 17.87. Robust standard errors clustered by station in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 6. Famine severity and individual entrepreneurship: IV estimates by birth cohort

VARIABLES	(a) Before 1940	(b) 1941-43	(c) 1944-46	(d) 1947-49	(e) 1950-52	(f) 1953-55	(g) 1956-58	(h) 1959-61
Relative cohort loss rate	0.026***	0.058***	0.097***	0.121***	0.116***	0.160***	0.149***	0.236***
	(0.006)	(0.015)	(0.024)	(0.027)	(0.029)	(0.035)	(0.038)	(0.051)
Observations	174,440	42,583	49,555	60,936	73,683	87,886	86,651	63,709
Counties	2,207	2,192	2,193	2,207	2,207	2,208	2,208	2,203
F statistic	15.09	19.65	17.36	20.06	23.12	19.76	19.17	17.33
Hansen J statistic	0.002	0.944	0.279	2.047	0.927	1.112	2.143	1.280
p-value	0.962	0.331	0.597	0.153	0.336	0.292	0.143	0.258
Elasticity	1.577	1.347	1.608	1.559	1.053	1.142	0.822	1.021
p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Notes: Sample: Local residents born before 1962 covered by the 2005 Population Census; Dependent variable is indicator of owning private enterprise or self-employed; Estimated by two-stage least squares weighted by population; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Elasticity is calculated as the coefficient of relative cohort loss rate multiplied by the sample mean of relative cohort loss rate and divided by the sample mean of the dependent variable; Columns (c)-(d): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$); F statistic is for first stage. Robust standard errors clustered by station in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 7. Increase in entrepreneurship beyond selective mortality: IV estimates by birth cohort

VARIABLES	(a) 1947-49	(b) 1950-52	(c) 1953-55	(d) 1956-58	(e) 1959-61
Relative cohort loss rate	2.993***	2.558**	2.784*	1.698	1.796
	(1.048)	(1.240)	(1.506)	(1.862)	(2.003)
Population (ln)	0.501	0.395	0.369	0.862	0.499
	(0.613)	(0.653)	(0.813)	(0.950)	(0.987)
Counties	477	477	477	477	477
Elasticity	1.059	0.905	0.985	0.601	0.635
p-value	0.004	0.039	0.065	0.362	0.370

Notes: Sample: Local residents born before 1962 covered by the 2005 Population Census; each column limited to counties with positive number of entrepreneurs (owners of private enterprises or self-employed) across all three-year birth cohorts between 1947-59 and 1959-61; Estimated by two-stage least squares weighted by population; Dependent variable is the logarithm of the number of entrepreneurs; Population is that of the respective birth cohort; Instruments are thermal agricultural productivity in famine period and the difference in thermal agricultural productivity between famine and normal periods; Elasticity is calculated as the coefficient of relative cohort loss rate multiplied by the mean of relative cohort loss rate; Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.05$), and F statistic = 12.87 and 4.16 for first stage regressions of relative cohort loss rate and logarithm of population respectively; Robust standard errors clustered by station in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 8. Risk tolerance: Individual analysis

VARIABLES	(a) First stage	(b) IV: Share investment	(c) IV: Financial investment	(d) IV: Owner
Relative cohort loss rate		0.091**	0.127**	0.050**
		(0.045)	(0.062)	(0.025)
Thermal agricultural productivity during famine	0.123***			
	(0.024)			
Difference in thermal agricultural productivity (famine - normal)	0.069			
	(0.260)			
Survey fixed effects	Yes	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes	Yes
Observations	21,965	21,965	21,965	21,965
R-squared	0.090			
Counties	248	248	248	248
Hansen J statistic		0.292	0.489	1.226
p-value		0.589	0.484	0.268

Note: Sample: Respondents to CFPS 2010, CHFS 2011, CGSS 2010, CGSS 2012, or CGSS 2013 born before 1962; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Column (b): Dependent variable is indicator of investment in share market; Column (c): Dependent variable is indicator of financial investment; Column (d): Dependent variable is indicator of owning private enterprise; Columns (c)-(e): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$), and first-stage F statistic = 13.96.; Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 9. Self-confidence: Individual analysis – IV estimates

VARIABLES	(a) Power	(b) Confidence	(c) Overcome difficulties	(d) Control own life	(e) Confidence
Relative cohort loss rate	1.136	-0.295	0.113	-0.143	-0.200
	(1.211)	(0.528)	(0.586)	(0.446)	(0.602)
Observations	1,671	1,628	1,652	1,683	1,680
Counties	47	47	47	47	47
F statistic	3.85	3.83	3.86	3.82	3.83
Hansen J statistic	2.177	0.977	0.933	2.577	0.124
p-value	0.140	0.323	0.334	0.108	0.724

Notes: Sample: Respondents to CGSS 2011 born before 1962; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Columns (a): Dependent variable is self-assessed power on 10-point Likert scale with higher score representing greater power; Column (b): Dependent variable is response to question, “Have you ever lost confidence in the past four weeks?”on 5-point Likert scale with lower score representing stronger agreement; Column (c) Dependent variable is response to question, “Have you encountered any difficulties that you cannot overcome in the past four weeks?”on 5-point Likert scale with lower score representing stronger agreement; Column (d): Dependent variable is agreement with statement, “I cannot control my life”on 5-point Likert scale with lower score representing stronger agreement; Column (e): Dependent variable is agreement with statement, “I am confident that I can handle problems in my life”.on 5-point Likert scale with higher score representing stronger agreement; Columns (a)-(e): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.05$); F statistic is for first stage. Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 10. Tenacity: Individual analysis – IV estimates

VARIABLES	(a) Even if not well	(b) Even if long time
Relative cohort loss rate	-0.483	-0.346
	(0.308)	(0.267)
Wave fixed effects	Yes	Yes
Observations	3,343	3,289
Counties	116	116
F statistic	8.79	8.72
Hansen J statistic	0.017	0.001
p-value	0.895	0.979

Notes: Sample: Respondents to CGSS 2008 or CGSS 2011 born before 1962; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Columns (a): Dependent variable is agreement with statement, "I will exert effort to finish the task even when not feeling well", on 4-point Likert scale with higher score representing stronger agreement; Column (b): Dependent variable is agreement with statement, "I will exert effort to finish the task even it takes a long time to finish", on 4-point Likert scale with higher score representing stronger agreement; Columns (a)-(b): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$); F statistic is for first stage. Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 11. Resilience: Individual analysis – IV estimates

VARIABLES	(a) Health	(b) Health	(c) Health	(d) Tenacity	(e) Tenacity	(f) Tenacity
Relative cohort loss rate	0.787		1.634**	0.207		0.131
	(0.613)		(0.664)	(0.303)		(0.295)
Send-down		0.774	3.519		0.423	-0.071
		(0.815)	(2.358)		(0.512)	(0.815)
Relative cohort loss rate			-11.196			3.328
x send-down			(9.772)			(3.689)
Observations	1,544	1,544	1,544	1,530	1,530	1,530
Counties	73	73	73	73	73	73
F statistic	5.35	28.65	5.96	5.29	29.24	5.91
Hansen J statistic	1.985		2.509	1.347		1.860
p-value	0.159		0.285	0.246		0.394

Notes: Sample: Respondents to CGSS 2008 born before 1962; Estimated by two stage least squares; Instruments for relative cohort loss rate are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Instrument for send-down is birth cohort times registration type; "Send-down" as reported by survey respondent; Columns (a)-(c): Dependent variable is self-reported health condition on 5-point Likert scale with higher score representing better health; Columns (d)-(f): Dependent variable is agreement with statement, "I will exert effort to finish the task even when not feeling well", on 4-point Likert scale with higher score representing stronger agreement; Columns (a)-(f): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.05$); F statistic is for first stage regression of relative cohort loss rate. Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Table 12. Craftiness, opportunism, obedience to authority:
Individual analysis – IV estimates

VARIABLES	(a) Use of *guanxi*	(b) Use of bribery	(c) Police (ln)
Relative cohort loss rate	-0.468	0.223	-5.494***
	(1.102)	(0.463)	(1.684)
Population (ln)			1.663***
			(0.324)
Observations	1,506	5,634	2,180
Counties	73	83	2,180
F statistic	6.12	7.25	21.06
Hansen J statistic	2.696	3.644	
p-value	0.101	0.056	

Notes: Estimated by two-stage least squares with instruments being thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal periods; Column (a): Sample comprises respondents to CGSS 2008 born before 1962; Dependent variable: The principal component of agreement with six statements regarding use of *guanxi* to gain advantage, where higher score represents stronger agreement; Column (b): Sample comprises respondents to CFPS 2010 born before 1962; Dependent variable: Agreement with using bribery to achieve success on 5-point Likert scale with higher score representing stronger agreement; Column (c): Sample comprises counties with any person reporting police as occupation in 2000 Population Census; Dependent variable is logarithm of number of police; Relative cohort loss rate and population specified as endogenous; Columns (a)-(c): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.05$); F statistic is for first stage. Robust s.e. clustered by weather station (* p<0.1; ** p<0.05; *** p<0.01).

Table 13. Alternative explanations: Children, State capacity

VARIABLES	(a) IV: Entrepre-neurship	(b) IV: Entrepre-neurship	(c) IV: Entrepre-neurship	(d) OLS: Road density	(e) IV: Entrepre-neurship	(f) IV: Entrepre-neurship
Relative cohort	0.071***	0.057**	0.076***	-1.777***	0.080***	0.056***
loss rate	(0.019)	(0.023)	(0.019)	(0.425)	(0.019)	(0.018)
Daughter		-0.018***				
		(0.006)				
Relative cohort loss		0.022				
rate x daughter		(0.018)				
No son			0.016*			
			(0.008)			
Relative cohort loss			-0.013			
rate x no son			(0.025)			
Gradient				-0.359***		
				(0.056)		
Road density					-0.011	-0.006
					(0.007)	(0.022)
Relative cohort loss						-0.004
rate x road density						(0.062)
Observations	228,535	228,535	228,535	2,171	629,805	629,805
Counties	2,211	2,211	2,211	2,171	2,171	2,171
F statistic	19.39	14.96	20.31	37.85	18.70	14.34
Hansen J statistic	0.026	0.102	0.131	8.131	0.953	12.57
p-value	0.872	0.950	0.937	0.004	0.329	0.002

Notes: All columns except (d): Estimated by two-stage least squares with instruments being thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal periods; Dependent variable is indicator of owning private enterprise or self-employed. Columns (a)-(c): Sample: Local female residents born after 1940 and before 1962 covered by 2005 Population Census; Column (b): Additional instrument -- thermal agricultural productivity in famine period interacted with daughter; Column (c): Additional instrument -- thermal agricultural productivity in famine period interacted with no son; Column (d): Estimated by ordinary least squares; Dependent variable is road density; Columns (e)-(f): Sample: Local residents born before 1962 covered by 2005 Population Census; Column (e): Additional instrument – county average gradient; Column (f): Additional instruments – county average gradient and thermal agricultural productivity in famine period interacted with gradient; Columns (a) and (e): Kleibergen-Paap rk LM statistics reject under-identification ($p < 0.001$); F statistic is for first stage regression of relative cohort loss rate. Robust standard errors clustered by station in parentheses (* $p<0.1$; ** $p<0.05$; *** $p<0.01$).

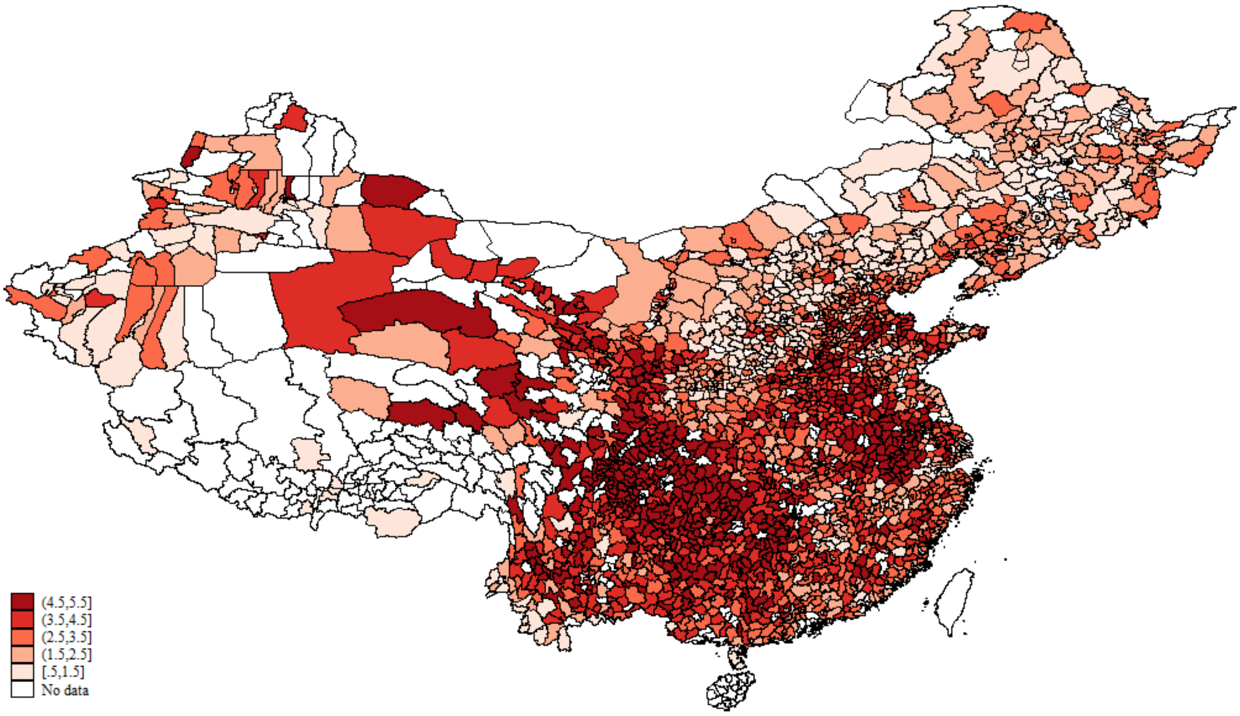


Figure 1. Famine severity: Variation by county

Notes: Quintiles of severity of famine as represented by the county relative cohort loss rate (difference between famine and normal periods) based on the 1990 Population Census; Darker color represents more severe famine.

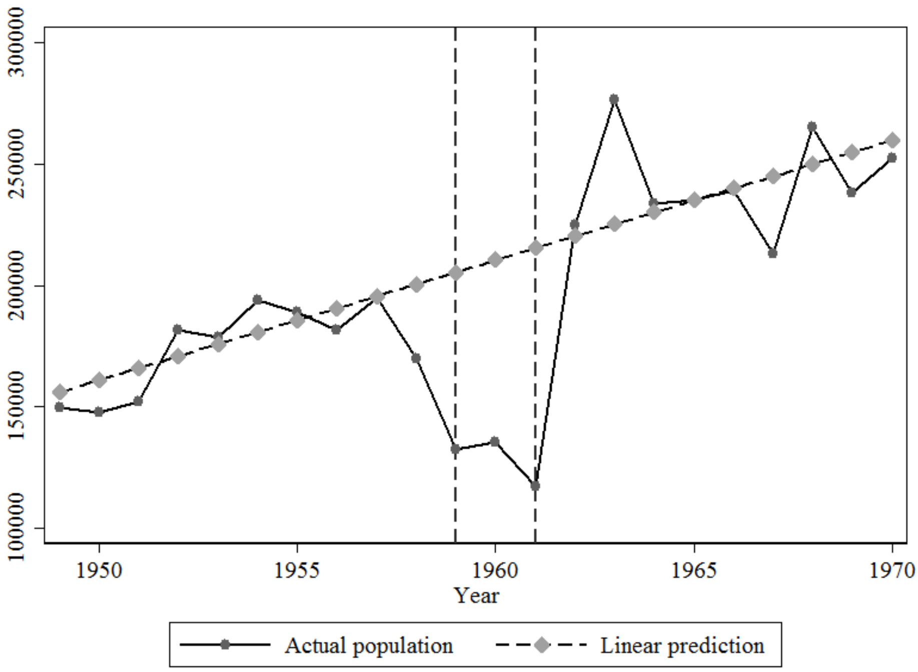


Figure 2. Population by birth cohort: Projected vis-a-vis observed

Note: Projected based on years 1949-57 and 1963-70 using the 1990 Population Census.

Figure 3. Thermal agricultural productivity

Notes: Graph depicts thermal agricultural productivity, or average sum of degree days between April 1 and September 30 (equation (4)) over all counties by year; the horizontal line depicts average thermal agricultural productivity over the period 1951-70.

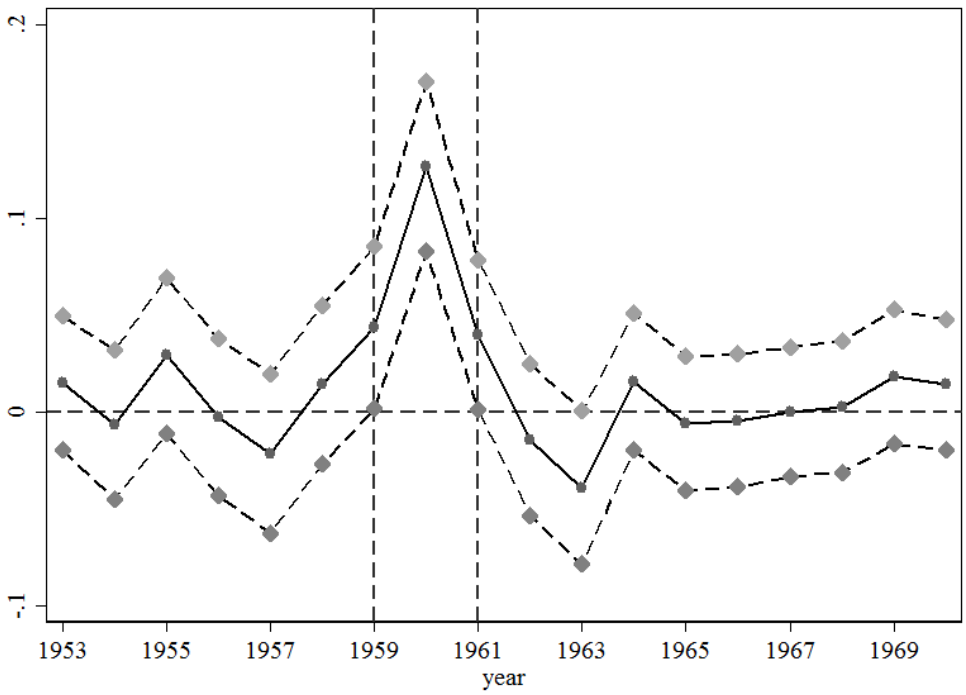


Figure 4. Cohort loss rate and thermal agricultural productivity

Notes: Graph depicts the estimated year-specific coefficients from regression of rate of cohort loss on previous year's thermal agricultural productivity (Table 2, column (b)) and 95% confidence intervals.

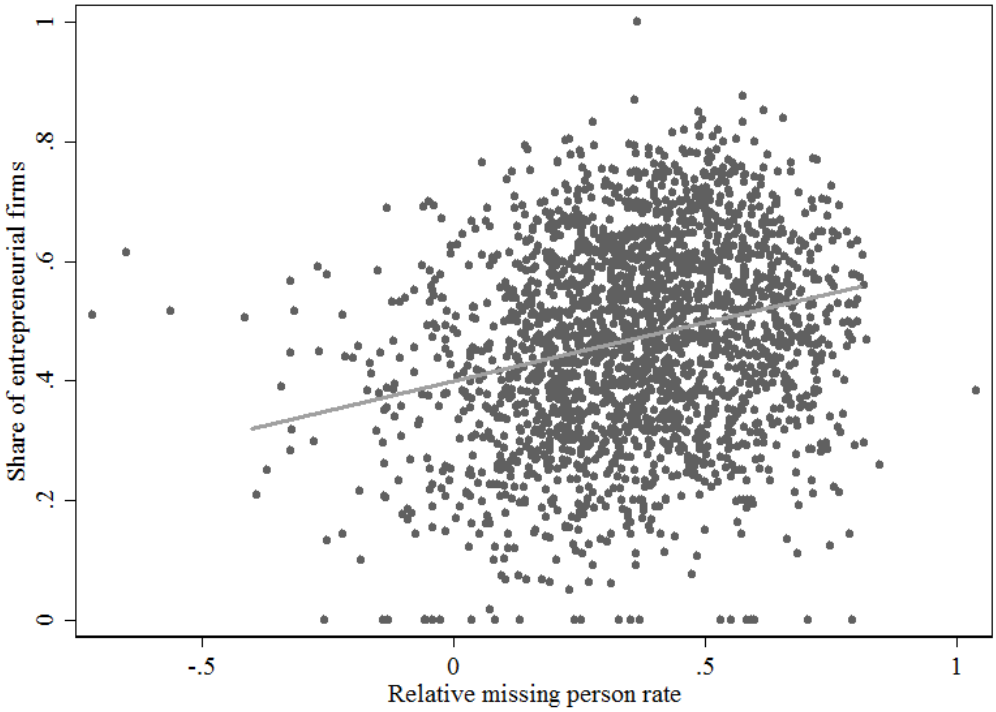


Figure 5. Famine severity and entrepreneurship
Notes: Based on 2004 Economic Census; Ratio of number of private enterprises with fewer than 100 employees and less than 10 years old to number of all enterprises; Each dot represents one county; the line depicts the linearly fitted values; slope coefficient 0.147 (s.e. 0.020).

Figure 6. Famine severity and entrepreneurship: Individual analysis by cohort
Notes: Figure depicts coefficients from Table 6, IV estimates of individual entrepreneurship (owner of private enterprise or self-employed) on relative cohort loss rate by birth cohort, and 95 percent confidence intervals.

Entrepreneurship and the School of Hard Knocks: Evidence from China's Great Famine Appendices

Appendix A. Famine severity and agricultural productivity

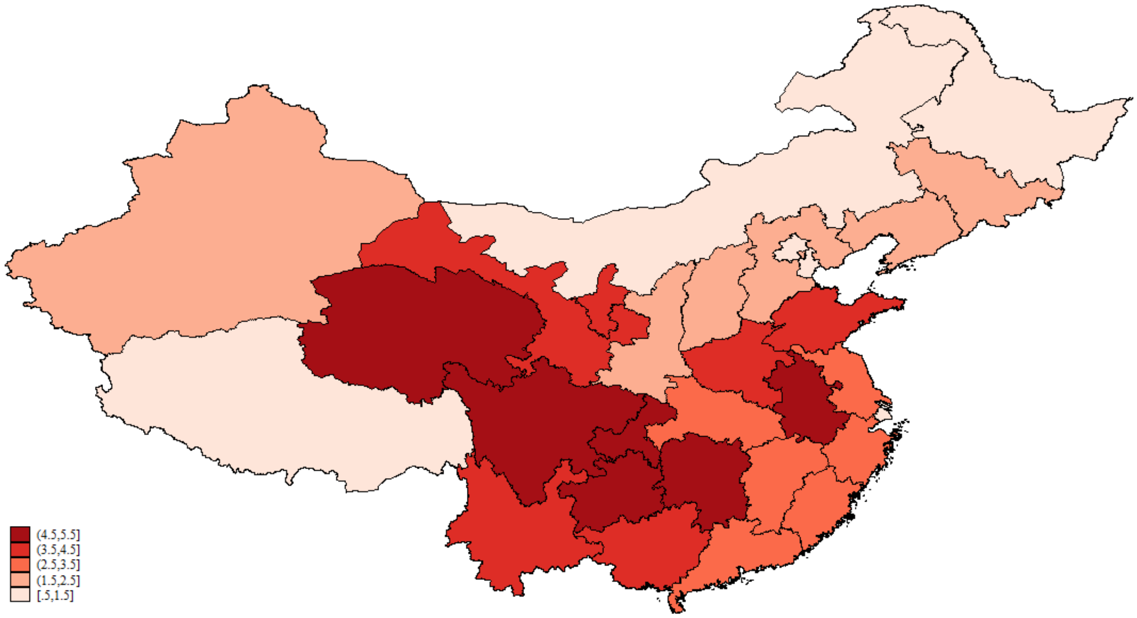
A.1 Severity of Great Famine

We use the 1990 Population Census to construct the measure of the severity of the Great Famine. This census covers 2,600 counties and, for 93.95% of the population, their hukou registration is the same as their county of residence and they have lived in the county for five or more years. For each county and year of birth, we calculate the population and use 1949-1957 and 1963-1970 as basis years to project the counterfactual population. The 1990 Population Census is a 1% sample. We drop any counties with zero persons in any cohort in the period 1949-1970, so, reducing the number of counties in the sample to 2,545. Then, for each county and year of birth, we construct the cohort loss as the difference between the projected and actually recorded number of people. The rate of cohort loss (cohort loss divided by the projected population) in the famine birth cohorts represents the intensity of and hardship during the famine.

Similarly, we use the 2000 Population Census to construct alternative measures of the severity of the famine to use in a robustness check. This census covers 2,869 counties, among which we drop 186 which have zero persons in some cohorts, leaving 2,683 counties.

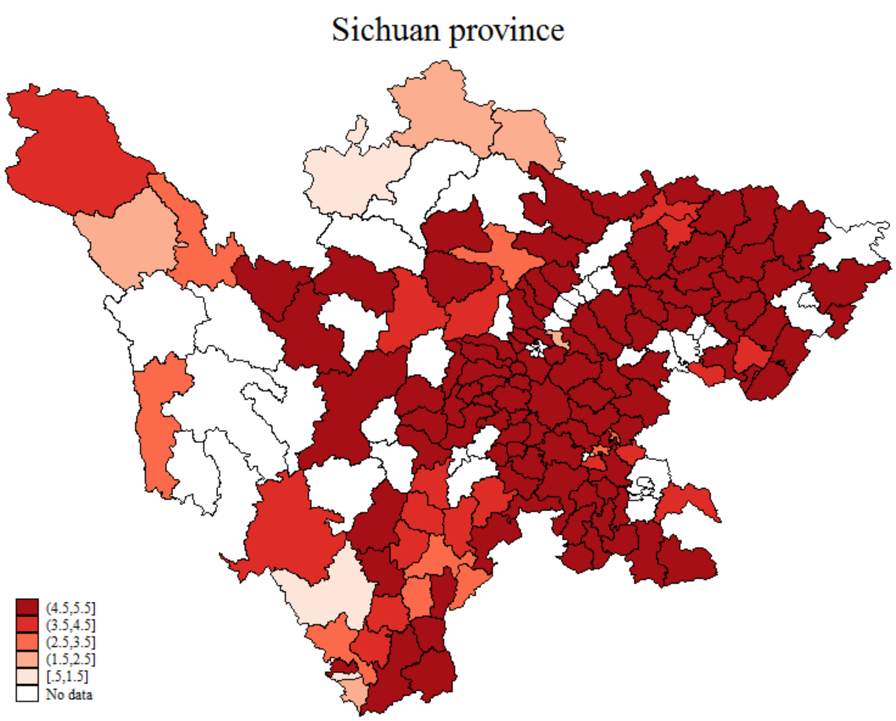
The severity of the famine during the Great Leap Forward differs across provinces and within provinces. Figure A1 depicts the severity by province, defined as the average rate of cohort loss in all counties in the province. Figure A2 depicts the severity by county in Sichuan province, one of the provinces that suffered most during the Great Famine. Apparently, there was substantial geographical variation in the severity of the famine.

Figure A1. Famine: Variation by province



Notes: Based on the quintiles of severity of famine as represented by the average of the county relative cohort loss rate (difference between famine and normal periods) in the province; Darker color represents more severe famine.

Figure A2. Famine in Sichuan: Variation by county



Notes: Based on the quintiles of severity of famine as represented by the county relative cohort loss rate (difference between famine and normal periods); Darker color represents more severe famine.

A.2 Agricultural thermal productivity

From the China Meteorological Administration, we download daily instrumental records for 727 weather stations over the years 1951-70. We measure the thermal agricultural productivity at each station and year by the sum of "degree days" according to equation (3), during the growing season, defined as April 1 to September 30. We exclude any station-year with incomplete coverage during the growing season. Table A1 reports estimates for the serial correlation in thermal agricultural productivity during 1951-1970. Columns (c) and (d) shows that, after controlling for station fixed effects, our construct of thermal agricultural productivity is a random walk.

Table A1. Thermal agricultural productivity: Serial correlation

	(a) Previous year	(b) Previous two years	(c) Previous year	(d) Previous two years
Thermal agricultural productivity (t-1)	0.987*** (0.001)	0.496*** (0.007)	0.001 (0.019)	-0.030 (0.020)
Thermal agricultural productivity (t-2)		0.496*** (0.007)		0.013 (0.017)
Station fixed effects	No	No	Yes	Yes
Stations	727	726	727	726
Observations	9,520	8,522	9,520	8,522
R-squared	0.977	0.983	0.000	0.001

Notes: Sample: All weather stations, years 1951-70; Estimated by ordinary least squares; Dependent variable: Agricultural thermal productivity; Robust standard errors clustered by station in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

In a robustness check to account for the harmful effects of extremely high temperatures, we follow Richie and NeSmith (1991) and compute degree-days by the following four-part piece-wise linear function:

$$H_{cd} = \begin{cases} 0, & T_{cd} < 0 \\ T_{cd} - 8, & 8 < T_{cd} < 33 \\ \frac{25}{8}[41 - T_{CD}], & 33 \le T_{cd} < 41 \\ 0, & T_{cd} \ge 41 \end{cases} \qquad \text{(Eq. A1)}$$

Then, using (4), we compute the thermal agricultural productivity for this alternative measure. The average thermal agricultural productivity for the counties in the 2005 Population Census is 2527.49, which is similar to the average, 2527.43, using our baseline formula for degree-days.

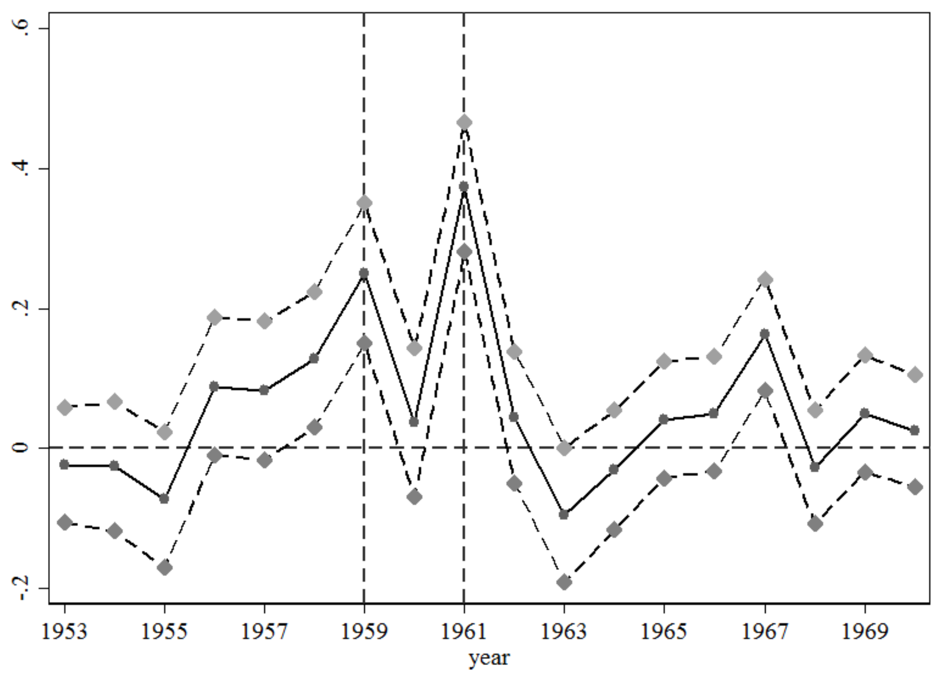
A3. Thermal agricultural productivity and cohort loss rate

To associate the weather stations with the counties, we collect the longitude and latitude of the centroid of each county in the year 2000 from *Michigan China Data Center*, and then match each county with the nearest weather station by the Euclidean distance from the county centroid. On average, the distance between the weather station and county centroid is 29.51 km. To avoid measurement error, we limit our analysis to counties for which the distance from the centroid to the nearest station is less than 100km.

The county level cohort loss rate constructed from 1990 Census is then linked to thermal agricultural productivity by the county identifier in the year 2000. We match to county in 1990 and 2000 based on the record in China's Administrative Division History. [1]After dropping the counties that changed border during this period (and counties with incomplete weather information), the 1990 Census covers 2,326 counties, the 2000 Census covers 2,658, and the combined 1990 and 2000 Censuses cover 2,257 counties.

Figure A3 illustrates the coefficient of year indicators, η_y, in Table 2. The thermal agricultural productivity in the preceding year explains the variation in missing persons except for the unusual years 1958, 1959, 1961, and 1967.

Figure A3. Coefficients of year indicators, η_y



Notes: Graph depicts the estimated year-specific coefficients, η_y, in regression of rate of cohort loss on year dummies (Table 2, column (c)).

[1] www.gov.cn/test/2007-03/23/content_559267.htm contains detail information for us to track changes in administrative division at county level during 1949-2006.

To further validate our construction of the severity of famine, Table A2 presents the estimates of equation (5) with the rate of cohort loss based on the 2000 Census and combined 1990 and 2000 Censuses.

Table A2. Thermal agricultural productivity and rate of cohort loss

	Estimates with 2000 Census			Estimates for combined census		
VARIABLES	β_1	λ_y	η_y	β_1	λ_y	η_y
Thermal agri prody	-0.056**			-0.044**		
previous year	(0.023)			(0.020)		
Year 1953		-0.015	0.068		-0.001	0.027
		(0.018)	(0.044)		(0.013)	(0.031)
Year 1954		-0.028	0.044		-0.015	0.009
		(0.019)	(0.047)		(0.016)	(0.039)
Year 1955		0.016	-0.029		0.023	-0.048
		(0.019)	(0.047)		(0.017)	(0.041)
Year 1956		0.001	0.071		0.002	0.078*
		(0.019)	(0.046)		(0.017)	(0.042)
Year 1957		-0.036*	0.124***		-0.025	0.099**
		(0.019)	(0.048)		(0.018)	(0.043)
Year 1958		-0.001	0.137***		0.008	0.140***
		(0.020)	(0.049)		(0.018)	(0.043)
Year 1959		0.042**	0.249***		0.040**	0.262***
		(0.020)	(0.050)		(0.019)	(0.046)
Year 1960		0.130***	-0.010		0.132***	0.015
		(0.021)	(0.052)		(0.020)	(0.048)
Year 1961		0.048**	0.343***		0.035**	0.392***
		(0.019)	(0.048)		(0.017)	(0.041)
Year 1962		-0.011	0.006		-0.021	0.058
		(0.019)	(0.048)		(0.017)	(0.040)
Year 1963		-0.025	-0.150***		-0.039**	-0.091**
		(0.019)	(0.047)		(0.017)	(0.041)
Year 1964		0.004	-0.025		0.000	0.005
		(0.017)	(0.043)		(0.015)	(0.036)
Year 1965		0.006	-0.010		-0.014	0.059*
		(0.017)	(0.042)		(0.015)	(0.035)
Year 1966		-0.001	0.058		-0.011	0.076**
		(0.016)	(0.040)		(0.014)	(0.032)
Year 1967		0.009	0.152***		-0.007	0.189***
		(0.016)	(0.040)		(0.014)	(0.033)
Year 1968		0.006	-0.039		-0.003	-0.008
		(0.016)	(0.040)		(0.014)	(0.033)
Year 1969		0.021	0.072*		0.008	0.097***
		(0.016)	(0.040)		(0.014)	(0.034)
Year 1970		0.041**	-0.045		0.022	0.012
		(0.016)	(0.040)		(0.014)	(0.033)
Counties	2,658			2,257		

R-squared	0.368	0.477

Notes: This table reports the OLS estimate of equation (5); Dependent variable: Cohort loss rate; with county fixed effects; Column (a) and (d): Coefficient of previous year's thermal agricultural productivity, β_1; Column (b) and (e): Coefficient of previous year's thermal agricultural productivity interacted with year, λ_y; Column (c) and (f): Coefficient of year indicator, η_y; Robust standard errors clustered by county in parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Appendix B. Effect of famine on entrepreneurship: Robustness checks

B1. County analysis

The 2004 Economics Census was the first National Economic Census covering all economic sectors. This census contains detailed financial information for 1,375,148 business entities in 2,860 counties. Among private enterprises, 719,610 are less than 10 years old and have fewer than 100 employees, which we classify as entrepreneurial. Using this definition, we calculate the entrepreneurial firms' share in all enterprises, entrepreneurial firms' share of all enterprise sales, entrepreneurial firms' share of all enterprise sales, median age of private enterprises, and entrepreneurial firms' share of all enterprise sales for each county. After dropping counties with incomplete information for some of the above measures of entrepreneurship and those that changed boundaries between 1990 and 2004, our sample for county-level analysis comprises 2,194 counties.

Table A3 reports checks of robustness of the county-level analysis of the relation between famine severity and entrepreneurship presented in Table 4. Table A3, columns (a)-(e) present IV estimates of (8) with the dependent variable being the entrepreneurial share of all enterprises. Table A3, column (a), is the preferred estimate from Table 4, column (c). Table A3, column (b), includes controls for county level gender ratio, Han ethnicity, migration rate, urbanization, average years of education, relative size of each age cohort (based on the 2000 Population Census). Table A3, column (c), includes control for government expenditure.[2] This accounts for the possibility that the government's policy towards entrepreneurship is endogenously related to the Great Famine. Table A3, column (d), reports an IV estimate accounting for spatial correlation in errors. In Table A3, column (e), firms are defined as entrepreneurial if they are less than 10 years old and have fewer than 500 employees. Table A3, Column (f), reports an IV estimate of (8) with the dependent variable specified as the entrepreneurial firm share of employment of persons with at least high school education in the private sector.

[2] The data is from the Ministry of Finance's Statistical Reports of All Prefectures, Cities, and Counties (Quanguodishixian caizheng tongji ziliao) for the year.

Table A3. Famine severity and county-level entrepreneurship: Robustness checks

VARIABLES	(a) Preferred: Entre. share of private enterprises	(b) Entre. share of private enterprises: Incl. county controls	(c) Entre. share of private enterprises: Incl. govt exp	(d) Entre. share of private enterprises: Spatial correlation	(e) Entre. share of private enterprises	(f) Entre. share of private enterprise high school educated employees
Relative cohort	0.658***	0.713***	0.608***	0.723***	0.698***	0.220***
loss rate	(0.119)	(0.231)	(0.110)	(0.161)	(0.127)	(0.065)
Counties	2,194	2,194	2,172	2,194	2,194	2,194
Hansen J stats	0.56	0.01	0.05		0.35	0.72
p-value	0.45	0.93	0.83		0.56	0.40

Notes: Sample: Enterprises covered by 2004 Economic Census; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Dependent variable is entrepreneurial firm share in all private enterprises, except column (f); Column (a): Baseline estimate from Table 4, column (c); Column (b): Including county-level controls; Column (c): Including county government expenditure; Column (d): Controlling for spatial correlation in errors; Column (e): Entrepreneurial firm is defined as less than 10 years old with fewer than 500 employees; Column (f): Dependent variable is entrepreneurial firm share of employment of persons with at least high school education in the private sector; Robust standard errors clustered by weather station in parentheses (* $p<0.1$; ** $p<0.05$; *** $p<0.01$).

B2. Individual analysis

Table A4 reports robustness checks of the individual-level analysis presented in Table 5. Table A4, column (a), presents the preferred estimate from Table 5, column (c). Table A4, columns (b)-(e), report estimates that variously omit the year 1963 from the base years used to project the counterfactual population, project the counterfactual population by a quadratic model, project by an exponential growth model, and measure the intensity of the famine by the gender ratio rather than cohort loss, project the counterfactual population using 2000 census respectively. Table A4, column (f), reports an estimate using agricultural thermal productivity measured by equation (A1) to account for the negative effect of extremely high temperatures. Table A4, columns (g)-(i), reports estimates with estimated standard errors clustered by province, prefecture, and county respectively. Table A4, columns (j)-(k), include controls for government expenditure and ethnicity, gender, age, urbanization, and education respectively. Table A4, columns (l)-(m), account for spatial correlation in errors and estimate using probit regression with instrumental variable rather than the linear probability model.

Table A4. Famine severity and individual entrepreneurship: Robustness checks

VARIABLES	(a) Preferred	(b) Exclude 1963 from base years	(c) Quadratic projection	(d) Exponential projection	(e) Gender ratio	(f) Projection using 2000 census	(g) Extreme temp. degree day	(h) Instruments: Rainfall
Relative cohort	0.088***	0.087***	0.102***	0.087***	0.307	0.092***	0.088***	-0.011
loss rate	(0.020)	(0.020)	(0.023)	(0.020)	(0.191)	(0.021)	(0.020)	(0.018)
Observations	639,443	639,443	639,443	639,443	635,256	725,214	639,443	639,443
Counties	2,212	2,212	2,212	2,212	2,150	2,527	2,212	2,212
Hansen J stats	1.257	1.150	0.481	1.179	3.413	1.144	1.205	18.75
p-value	0.262	0.283	0.488	0.278	0.0647	0.285	0.272	0.000

Table A4 cont'd

VARIABLES	(i) Std error clustered by province	(j) Std error clustered by prefecture	(k) Std error clustered by county	(l) Government expenditure	(m) Individual controls	(n) Spatial correlation	(o) Probit
Relative cohort	0.088	0.088***	0.088***	0.081***	0.134***	0.0919***	1.843***
loss rate	(0.058)	(0.024)	(0.013)	(0.018)	(0.024)	(0.0331)	(0.281)
Observations	639,443	639,443	639,443	639,443	639,443	2,212	639,443
Counties	2,212	2,212	2,212	2,212	2,212	2,212	2,212
Hansen J stats	0.575	1.583	3.819	0.0964	0.465		
p-value	0.448	0.208	0.051	0.756	0.495		
Marginal effect							1.843

Notes: Sample: Local residents born before 1962 covered by 2005 Population Census; Estimated by two-stage least squares weighted by population except column (o); Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal periods, except column (g); Dependent variable is indicator of owning private enterprise or self-employed; Robust standard errors clustered by weather station in parentheses, except columns (i)-(k) (*** p<0.01, ** p<0.05, * p<0.1). Column (a): Preferred estimate from Table 5, column (c); Column (b): Exclude 1963 from base years for projection of counterfactual population; Column (c): Project counterfactual population by quadratic model; Column (d): Project counterfactual population by exponential growth model; Column (e): Represent intensity of famine by gender ratio rather than cohort loss rate; Column (f): Project counterfactual population using 2000 Census; Column (g): Represent thermal agricultural productivity by sum of degree days according to equation (A1); Column (h): Instrument by rainfall during famine period and difference in rainfall between famine and normal periods; Column (i): Cluster estimated standard errors by province; Column (j): Cluster estimated standard errors by prefecture; Column (k): Cluster estimated standard errors by county; Column (l): Control for county-level government expenditure; Column (m): Control for ethnicity, gender, age, urbanization, and education; Column (n): Account for spatial correlation in errors (estimated by Stata routine, spreg); Column (o): Estimate by probit with instrumental variables.

Appendix C. Individual psychology: Details

C1. Risk tolerance

We use the China General Social Survey (CGSS) 2010, 2012, and 2013, China Family Panel Studies (CFPS) 2010, and the China Household Finance Survey (CHFS) 2011 to derive measures of risk tolerance.

The CGSS is the first comprehensive national wide social survey project in China and is administered by Renmin University of China and the Hong Kong University of Science and Technology. It covers 125 counties from 31 provinces. The survey was conducted twice, during the periods, 2003-2008 and 2010-2019, with 5 waves in each survey. The questions asked differ by wave. Only the CGSS 2010, 2012 and 2013 contain the question on financial investment. We collect and compile the measure of investment from the answers to the following questions:

Are you currently engaging in the following investment activities?
1.Stock 2. Fund 3. Bond 4. Futures 5. Options
6. Real estate 7. Foreign currency 8. Others 9. None of the above

The CFPS is a large-scale, almost nationally representative panel survey conducted by the Institute of Social Science Survey at Peking University. The baseline wave started in 2010 and covered 157 counties from 25 provinces of China. We collect and compile measures of investment from the answers to the following question:

At the end of last year, did you hold any of the following financial products?
1. Stock 2. Fund 3. Bond 4. None of above

The CHFS, administered by the Southwestern University of Finance and Economics, is the only nationally representative survey in China that elicits detailed financial information about household finance and assets, including housing, business assets, financial assets, and other household assets. The first wave of the survey was conducted in summer 2011 with a sample size of 8,438 households and 29,500 individuals. We collect and compile the measure of investment from the answers to the following questions:

Does your family have any stock accounts?
1. Yes 2. No
Which of the following assets does your family own? (Can choose multiple answers)
1.Bonds 2.Mutual Funds 3.Derivatives 4. Wealth Management Products 5.None

We combine responses to these surveys, limiting to individuals born before 1962. After dropping the counties which changed boundaries between 1990 and the time of the survey, our sample comprises 21,965 individuals resident in 248 counties. Of these, 9,596 are from the CGSS, 9,560 are from the CFPS, and 2,809 are from the CHFS.

As investment is a binary variable, Table A5 presents the robustness checks for risk tolerance using probit model.

Table A5. Risk tolerance: Robustness checks

VARIABLES	(a) ivprobit: Share investment	(b) ivprobit: Financial investment	(c) ivprobit:: Owner
Relative cohort loss rate	1.141**	1.065*	0.993*
	(0.578)	(0.551)	(0.510)
Survey fixed effects	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes
Observations	21,965	21,965	21,965
Counties	248	248	248

Note: Sample: Respondents to CFPS 2010, CHFS 2011, CGSS 2010, CGSS 2012, or CGSS 2013 born before 1962; Estimated using STATA routine *ivprobit;* Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years(* p<0.1; ** p<0.05; *** p<0.01).

C2. Self-confidence and craftiness

The CGSS 2011 includes five questions that gauge confidence, viz.,

1. On a basis 1 to 10, where 10 represents highest level of power, please rate your power.
2. Have you ever lost confidence in the past four weeks?
   1. Very frequent 2. Frequent 3. Sometimes 4. Rarely 5. Never
3. Have you encountered any difficulties that you cannot overcome in the past four weeks?
4. 1. Very frequent 2. Frequent 3. Sometimes 4. Rarely 5. Never
5. In the past four weeks, have you have the ever had the following feeling: "I cannot control my life"?
6. 1. Very frequent 2. Frequent 3. Sometimes 4. Rarely 5. Never
7. In the past four weeks, have you have the ever had the following feeling: "I am confident that I can handle problems in my life"?
   1. Never 2. Rarely 3. Sometimes 4. Frequent 5. Very frequent

As self-confidence is measured on integer Likert scales, Table A6 presents the robustness checks for self-confidence using ordered probit model.

Table A6. Self-confidence: Robustness checks

VARIABLES	(a) Power	(b) Confidence	(c) Overcome difficulties	(d) Control own life	(e) Confidence
Relative cohort loss rate	0.957**	-0.434	0.035	-0.386	-0.205
	(0.387)	(0.403)	(0.395)	(0.397)	(0.377)
Observations	1,671	1,628	1,652	1,683	1,680
Counties	47	47	47	47	47

Notes: Sample: Respondents to CGSS 2011 born before 1962; Estimated using STATA rountine *cmp*; Instruments are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years parentheses (* p<0.1; ** p<0.05; *** p<0.01).

Tenacity and resilience are also measured on integer Likert scales, Table A7 shows the estimates of tenacity and resilience using ordered probit and poison model.

Table A7. Tenacity and Resilience: Robustness checks

VARIABLES	(a) Even if not well	(b) Even if long time	(c) Health	(d) Health	(e) Health	(f) Tenacity	(g) Tenacity	(h) Tenacity
Relative cohort loss	-0.933***	-0.734**	0.803**		0.502***	0.333		0.039
rate	(0.281)	(0.288)	(0.368)		(0.142)	(0.404)		(0.098)
Send-down				0.754	0.737***		0.795	-0.016
				(0.755)	(0.270)		(0.813)	(0.261)
Relative cohort loss					-2.359			0.826
rate x send-down					(1.725)			(0.967)
Observations	3,343	3,289	1,544	1,544	1,544	1,530	1,530	1,530
Counties	116	116	73	73	73	73	73	73

Notes: Sample: Column (a) and (b): Respondents to CGSS 2008 or CGSS 2011 born before 1962; Column (c)-(h): Respondents to CGSS 2008 born before 1962; Instruments for relative cohort loss rate are thermal agricultural productivity in famine period and difference in thermal agricultural productivity between famine and normal years; Instrument for send-down is birth cohort times hukou registration type; Column (a)-(d) and (f)-(g) are estimated using STATA routine *cmp*; Column (e) and (h) are estimated using STATA routine *ivpoisson* as STATA cannot implement IV-oprobit with multiple endogenous variables; (* $p<0.1$; ** $p<0.05$; *** $p<0.01$).

The CGSS 2008 includes six questions to elicit respondents' views on the use of guanxi (personal connections) to gain advantage. The questions asked are:

It is very common to use personal connections to gain advantage, do you agree with the following statements based on your own experience:	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
a. Everybody uses guanxi to gain advantage	1	2	3	4	5
b. Using guanxi to gain advantage is a Chinese tradition	1	2	3	4	5
c. People use guanxi to gain advantage only if they have limited power	1	2	3	4	5
d. Using guanxi to gain advantage does not violate the rule of fairness	1	2	3	4	5
e. Using guanxi to gain advantage, the earlier the better	1	2	3	4	5
f. Using guanxi to gain advantage, the closer the relationship, the better	1	2	3	4	5

We use principal components analysis of the responses to the above questions to derive a single measure of attitudes towards the use of personal connections.