The Great Famine and Household Saving in China*

Heng Chen† Maëlys de la Rupelle‡

March 15, 2018

Abstract

The Great Famine in China (1959-1961) is one of the most dramatic tragedies in history, which may have long term consequences for economic behaviors of Chinese population. In particular, we explore whether it continues to have impacts on household saving choices. Employing a dataset across 122 Chinese counties, we find that the saving rates of rural households in 2002 tend to be higher in counties where the famine was severer; and the impact of the famine is even larger, once its severity is instrumented. Evidence from individual preference data shows that people are more willing to cultivate thrift in children in provinces more severely affected. These findings are consistent with the hypothesis that the Great Famine altered the thrifty attitude of survivors and subsequent generations.

Keywords. The Great Famine, saving rate in China, thrift, endogenous preferences

JEL Classification. D14, D91, Q54

*Our research is funded by the Research Grants Council of Hong Kong (Project No. HKU 742211B) and has benefited from the support of the CEPREMAP. The research is part of the project Labex MME-DII (ANR11-LBX-0023-01). We thank François Libois who provided us with the satellite rainfall data for China. We also thank Ying Bai, Christelle Dumas, Yi Huang, Lore Vandewalle, Vincenzo Verardi and seminar participants at UI-UC, University of Hong Kong, Hong Kong Polytechnic University, University of Cergy, Cepremap India-China Conference at the Paris School of Economics, University of Namur, 2016 EEA conference for comments and suggestions.

†School of Economics and Finance, The University of Hong-Kong. Address : Pokfulam Road, Hong Kong. Email : hengchen@hku.hk, Phone : (852) 2857-8506

‡THEMA, University of Cergy-Pontoise. Address : THEMA, 33 bd du Port, 95011 Cergy-Pontoise. Email : maelys.delarupelle@u-cergy.fr, Phone : (33) 134256042

1. Introduction

Historical catastrophic events have shaped our norms, cultures and preferences, which may be passed forward to younger generations and continue to have impacts on their economic behaviors even decades later. In this paper, we show that the China's Great famine (1959-1961), which caused an estimated 30 million death in rural areas, may still influence the consumption and saving choices of Chinese households now. Specifically, our hypothesis is that the Great Famine may alter the thrifty attitude of the generation who experienced it and that of the subsequent generations who did not, which contributes to the high saving rate in the household sector in China.

It is a striking fact that Chinese households save an even larger fraction of their incomes than those in most developed economies.[1] One popular conjecture for such high saving rates is that Chinese households are particularly thrifty, because of Confucian cultural traits.[2] However, Modigliani and Cao (2004) dismiss this type of explanation as "fundamentally baseless" because the household saving rate in China was not always high and indeed was quite low before the end of 1970's.

Distinct from both arguments, we argue that thrift may indeed constitute one important determinant of the high saving propensity of Chinese households, but the thrifty attitude may not necessarily be shaped by the commonly believed cultural factors. Instead, our hypothesis is that the Great Famine in China that lasted three years from 1959 to 1961, may cause Chinese households to be more thrifty (i.e., patient and prudent).[3] First, the hunger experience may install fear of the lack of foods into survivors, so that they behave in a more thrifty manner, even when resources were abundant years later and when the chances of experiencing hunger in the future seemed to be slim. We label it "experience mechanism." Second, adult famine survivors who cherish the value of thrift may indoctrinate their children to be thrifty. Subsequent younger generations also acquire the thrifty attitude towards expenditure from observing the consumption choices of the older generation who experienced the famine.[4] We label it "transmission mechanism." Both mechanisms raise the saving propensity of Chinese population. The behavioral responses hypothesized here may not be ra-

[1] For example, in 2004, the survey data suggest that the household saving rate was 26 percent and 24 percent in rural and urban areas respectively; as a share of GDP, household saving is around 16 percent, significantly higher than that in OECD countries (Kuijs 2006).

[2] Chinese are often described as "notorious savers" and the strong saving propensity is often attributed to the "Confucian values of thrift and frugality" (Moosa 2012). For example, Zhou Xiaochan argue that Chinese are "influenced by Confucianism, which values thrift, self-discipline, ...," therefore the household saving rate is high.

[3] Patience characterizes the attitude towards the weight given to the future utility flow, while prudence characterizes the attitude towards the inter-temporal risk. Both higher patience and prudence contribute to higher saving rates, other things being equal. In our data, we cannot separate the two types of conceptually distinct preference. In this paper, we refer to them as "thrift."

[4] The transmission of the thrifty attitude can take place within family or through local norms.

tional choices made by individuals. In the same vein, Callen (2015), Malmendier and Nagel (2011) and Alesina and Fuchs-Schündeln (2007) demonstrate how individuals' preferences can be shaped by the experiences of natural disasters and macroeconomic shocks as well as political environments.

In this article, we conduct a number of empirical exercises and provide evidence consistent with our conjectures. Specifically, we exploit the county-level and province-level variations in the famine severity to approximate famine experiences and study its effect on the variations of thrifty preferences at the individual level, saving rates at the household level and deposits at the county level. Following Meng, Qian, and Yared (2015), we use the population census data to construct a measure for the famine survival likelihood at the county level. Following Chen and Zhou (2007), we use the provincial death rate data to measure the famine severity at the province level.

Some institutional features allow our approach to be feasible. First, there exists sufficient geographical variation in famine severity (Meng, Qian, and Yared 2015). Second, the disastrous effects of the Great Famine at the aggregate level were unknown to individuals during the famine years and afterwards. Individuals may regard the local severity as a proxy for that of the Great Famine itself and therefore the local variation in famine severity can be useful for identifying its effects.[5] Third, migration had been highly controlled during the famine years and afterwards (Dikötter 2010).[6]

Our first step of investigation is to test whether the famine event still has lingering effects on the saving behaviours of Chinese rural households now. The dataset from Chinese Household Income Project in 2002 offers detailed information on the expenditures and incomes of rural households in 122 counties, which allows us to measure the saving rates at the household level.[7] In the baseline estimations, we study the impact of county-level famine severity on the household-level saving rates, by controlling for a wide range of characteristics including pre-famine political and agricultural conditions. In all specifications, households tend to have higher saving rates in counties where the famine was severer. A one standard deviation decrease in the famine survival index raises the saving rate by 5 percentage points. The economic magnitude of this effect is large, considering the average household saving rate in this sample is 24 percent.

[5]The consequences of the Great Famine were classified information and not accessible to the rural households during or after the famine. There is no consensus in the total famine casualties even nowadays. The earliest estimate of the total death toll is 16.5 million, according to Coale (1981), and subsequent studies constantly revise the estimate up. For example, according to Banister (1991) and Dikötter (2010), the total death tolls were 30 million and 45 million, respectively

[6]Both the migration and the common knowledge of rural households about the severity of the Great Famine may bias our estimate towards zero.

[7]We choose to conduct our investigation in the early 2000s, in order to avoid as much as possible migration-related attrition. At that time, rural-urban migration was mostly restricted to individuals but not entire households.

Once we restrict the sample to households whose oldest member was born after 1964, or a couple of years after the famine ended, we find that the household saving rate still tends to be higher in places with higher famine severity. It may suggest that younger cohorts may be still affected, even though they did not have direct exposure of the famine, which is consistent with our conjecture that the saving propensity can be cultivated and transmitted across generations.

Despite the fact that we use a rich set of controls at the household, village, and county levels, combining survey data with historical information and satellite data, we still cannot remove completely the concern that some pre-famine characteristics may drive both the severity of the Great Famine in late 1950's and the current household saving pattern in 2002. Further, our estimate for the famine severity may also suffer from measurement errors.

To deal with those concerns, we employ a set of instruments combining meteorological conditions and political characteristics at the county level. First, we hand-collect a set of historical data regarding various natural disasters and rainfall shocks before and during the famine years. Even though the Great Famine was primarily a man-made disaster, natural disasters may have worsened the situation and exacerbated the survival likelihood. Second, we take advantage of the variation in *the timing* of radical actions taken by local leaders during the Great Leap Forward, a political movement that preceded the famine. We construct a measure for the radicalism of county leaders' in late 1950's and assume that, conditioning on that, the timing of radical action may be exogenous to the household propensity to saving four decades later. A test of overidentifying restrictions does not contradict our claim. When the famine survival index is instrumented, the impact of the Great Famine on saving rates becomes larger.

Our 2SLS strategy relies on the fact that higher ranked leaders had imperfect information on the people they appointed, and could not anticipate fully what would be the timing of the decisions of local county leaders. It relies as well on the fact that county leaders took decisions partly to answer intrinsic motives regarding their career and their ranking within the party.

Further, we then investigate the impact of the Great Famine on the relationship between income and savings at the county level, which complements our study at the household level. Specifically, we examine whether the heterogeneous famine experiences in late 1950's still affect how local bank deposits, i.e., a proxy for savings in rural counties, respond to rising incomes. With a 14-year panel dataset of more than 1600 rural counties, we find that the magnitude of the increase in bank deposits in response to income growth is substantially larger in counties that were affected the most by the

Great Famine. The results are robust, once we control for the year and county fixed effects and correct for spatial and serial correlations.

The specific channel through which the famine has an impact on current household saving decision is of great interest. We exploit the World Value Survey which provides detailed information about individuals' beliefs and attitudes to verify our conjectured mechanism. That is, the attitude of Chinese households towards thrift may be affected by the famine event. The World Value Survey provides information about how important the respondents think the characteristic of "thrift" is for their children to learn. Using this survey conducted in 1990, 2001, 2007 and 2012, we find that individuals in provinces with higher famine severity tend to value thrift more and are more willing to transmit the value of thrift to their children than those in counties less affected. But we do not find evidence that the famine experiences matter for other characteristics such as hard-work, tolerance and independence, etc. One may conjecture that the Great Famine could be less severe in places where the population were more thrifty. However, such a mechanism predicts that the famine severity should be negatively correlated with the value individuals attach to the thrift characteristic, which is exactly the opposite to what we find.

This paper connects to three strands of literature. First, we propose a new explanation for the determinants of high household saving rates in China. It is complementary to the existing ones, including rising income, increased income uncertainty, demographical changes, sex ratio imbalance and one-child policy, etc., which we review in the following section. Second, our paper enriches the endogenous preference literature by offering some evidence on the preference transmission and shows that historical events may have long lasting effects on economic behaviors of subsequent generations. Third, this study illustrates how the Great Famine can affect the economic choices of Chinese population, which is a new addition to the famine literature which focuses on the determinants of famines and their long-term consequences on health and political beliefs.

The remainder of this paper is organized as follows. In Section 2, we review the related existing literature on the consequences of the Great Famine, on endogenous preferences and on household saving rates in China. In Section 3, we provide a brief background of the Great Famine and explain in details data used in this paper. In Section 4, we present our results on the effect of the Great Famine on saving choices with three empirical exercises. The last section concludes.

2. Literature Review

Our study provides an explanation for the determinants of the high saving rate in the household sector in China, which complements the existing ones. Yang, Zhang, and Zhou (2012) document a robust relation between the rising saving rate in urban households sector and the rapid income growth. Wei and Zhang (2011) argue that household may save because of the competitive motive in the marriage market: the households with male offspring intend to save more to improve the marriage prospects of their children. Choukhmane et al. (2014) show that the implementation of one-child policy had raised the household saving rate in household sector. Song and Yang (2010) empirically demonstrate that the increase in household savings is due to the structural changes in the labor market. Chamon, Liu, and Prasad (2013) empirically shows that the rise in saving rate is related to the higher income uncertainty. Curtis et al. (2015) use a quantitative model to show that a large fraction of the increase in saving rate can be explained by demographic changes.

The hypothesis that we test in this paper is built on the existing endogenous preferences literature. The conjecture that individual life experiences of macroeconomic shocks or natural disasters may shape the preferences has been the focus of a set of empirical studies in different contexts and environments. Callen (2015) shows empirically how the experience of Indian Ocean Earthquake tsunami has shaped the time preference of those survivors. Malmendier and Nagel (2011) study the effect of the Great Depression experience on the risk preferences of individuals and demonstrate that the Great Depression generation was less inclined to take financial risks. Alesina and Fuchs-Schündeln (2007) show individuals political preferences are heavily influenced by political environments.[8]

The preference transmission mechanism has been characterized in Doepke and Zilibotti (2008 and 2014). Dohmen, Falk, and Sunde (2012) provide direct empirical evidence that risk preference can be transmitted from parents to children by using German datasets. Our paper illustrates that the two mechanisms, i.e., how historical events can alter preferences of individuals and how such effects can be carried forward to the subsequent generations through inter-generational interactions, may have profound impacts on the economic choices of current households.

This paper is mostly related to a line of studies on the consequences of the Great Famine. Chen and Zhou (2007) and Meng and Qian (2009) show that early childhood exposure to the Great Famine had large negative effects on the health conditions of

[8]Studies in different contexts have also offered concrete evidence suggesting that natural disasters (e.g., tsunamis, earthquakes) and devastating events (e.g., violent conflicts) may help shape individuals preferences, including Callen et al. (2014), Eckel et al. (2009) and Voors et al. (2012).

survivors. Peng (1987) focuses on the demographic consequences of the Great Leap Forward by analyzing the massive fertility deficits and excess deaths that occurred during and immediately after the Leap. The findings of Bai and Kung (2014) suggest that the famine event in China may have undermined people's belief about the efficacy of collectivization. Our work has a distinctive focus: we study its impact on the thrift attitude of Chinese households and its long lasting effects on the saving patterns in China half century later.

A similar idea that the Great Famine may have impacted the households savings decisions has been explored in the Chinese literature, i.e., Cheng and Zhang (2011), who conjectured a cohort effect that the generation who exposed to the famine when they were young would be affected and tend to save more than others because of such experiences, and demonstrate it by comparing the famine generation and other cohorts.[9] Our work shows that both the experience and transition mechanisms may matter and the younger generations who did not experience famine may be affected indirectly. Our identification approach is also entirely different. In the investigation at the household level, we complement the CHIP data with an enriched set of data characterizing agricultural production conditions and weather shocks at both the village and county levels, the pre-famine condition at the province level, as well as socio-economic and political conditions. We address the endogeneity issue with a set of novel instruments that we constructed. We also demonstrate our conjecture with bank deposits data at the county level and explore the mechanisms with data at the individual level.

3. Measurement and Data

3.1. Historical Background

According to Ravallion (1997), famines were characterized by both high mortality risks and unusually severe threat to the food consumption. The twentieth century witnessed several severe famines with high casualties in Asia (e.g., Bengal Famine in 1943-44), Africa (e.g., Sudan and Uganda famines in 1980s), the former Soviet Union and Holland (1944-45). The Great Famine (1959-1961) in China was the ultimately worst one with much longer duration and more severely curtailed food availability.

Prior to the famine years, the agricultural production and food distribution system had been collectivized over a short period of time from 1955 to 1957. When Mao

[9] In their investigation, they categorize households with the age of household heads and divid them in to four cohorts, i.e., one cohort in formative years during the famine, two older and one younger cohorts. We also find some evidence for heterogeneous effects across cohorts, which is broadly consistent with their findings. But we adopt an different approach, in which the cohort categorization is endogenously arise from our empirical investigations.

launched the "Great Leap Forward" campaign in 1958, all rural households were organized in the form of people's communes. The traditional organization of agricultural production with family as the basic production unit had been replaced completely. Farmers worked collectively in production teams organized and led by the village officials, without rights to withdraw or work separately (Lin 1990). They had no control over the agricultural output and could only consume in communal kitchens managed by the village (Thaxton 2008). Famine occurred when there was a sudden, sharp drop in grain output of 15 percent in 1959 and the grain output continued to decline in the following two years, which worsened the lack of food (Chen and Zhou 2007). The grain redistributed to the village level was below the villagers' subsistence for many reasons, which triggered the massive starvation (Meng, Qian, and Yared 2015). In 1961, the famine ended when the government temporarily increased the amount of grain delivered to the rural areas and reverted some radical policies pursued (Meng, Qian, and Yared 2015 and Walker 2010). The growth rate of population in China recovered and reached 3.57 percent in 1962, which signaled that the threat to the food consumption was alleviated (Ashton et al. 1984).

The official account for this disaster provided by the government, during or after the demographical crisis, was unfavorable weather conditions.[10] Large-scale propaganda campaigns after the famine were intended to convince the local farmers and survivors in rural areas that the bad weather was the primary reason for the observed high mortality rate. This famine period (1959-61) was typically referred to as "three-year natural calamities" in most of the official historical records.

In contrast, recent studies reveal that natural disasters only accounted for a small fraction of the reduction of grain output during the famine years and the high mortality was mainly caused by policy-induced decline in agricultural output and defects of the procurement system. Li and Yang (2005) argue that the major cause of the large scale of death is that the government diverted agricultural resources to expedite industrialization and imposed excessive grain procurement burden on peasants. Meng, Qian, and Yared (2015) show that the inflexibility of the government procurement policy was the major driver for the observed positive correlation between rural mortality rates and per capita food production and it was responsible for the documented severe total famine mortality.[11] Kung and Chen (2011) show empirically that the variation in political radicalism at the province level may explain a large fraction of the variation

[10]For example, the Central Committee of Chinese Communist Party released an official document which provided an explanation for the cause of the food shortage during the famine years and mainly blamed the bad weather for the abnormal mortality rate, see *Decisions on Several Historical Issues of the Communist Party of China since the Founding of the Republic, 1981*.

[11]The level of procurement from a given rural region was typically estimated based on a production target established earlier, which in turn was calculated from past production. And more importantly, such a procurement target did not respond to contemporaneous harvests and production.

in excessive grain procurement.[12] Lin and Yang (2000) quantify the extent to which the urban bias and the decline in food availability contributed to the demographic crisis during the Great Famine. Lin (1990) argued that the key driver of this tragedy is collectivization, which led to a sharp decline in labor productivity.

3.2. Data Construction

Savings of rural households To measure the savings of rural households, we employ the 2002 wave of Chinese Household Income Project (CHIP). Using this dataset has three main advantages. First, its rural sample provides very detailed information about items related to incomes and expenditures of rural households.[13] Second, the sample size of this survey is fairly large: in total, 9200 households from 961 villages and 122 counties were surveyed, which were well distributed across the Chinese territory. Third, it was conducted in 2002, when rural-urban migration of entire households was almost non-existant, so that the saving pattern of rural areas observed is not driven by the departure of migrant families.[14] Even if rural-urban migrants were absent from their place of residence at the time of the survey, it is still likely that their households were surveyed and provided information about their situations.

Regarding incomes, the survey reports both the self-declared income value by households and disaggregate income items from various sources, including individual wages, income from collective welfare fund, farming net income, business net income, income from rent and interests, subsidies, gifts and remittances. In the paper, we use the self-declared income value in our baseline regressions and check the robustness of our results with the sum of incomes from all the reported sources. Regarding expenditures, the survey provides information on items including staple and non staple food, other food, clothing, transport and communication, daily use goods, medical care paid by self and by government, educational expenditure, durable goods, housing repair, other expenditures. In the baseline regressions, we compute the value of expenditures by aggregating all the survey items. Savings of rural households are defined by the difference between incomes and expenditures. To avoid dealing with negative savings, the main dependent variable that we use is the logarithm of income to expenditure ratio, namely household disposable income divided by household expenditures[15]. To

[12] Their investigation shows that among provincial leaders, the alternate members of central committee who aspired for promotion were relatively more radical than full members.

[13] The survey was conducted by the Chinese Academy of Social Sciences in 2003 and inquired rural households about their situation during the preceding year.

[14] In the 1990s, rural migrants in cities without legal documents not only faced prohibitive costs when accessing social services, but also were subject to detention, repatriation and mistreatment. As a result, even though millions of rural emigrants worked outside their place of origin, many of them were temporary migrants whose formal administrative residence was located in hometowns.

[15] In all regressions regarding saving behaviors, we use the logarithm of income to expenditure ratio instead of saving rates. Given this ratio is a monotone transformation of saving rate, we refer to saving

ensure that our results are not driven by the definition of incomes and expenditures, we perform robustness checks by computing this ratio with alternative measures of incomes and expenditures[16].

Famine severity. Ideally, the variations of food availability during the famine years should be the closest approximation for famine experiences. However, such data do not exist at the county or village level. The primary approach in the literature to measuring the variations in famine severity at the county-level is to extract information from the demographical structures. Both Huang et al. (2010) and Meng, Qian, and Yared (2015) used a 1 percent sample of China's 1990 population census to derive a measure of famine intensity at the county level based on the relative size of birth cohorts.[17] This approach is justified because both reduced fertility and increased mortality during the famine years may lead to a smaller size of the cohorts born during the famine years, relative to other cohorts. In our paper, we follow the same approach and compute a *survival index* at the county-level. It is the ratio of the birth-cohort size of famine cohort (1959-1961) to that of non-famine cohort (1954-1957), among the agricultural population. That index is correlated with the probability of survival: the smaller is the index, the more severe is the famine in a given county.

We also exploit variations of famine severity at the provincial level to test our hypothesis. Given the data on province-level death rates are available, we follow Chen and Zhou (2007) and measure the severity by using the excess death rate, i.e., the difference of death rates between 1959 and 1958. Investigations on the provincial level with an alternative measure of famine experiences may corroborate the robustness of our analysis at the county-level.

Political Radicalism Given the famine was largely a product of Mao's radicalism and the political system, it is important to control for the characteristics of political leaders at the local level. We take advantage of information about political leaders at three levels. At the provincial level, we use a measure for the radicalism of the provincial first party secretary, by following Kung and Chen (2011). They show that the first party secretaries who were alternate members of central committee were the most radical and least flexible with procurement target, because they aspired to be promoted, reaching the rank of full members. We include a dummy indicating whether

rate in our discussions unless it causes any confusion.

[16]For instance, households with migrants members are households for whom income and consumption variables might be subject to more measurement error. We perform robustness tests by including or not gifts and remittances in the income variable.

[17]The famine cohort in Huang et al. (2010) refers to the cohort of women born during the famine years (1959-1961) and the non-famine cohort refers to the cohort of women born during the 3 years immediately before the famine (1956-1958) and the 3 years immediately after the famine (1962-1964). In Meng, Qian, and Yared (2015), the famine cohort is defined by rural population that were born during the famine years (1959-1961) and the non-famine benchmark is the average county birth-cohort size over the period 1949-1966.

the provincial first party secretary was an alternate member. At the village level, the historical strongholds for the Communist Party are likely to exhibit particular trends of both the famine severity and local economic decisions. We control for whether the village was a former CCP base in the 1940s; and one fifth of the villages in the CHIP data were.

At the county level, we construct a measure for the county leader's radicalism by using information about building of irrigation projects during the Great Leap Forward, a political movement prior to the famine. In 1950's, it was widely believed that natural disasters were the major sources of agricultural output fluctuations, and that building irrigation projects and water conservation programs, as part of the industrialization in the rural areas, was an effective response to alleviate the situation (People's Daily, 1957, Dec 22nd). In late 1955, Mao made an official comment and encouraged all the counties to build irrigation projects and water conservation programs to "secure the growth of agricultural output." [18] In September of 1957, Mao further called for a movement of building irrigation projects in rural areas of China and instructed that those projects would not be funded by the government and they had to be built by the mass (Chen 2005, page 201). A media campaign was waged immediately and local government officials responded by proposing and building irrigation projects of all kinds.

During this movement, in response to the call of Mao, officials at the county level, who wanted to demonstrated loyalty, diverted a substantial amount of agricultural inputs to build irrigation and water reservation projects. On the one hand, those projects were built at the expense of a substantial diversion of the manpower and other resources from agricultural production. In October of 1957, roughly 30 millions of farmers were mobilized to participate. The number of participants reached around 70 millions at the end of 1957 and 100 millions in January of 1958 (Bo 1993, page 681). Li and Yang (2005) found that the resources diversion was one major cause of the fall of grain output that triggered the famine. On the other hand, those irrigation projects undertaken during the political movement were largely showcases and turned out to be white elephants, which were useless for economic development. Some county gazetteers even explicitly document that those projects were "a pure waste," and were destroyed or abandoned quickly after they were built because of their poor quality. To some degree, the number or frequency of attempting to build those showcase projects should be correlated with the effort local officials had taken to signal political loyalty. Those local officials who were radical during the Great Leap Forward movement were likely to be radical in implementing policies that led to the famine.

[18] Details of this call for movement can be found in "Mao Zedong Collected Works" reprinted in 1999, volume 6, page 451.

We have collected the details of those large irrigation projects built during the Great leap forward in the section of *Major Events* in *County Gazetteers* or in *Irrigation Gazetteers*. We have information about the number of the irrigation projects built in 1956, 1957, and 1958 in each county, whether they had been registered explicitly as wasteful projects, and whether these projects had been abandoned shortly after being built. We consider a county leader to be radical if he built an irrigation project in 1956 or 1957 which proved to be a waste, or if he built more than two irrigation projects in 1958. One third of the counties in our sample is identified as radical.

Soil conditions It is crucial to properly account for soil conditions, because we expect that they account for, to some degree, variations in agricultural production now and at the time of the famine. We compute terrain conditions and soil properties using the GIS Soil and Terrain database (SOTER) of China.[19] This database provides information on topography and on soil composition. It combines a high resolution digital elevation model with data on soil profound physical characteristics, such as the sand, silt and clay fractions. Such soil properties are likely to persist in the long run. At the county level, these characteristics contribute to explain crop suitability and the soil's response to environmental conditions, but are expected to have been affected neither by the severity of the Great Famine nor by agricultural activity. We construct county-level variables on average elevation, ruggedness, and the share of plain terrain. As regards soil characteristics, we classify the soil into different types, like "sandy soil" or "clayey soil", according to the proportion of sand, silt and clay, as defined by the soil texture classifications established by the US Department of Agriculture[20]. We then compute the share of these different types of soil in each county. Last, we also use CHIP survey information on village terrain condition.

Climate conditions We combine two data sources to construct variables related to precipitations: the recent and precise data from the Tropical Rainfall Measuring Mission (TRMM), and the historical data compiled by the University of Delaware.

The "Tropical Rainfall Measuring Mission"offers very accurate satellite data from 1997.[21] The measures are provided for 0.25 x 0.25 degree grid squares (around 25 km x

[19]SOTER GIS information have been originally compiled by the Institute of Soil Science, Chinese Academy of Science (ISSAS) and ISRIC-World Soil Information within the framework of the Land Degradation Assessment in Drylands (LADA, GLADA) project. The primary data were compiled using the SOTER methodology (ISSS, 1986). SOTER unit delineation was based on the raster format of the soil map of China, correlated and converted to FAO's Revised Legend (1988), combined with SOTER landform characterization derived from Shuttle Radar Topographic Mission (SRTM) 90 m digital elevation model (DEM).

[20]Carranza (2014) shows that soil texture accounts for various patterns in agricultural activities but is not affected by land management practices.

[21]The TRMM is a joint project between the NASA and the Japanese Aerospace Exploration Agency which has been launched in 1997 to study tropical rainfalls. The satellite employ a set of five instruments to construct gridded rainfall rates at very high spatial and temporal resolution. Various technological

25 km), which allows us to construct very precise climatic variables at the village level. They are recognized to be one of the most accurate precipitation data, as they combine satellite measures with monthly terrestrial rain gauge data.

We compute long term averages and standard deviations of precipitations using the above described satellite data for the 1998-2014 period with the monthly gridded time series provided by the Department of Geography of the University of Delaware for the 1901-1997 period.[22] We also compute yearly rainfall z-scores by subtracting the long-run mean from rainfall value and dividing the rest by the long-run standard deviation.[23] This allows us to create rainfall anomalies at the village level for the years preceding the survey (1998 - 2001) and at the county level for earlier periods. Finally, the average of rainfall anomalies equates zero over the XXth century, but some counties have faced more positive (or negative) rainfall anomalies during the last decades of the XXth century. If the vulnerability to climate shocks was correlated to some drivers of the mortality during the famine, our results may be biased. Therefore, we control for the weather shocks history the 1980s and the 1990s, by averaging the rainfall anomalies over these two 10-year periods .

Socio-economic conditions The CHIP survey provides a wide range of socio-economic characteristics at the household and village levels, which are important determinants of savings. They may also correlate with local institutions, norms and preferences that have persisted since long before the famine event. At the household level, we include demographic structure, wealth, income perspectives, educational choices. At the village level, the survey provide information on village remoteness, population dynamic as well as economic growth. These variables are measured at the time of the survey, and they are likely to have been affected by the severity of the famine, if the intensity of the famine had long lasting consequences on population's demography and economic environments. To proxy for GDP and its growth at the county level, we use nighttime light intensity of the main county city in 2001, as well as the growth in nighttime light intensity from 2000 to 2001, which is measured by the SMSP/OLS F-15 satellite. Tables 8 and 9 report summary statistics of variables that we have described.

innovations have been used to increase the accuracy of the climatic measures, including a precipitation radar, flying for the first time on an earth orbiting satellite, and the low flying altitude of the satellite. Technical details about TRMM can be found at `https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission`.

[22]This dataset can be downloaded from `http://climate.geog.udel.edu/~climate/html_pages/download.html`.The size of the grid (2.5x2.5 degree grid squares) is closer to the area of a whole county.

[23]A positive value for year t in county j means that precipitations in year t in county j are higher than long-term average. Conversely, a negative value means that precipitations have been lower than that.

4. Empirical Strategies and Results

In this section, we first explore the impact of the Great Famine on saving choices with two datasets. We estimate the impact of famine severity during the famine years (1959-1961) on the saving rates of rural households in 2002. To deal with the potential endogeneity issues, we use a set of instruments for the famine severity. Both strategies provide highly significant and consistent estimates. We then investigate the relationship between the saving pattern and the impact of the famine at the county level, and demonstrate a consistent result with a panel dataset in the period of 1997 to 2010 for more than 1600 rural counties. In the final section 4.6, we provide further evidence that corroborates our conjecture that the thrifty attitude of Chinese households could have been affected by the famine experiences, using data from the World Value Survey.

4.1. Linking the Great Famine to current household saving rate

Our baseline estimations are based on the CHIP dataset. We examine the link between famine severity and saving rates using the following specification:

$$\log(Y/C)_{ikj} = \beta S_j + \eta_i Z_i + \eta_k Z_k + \eta_j Z_j + \epsilon_{ikj} \tag{1}$$

where $\log(Y/C)_{ikj}$ is the log value of income to consumption ratio for household i in village k and county j; S_j is the survival index for county j; Z_i, Z_k and Z_j capture three sets of controls, at the household, village and county levels, respectively; and ϵ_{ikj} is the error term.[24] We cluster the standard errors at the county level.

First, to control for historical conditions, we include variables capturing long run environmental conditions, political radicalism and pre-famine agricultural conditions. Long run conditions comprise of properties of terrain and soil, both at village and county level; village remoteness, long term average and standard deviation of precipitations at the county level between 1901 and 2002, a very rough proxy for exposure to "climate change" (10-year average of rain shocks in the 1980s and in the 1990s), location in a coastal province or in a western one. Political radicalism is measured, at the province level, by the political rank of the province leader within the party; at the county level, by a dummy which takes the value 1 if useless irrigation projects were built in 1956 or in 1957, or if more than two irrigations projects were built in 1958; and at the village level, by a dummy which takes the value 1 if the village was a CCP base in the 1940s. To account for pre-famine condition, we include the share of agricultural output in provincial total output in 1957.

Second, contemporaneous controls include precipitations anomalies at the village

[24] In what follow, we label Y/C as "savings ratio", while "savings rate" refers to $(Y-C)/Y$.

level in 1999, 2000 and 2001, and various household characteristics: household size, age of household head, age of household head squared, the proportion of dependent members, the proportion of women in the household, age of eldest member in the household, average educational attainment, forms of energy use and access to water.

Third, we also control for economic factors that may drive the saving decision, i.e., income, growth of income, a proxy for GDP of the main county city (and its growth) with the nighttime light intensity, whether the village implemented the rural tax reform, the village population growth, the village cadres' salary growth intended to capture the growth at the village level, the income quartiles to which the household belongs, the amount of land cultivated by households in the village, the number of migrant members of the household. In our investigation, we always include the first two categories of controls and present results with and without the third set.

We report the impact of the Great Famine on the saving rate in Table 1. The baseline result is presented in column (1). It is implied that higher chances of survival during the famine years are associated with lower household saving rates in 2002. In other words, households that reside in counties with a higher mortality rate during the famine years, tend to have higher saving rates even more than 40 years later. The magnitude of the impact is sizable: a one standard deviation decrease in the survival index raises the household saving ratio by roughly 7% or the savings rate by 11 percentage points.

Column (2) shows results from the specification with the village level and county level proxies for economic growth. The estimate of the impact of famine survival is quite similar. Column (3) includes economic controls at the household level. The estimated impact of famine survival decreases slightly to 5%.

In column (4), we investigate heterogenous impact across income quartile categories. It appears that the impact of famine severity is stronger for the poorest households. Even with a large set of controls, it is unlikely that we exhaust the relevant common factors that matter for both the pre-famine characteristics and post-famine development paths of the Chinese villages. To ensure that the estimates obtained do not simply capture unobserved heterogeneity at the village level, we include village fixed effects, and compare the effect of the famine across income quartile categories. In column (5) of Table 1, we present results obtained from a village fixed effect model, where we interact the famine survival index with income quartile dummies. The result suggests that the poorest rural households in 2002 tend to be the most responsive to the across-county variations in the survival index. A decrease in the survival index raises the saving rates of the poorest households to the largest extent. Such a finding is consistent with our conjecture of the impact of the famine. The poorest rural house-

holds typically strive for maintaining consumption at the subsistence level and they save mainly for the reason of building a buffer to cushion the future income shocks. In contrast, richer households may have accumulated a buffer stock to smooth income shocks and they may save out of alternative motives, such as bequests or conspicuous consumption (e.g., a fancy wedding). Therefore, the results indicate that a shift in thrifty attitude by the famine is likely to play a role.

In column (6) of Table 1, we report the result by using the provincial excess death rate as a measure for the famine severity. It indicates that one standard deviation increase in excess death rate raises household saving rates by 3.3 percentage points.

One may argue that the famine could be less severe in places where the population (and their offspring) was characterized by higher patience and prudence, so that they had accumulated a larger buffer to cushion the fall of the agricultural output during the famine years. On the one hand, this mechanism, if exists at all, predicts that the saving rates should be higher in counties or provinces with less severe famine, which may bias our *ols* estimates towards zero. On the other hand, as we documented in Section 3, the food consumption in the rural areas had been collectivized prior to the famine years and the rural households were not allowed to store food individually. Therefore, it may be reasonable that famine mortality should be explained by various institutional causes and weather conditions, instead of the propensity to save of the population.

Further, if the transmission mechanism discussed in Section 1 is partially responsible for the saving pattern, we should expect that the younger generations who were not exposed to the famine directly were also affected by the famine indirectly. To investigate this, we restrict our sample to households whose eldest member was born after the end of the Great Famine. Column (1) and (2) in Table 4 present the results. The association between the saving propensity and famine survival index is still significant for the sample with younger generations, which is consistent with our conjecture.

4.2. Results from two stage least squares estimation (2SLS)

One concern with the baseline estimations is that they may suffer from the bias of omitted variables. Some pre-famine characteristics may drive both the severity of the Great Famine in the late 1950's and the household saving pattern in 2002. Another concern is that our measure of famine survival index using the 1990 census may be subject to measurement errors. Such a measurement error should be particularly problematic in counties where the census samples only a few individuals born in 1950s. This could happen, if the county population is small, if there has been a large emigration between the famine and the census, or if the child mortality was too high in the 1950s. To deal with both concerns, we collect and employ a set of instruments for the survival index

at both the county and province levels.

The set of instruments comprises of both meteorological conditions and the timing of actions motivated by political radicalism of the county leaders. For meteorological variables, we take advantage of the fact that climatic conditions preceding the famine and contemporaneous to the famine had played a contrasting role. Favorable meteorological conditions in 1957 may have led to a higher gain output in 1957, thus to a higher grain procurement target or expectation in 1958, which could have increased the severity of the famine. Adverse conditions during the famine years may have aggravated the mortality and reduced the likelihood of survival during the famine years. Li and Yang (2005) has empirically established that the bad weather indeed contributed to the fall of grain outputs.

As for data on meteorological conditions, we use data from three sources. First, we have collected detailed information regarding natural disasters at the county level from the *Natural Disasters* section of the *county gazetteers*, which documented four types of major natural disasters that happened often in rural China, i.e., hailstone, drought, pests invasion and flood. Second, we have a measure of the proportion of affected areas by natural disasters in each province, following Kung and Chen (2011).[25] Third, we utilize the historical precipitation data of the University of Delware to build a standardized measure of yearly rainfalls (z-score) at the county-level.

The second category of instruments is related to political radicalism. As detailed in section 3.2, we proxy the degree of political radicalism during the Great Leap Forward with the number of wasteful irrigation projects and one third of the counties in the sample were labelled as more radical than the rest. While we control for the political radicalism in our baseline regressions, we maintain the assumption that the radical behaviors of county leaders were not entirely predictable and some aspects of their choices can be potentially exogenous to the long term county characteristics that are related to household savings in 2002. We use as instruments two variables describing the timing of the irrigation projects signaling radicalism: two dummies indicating whether a useless project was built in 1956 or 1957[26]. One such project that was carried out later could be more damaging for the survival chances during the famine years that no one has foreseen to follow almost immediately.

To verify the relevance of this set of variables, we run a county-level regression of the famine survival index on several measures of climatic conditions in 1957, 1958 and 1959, and on variables related to the timing and intensity of county leaders' radicalism. Results are shown in Table 10. We find that adverse meteorological conditions in 1957

[25]It is calculated as *shouzai index* in Kung and Chen (2011), who elaborate the data construction and source.

[26]These two dummies are subcategories of the dummy controlling for radicalism in all specifications.

had a positive or no impact on famine survival index. âĂŃIt could be the case that the target for procurement in 1958 or 1959 âĂŃwas lower, because the occurrence of a flood or a drought in 1957 reduced the benchmark.âĂŃ[27] Adverse meteorological conditions in 1958 and 1959 had a negative impact on the famine survival index, as expected.

Regarding the timing of radicalism, counties which implemented radical policies in 1956 rather than in 1957 or in 1958 are associated with a higher famine survival index. It implies that the resource diversion that took place earlier is likely to have had a smaller impact on the severity of the famine.

To reduce the number of excluded instruments, we select the most relevant ones, and interact some of them to increase their joint significance. We consider as excluded instruments the following 8 variables: the 1958 county rain shock, the interaction between county rain shock and the provincial area hit by disasters in 1958 and in 1959, the province area hit by disasters in 1959, the number of disasters which occurred in a given county in 1956 and 1957, the number of disasters which occurred in a given county in 1958 and 1959, whether a wasteful irrigation project was built in 1956, whether a wasteful irrigation project was built in 1957. We report the summary of statistics of the instruments in Table 11 and the first stage regressions in Table 12.

In Table 2, we report the results from the second stage regressions. For comparison of the magnitude of estimated coefficients, we present also results from OLS regressions when we restrict the observations to the sub-sample used in the IV regressions.

In column (2) of Table 2, with the basic set of controls, the magnitude of our estimate becomes larger than that in OLS regressions, which is consistent with the concern of measurement error discussed at the beginning of this section. It implies that in a county with a famine survival index one standard deviation below the mean, households raise their saving ratio by about 9% or the saving rate by 14 percentage points.

In columns (4) and (6) of Table 2, we report our IV estimates when controlling for growth, income, land per household and migration, etc. The magnitude of the coefficient remains stable. In column (8), the specification with the village fixed effects and the interaction between the famine survival index and income quartile of the households indicates that the effect of famine severity on savings is probably driven by the poorest households. Column (3) in Table 4 demonstrates that the impact of the famine on younger cohort without famine experiences still exists, once we instrument the survival index, but it is smaller in magnitude and less significant.

In all estimations, we have checked the relevance and exogeneity of instruments

[27] Note that as the number ofâĂŃâĂŃ concomitant disasters increases, âĂŃthe impact of such a mechanism âĂŃmay diminish, suggestive of non linear effects. âĂŃ

used. Since we instrument the famine survival index with multiple variables, we perform the Hansen's J test using overidentifying restrictions. As our set of instruments includes at least one variable whose exogeneity is hard to question, the rainfall shocks (z-score) in 1958, we can test the exogeneity of other instruments. We report the p-value of the Hansen J Test in each regression, which is consistent with our exclusion restriction. We also report the first stage F-statistics, which implies that our instruments are not weak.

4.3. Robustness checks and interpretation

Less than two decades after the Great Famine, the one child policy that was implemented in China not only affected the demographical structure of Chinese population and family size profoundly, but also has had an impact on the saving decisions of Chinese households (Choukhmane et al. 2014). The effort of county officials to implement the population control may be related to the severity of famine experiences. We construct a measure of the implementation of the one child policy by using the 1990 population census and by dividing the size of the cohorts born during 1978-1981 by the size of the cohorts born during 1975-1977. Such a construction is rather similar to that of the famine survival index. The higher is the one-child policy index, the less rigorously was the policy implemented. As shown in the column (2) of Table 3, we find that in counties where the policy was more rigorously implemented, the saving rate is higher in 2002, which is consistent with the conjecture that such a policy led to higher household saving rates in China. However, the magnitude of the impact of famine on the saving rate is barely affected.

The famine could have also altered political beliefs of the population and therefore led to a change in the distribution of the CCP member in the county, which might be related to political resources and insurance network, and thus to the saving propensity as well. We include a dummy variable for the party affiliation of the household head. Column (3) of Table 3 shows that households with a head who is Communist party member save less, but our estimate for the famine impacts is not affected in terms of magnitude and significance.

Another obvious concern relates to health. The Great Famine could have had adverse long lasting effect on the health of famine survivors and their offspring while health might be an important motive for saving (Cheung and Padieu 2015).[28] We have some limited information about the health conditions of family members of the rural households surveyed. We report our results, when controlling for whether the family members were in bad health, in columns (4) and (5). The effect of famine survival likelihood does not change.

[28]See Chen and Zhou (2007) and Meng and Qian (2009).

Fertility patterns and sex ratios are also likely to have been affected by the Great Famine, either directly or indirectly. Wei and Zhang (2011) show that the sexual ratio imbalance may affect saving choices of Chinese households through the competition of marriage market. We report the results when controlling for the age and the gender of the eldest child (in column (6)), and the county level sex ratio (in column (7)). Our estimate for the famine impact is still robust.

The famine survival index has been computed for all the counties, despite the size of each cohort. However, very small cohorts (with one or two members) could lead to measurement errors in computing the index. To deal with this issue, we consider an alternative: once any birth cohort size is less than 10 members, we compute and use the survival index at the prefecture level instead and report the results in Table 13. Our results remain robust to alternative measures of the survival index.

4.4. The Selection Effect

One may worry that the impact we found is driven purely by the selection effect, i.e., some characteristics that drive the survival likelihood may also drive their saving propensity. That would be the case that the effect we observe was entirely attributed to variations in attrition and to the death of the individuals who were less inclined to save.

To deal with this concern, we first conduct a counter-factual exercise in which all provinces had the same excess death rate as the most affected province, i.e., Sichuan, which experienced the highest excess death rate (measured as the difference in mortality rate between 1959 and 1958). It amounts to eliminating from the sample observations of households with the lowest saving ratio. Depending on the initial excess death rate of the province, a fraction of 1.7% to 2.5% of observations are dropped in each province. When we re-do our estimation with this artificially trimmed sample, the results still hold.

One may also worry that selection was heterogenous across age cohorts. The distribution of age in the 1990 population census shows that the most severely affected cohorts by the famine are the ones born during the famine years. If selection is the only candidate for accounting for our results, we would expect to see that the magnitude of the famine impact on savings found should be strongest for those cohorts.

To investigate this conjecture, we examine and compare the effect of famine severity for two cohorts: the households whose eldest member was born after a cutoff year and those born before that. We include an interaction term between the cohort dummy (taking 1 for the younger cohort) and famine survival index in estimation (1) and repeat the estimation for a series of cutoff years. In Figure 1, we plot the size of the

estimated coefficient for the interaction term and its confidence interval against the cutoff value. To avoid outliers, we trim the sample from the two extreme percentiles (households whose eldest member was born before 1916 or after 1973).

We observe that only when the cutoff birth year of the eldest member was 1950 or later, there is a significant difference between the two cohorts. That is, if the eldest member was (roughly) under 10 years old or not born at the time of the famine, the famine is likely to have a bigger impact on its saving decision than that of the older cohorts. That result suggests that the famine shock may be particularly important for the young individuals at that time and their offspring.

One potential drawback of these regressions is that the results could be driven by some very old or very young households members in the sample. We run a set of regressions focusing on small samples, so that the difference between the oldest and the youngest eldest member of the compared households would be of only 10 years. To do so, we consider a cutoff value X. We restrict our sample to households whose eldest member was born between the years X-5 and X+5 and interact the famine survival index with a dummy for households whose eldest member was born between the year of X and X+5. In Figure 2, we plot the size of the estimated coefficient for the interaction term and its confidence interval against the value of X. We observe that there is no difference between households whose eldest member was born during the famine and those born before. Moreover, households whose eldest member was a child during the Great Famine (or born between 1951 and 1958) were affected by the survival likelihood significantly more than the preceding cohort.

Motivated by the observations, we define four groups of cohorts, distinguishing households whose eldest members was born before the famine, during the famine, or after the famine. In order to have groups of sufficient and comparable size, we define the following five groups: before 1950, 1951-1954, 1955-1958, 1959-1962, and after 1963.[29] We re-estimate equation (1) by including 4 interaction terms between the survival index and dummies of household whose eldest member was born between 1951 and 1954, between 1955 and 1958, between 1959 and 1962, and after 1963.

We estimate specifications with and without controls relating to growth, income, land and migration, and report our results in column (7) and (8) of Table 4, respectively. Consistent with the previous results, the cohort born during the famine years was not the most significantly affected by the famine. Rather those households whose eldest member was a child or not even born were affected more significantly. The

[29]We include cohorts born in 1962 in the group of eldest born during the famine, considering that pregnancies that started in 1961 were likely to count as highly affected by the famine. Restricting this group to households whose eldest member was born between 1959 and 1961 does not change the results. The first three groups are composed of around 1000 households, while the last one (born after 1963) comprises of around 1800 households.

evidence we provide does not suggest that the selection effect is not important, but it may mitigate or even rule out the possibility that the effect of famine is driven entirely by a selection effect, which leaves room for other mechanisms to play a role.

Another likely candidate mechanism that accounts for our results is the response of local institution. It is likely that the local institutions may have evolved and responded to the famine shock, and may have also affected the saving propensity of rural households a few decades later. However, such a mechanism is not likely to be the only driver for our results. While one can easily imagine that a change in institutions could account for a certain heterogeneity between young and old households, it is hard to rationalize the non-monotone pattern shown in Figure 2. Instead, the heterogeneous cohort effect and our main findings can be reconciled by our preferred hypothesis that the famine experiences shifted the thrifty attitude of the population. In Section 4.6, we provide some evidence for this mechanism.

4.5. Income Growth and Saving Deposits: the Impact of Famine

One may conjecture that current saving rates should be also higher at the aggregate level in areas which were hit harder by the Great Famine, if it is true at the household level. However, there exists much fewer data available at the county level regarding savings. But local bank deposits are typically regarded as an approximation for household saving decisions (Wei and Zhang 2011), given few of rural households have access to the financial market and possess other types of financial wealth. Rising income leads a larger amount of banking deposits in all rural counties, but how bank deposits evolve in response to income over time is heterogeneous. In this section, we investigate whether deposits respond to income differently in counties with heterogeneous famine exposure in 1950's. Towards this end, we collect data on bank deposits at the county level during the period of 1997 to 2010, for 1624 rural counties in China, from *National Bureau of Statistics*.

We consider a panel regression model with both year and county fixed effects. Specifically, we first regress the log of deposits per capita on log of GDP per capita in year $t-1$ and its interaction with a dummy of the lowest quartile of the famine survival index. As common or idiosyncratic unobserved characteristics of counties may induce serial and spatial correlation on error terms, we provide non-parametric estimates of the variance of the coefficients following Conley and Ligon (2002).[30] We

[30]The code for STATA has been developed by Hsiang (2010), based on Conley (1999). We modify the code so as to account for fixed-effects and correct for the subsequent loss of degree of freedom. Parameters are estimated by OLS, and standard errors are corrected accounting for serial correlation over 1 period and for spatial correlation up to a distance cutoff set at 200km, representing the mean value of the distance between the centroids of counties.

estimate the following equation :

$$\log(D_{jt}) = \alpha log(Y_{j(t-1)}) + \beta F_j \times log(Y_{j(t-1)}) + \zeta Z_{j(t-1)} + \gamma_t + \gamma_j + \epsilon_{jt} \quad (2)$$

Where $\log(D_{jt})$ is the log of deposit per capita in county j and year t, $log(Y_{j(t-1)})$ the log of GDP per capita in county j and in year $(t-1)$, F_j is the first quartile of the famine survival index, Z_{jt} is the standardised score of precipitations in county j and year $(t-1)$, γ_t and γ_j are year- and county-fixed effects, and ϵ_{jt} is the error term.

In column (1) of Table 5 reports the result. The elasticity of deposit to GDP in counties most severely affected by the Great Famine, i.e., the ones with the lowest 25 percent survival index, is roughly twice as much as in the rest of counties. This result is consistent with our conjecture that the Great Famine in late 1950's still has an impact on the current saving propensity of Chinese households.

One possibility is that those most severely affected counties tend to be underdeveloped and are likely to remain to be rural, which can drive the results. We add an interaction between the log of GDP per capita and a measure of development level (i.e., a dummy of the lowest quartile income group) in column (2), and an interaction between the log of GDP per capita and a measure of urbanization (i.e., a dummy of rural population exceeding 95 percent in the county) in column (3). The magnitude of the effect of famine severity is stable. In columns (5), we include a measure of past county growth, the growth rate between $(t-2)$ and $(t-1)$. The results are still robust. Further, results are also robust to using a conditional maximum likelihood estimator.[31]

4.6. Thrift and the Great Famine

Our hypothesis that the Great Famine still has an impact on the consumption and saving choices even now is based on the mechanism that the thrifty attitude can be affected by life experiences and can be even transmitted. In other words, we should observe that the famine has an impact on the thrifty attitude of the population. To examine whether this pattern exists, we explore the dataset of World Value Survey which provides detailed information about beliefs, values and preferences of individuals, including their attitude towards thrift. In one set of questions, the survey describes a list of ten characteristics to respondents and ask them to name five qualities that they consider to be especially important for children to learn at home. The list of qualities includes Independence, Hard work, Feeling of responsibility, Imagination, Tolerance and respect of other people, Thrift, saving money and things, Determination and perseverance, Religious faith, Unselfishness (generosity), Obedience, Self-expression. It is more likely that people who give priority to the quality of thrift are more thrifty and

[31] Table available upon request.

take more effort to cultivate thrift in their children. Therefore, we expect that people may also value thrift more in places that were subject to severer famine.

We start by examining whether such a relationship between the attitude towards thrift and the famine severity exists at the provincial level. The World Value Survey conducted in 1990, 2001 and 2007 reveal the respondents' location at the provincial level, which makes our investigation feasible.[32] Following Lin and Yang (2000) and Chen and Zhou (2007), we use the excess death rate to approximate the famine severity at the provincial level. Specifically, the excess death rate is defined by the difference between the provincial death rate in 1959 and 1958, which is supposed to be positively correlated with the famine severity. To mitigate the concern of endogeneity, we instrument the excess death rate by the province area hit by natural disasters in 1959.

Specifically, we estimate a simple linear probability model of the following specification:

$$Thrift_{ij} = \beta S_j + \alpha Z_i + \gamma Z_j + \epsilon_{ij}, \tag{3}$$

where $Thrift_{ij}$ is a dummy variable, taking the value of 1 if the respondent i in province j chooses "thrift"as one of the five qualities especially important for their children; S_j is the survival index for province j; Z_i and Z_j are the control variables at the individual and province levels, respectively; and ϵ_{ij} is the error term.

Table 6 presents the results of IV regressions by pooling four waves of the World Value Survey (conducted in 1990, 2001, 2007 and 2012). We control for characteristics of individuals (such as the age, gender and eduction, etc.), pre-famine conditions, provincial level income per capita and provincial leaders' characteristics as well as a wave fixed effect.[33] Given there might be some positive correlation across individuals within cohort, we cluster the standard errors at the cohort level. Column (1) of Table 6 reports the result, which suggests that people are more likely to give priority to thrift as one of the five vital qualities for their children in provinces with higher excess death rates. The magnitude of the coefficient implies that one standard deviation increase in the excess death rate leads to an increase of 6.1 percentage points in the probability of picking thrift.

We also perform the same exercise with the other nine qualities. Specifically, we create nine other dummies, which take value 1 when the corresponding quality is picked as one of the five priorities. Interestingly, we found that the impact of the Great

[32]The World Value Survey conducted in 1995 in China does not release information regarding the location of respondents.

[33]Information regarding respondents' education do not always exits, especially in the early waves of the survey. We therefore use the age of education completion as a proxy for the years of education and include age of education completion, whether the respondent complete education between 15 and 17 or after 23, and whether he or she receives education at all.

Famine is significant for none of them.

Further, we are interested in which cohorts of individuals in the sample tend to drive this relationship that we found. To do so, we re-estimate the equation (3), with a series of smaller samples, with a time window between the year X to $X + 10$. In Figure 3, we plot the size of the estimated coefficient and its confidence interval against the cutoff value X.

Two observations deserve comments. First, we find that the estimate is positive and statistically significant only when X takes a value between 1951 and 1964. Second, the magnitude of the estimate is the largest in subsamples with individuals born between 1951 and 1961 and it diminishes when younger subsamples are used. That implies that it is likely that the famine had stronger impacts on individuals who were born after 1950 and even after the famine years but less strong impacts on older individuals, which is consistent with what we found from in Section 4.4 using household level data.

5. Conclusions

Throughout the course of history, the world has witnessed a multitude of catastrophic events, many of which had tremendous impacts on our cultures, norms and even preferences and shaped the way we behave even many years later. This paper discovers the link between the Great Famine that struck China in late 1950's with the observed saving patterns of Chinese households in recent years. Our work suggests that one likely channel is the Great Famine may have altered Chinese households' thrift preferences and therefore may constitute one important determinant of their high saving propensity. Our results complements the existing explanations for saving patterns of Chinese households and also lend some support to the endogenous preference literature.

References

Alesina, A. and N. Fuchs-Schündeln. 2007. Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *The American Economic Review 97*(4): 1507–1528.

Ashton, B., K. Hill, A. Piazza, and R. Zeitz. 1984. Famine in China, 1958-61. *Population and Development Review* : 613–645.

Baland, J.-M., F. Libois, and D. Mookherjee. 2013. Firewood Collections and Economic Growth in Rural Nepal 1995-2010: Evidence from a Household Panel.

Bai, Y. and J. K.-s. Kung. 2014. The Shaping of an Institutional Choice: Weather Shocks, the Great Leap Famine, and Agricultural Decollectivization in China. *Explorations in Economic History 54*: 1–26.

Banister, J. 1991. *China's Changing Population*. Stanford University Press.

Bchir, M. A. and M. Willinger. Does the Exposure to Natural Hazards Affect Risk and Time Preferences? Some Insights from a Field Experiment in Perú. Working Paper.

Bo, Y. 1993. *Review of Some Important Decision-Makings and Matters*. Central Party School Publishing House.

Callen, M. 2015. Catastrophes and Time Preference: Evidence from the Indian Ocean Earthquake. *Journal of Economic Behavior & Organization*: 199–214.

Callen, M., M. Isaqzadeh, J. D. Long, and C. Sprenger. 2014. Violence and Risk Preference: Experimental Evidence from Afghanistan. *The American Economic Review 104*(1): 123–148.

Cameron, L. and M. Shah. 2013. Risk-taking Behavior in the Wake of Natural Disasters. Working Paper 19534, National Bureau of Economic Research.

Carranza, E. (2014). Soil endowments, female labor force participation, and the demographic deficit of women in india. *American Economic Journal: Applied Economics 6*(4), 197–225.

Cassar, A., A. Healy, and C. Von Kessler. 2011. Trust, Risk, and Time Preferences after a Natural Disaster: Experimental Evidence from Thailand. *Unpublished Manuscript*.

Castillo, M., M. Carter. 2011. Behavioral Responses to Natural Disasters. *Unpublished Manuscript*.

Chamon, M., K. Liu, and E. Prasad. 2013. Income Uncertainty and Household Savings in China. *Journal of Development Economics 105*: 164–177.

Chen, Y. and L.-A. Zhou. 2007. The Long-term Health and Economic Consequences of the 1959–1961 Famine in China. *Journal of Health Economics 26*(4): 659–681.

Chen, Y. 2005. *Collected Works of Chen Yun, Vol.3*. Central Documentary Press.

Cheng, L. and Y. Zhang. 2011. Does Famine Experience in Childhood Influence One's Saving Decision? A New Explanation of China's High Household Saving Rate. *Economic Research Journal (in Chinese)* (8): 119–132.

Cheung, D. and Y. Padieu. 2015. Heterogeneity of the effects of health insurance on household savings: Evidence from rural China *World Development 66*: 84-103.

Choukhmane, T., N. Coeurdacier, K. Jin. (2014). The One-Child Policy and Household Savings. *LSE Working Paper*

Coale, A. J. 1981. Population trends, population policy, and population studies in china. *Population and Development Review*: 85–97.

Conley,T. G. 1999. GMM estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Conley,T. G. and Ligon, E. 2002. Economic distance and cross-country spillovers. *Journal of Economic Growth*, 7(2):157–187.

Curtis, C. C., S. Lugauer, N. C. Mark, et al. 2015. Demographic Patterns and Household Saving in China. *American Economic Journal: Macroeconomics 7*(2): 58–94.

Dikötter, F. 2010. *Mao's Great Famine: the History of China's Most Devastating Catastrophe, 1958-1962*. Bloomsbury Publishing USA.

Doepke, M. and F. Zilibotti. 2008. Occupational Choice and the Spirit of Capitalism. *Quarterly Journal of Economics 123*(2).

Doepke, M. and F. Zilibotti. 2014. Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. Working Paper 20214, National Bureau of Economic Research.

Dohmen, T., A. Falk, and U. Sunde. 2012. The Intergenerational Transmission of Risk and Trust Attitudes. *Review of Economic Studies 72*: 645–677.

Eckel, C. C., M. A. El-Gamal, and R. K. Wilson. 2009. Risk Loving after the Storm: a Bayesian-Network Study of Hurricane Katrina Evacuees. *Journal of Economic Behavior & Organization 69*(2): 110–124.

Fetzer, T. 2014. Can Workfare Programs Moderate Violence? Evidence from India. *The London School of Economics and Political Science, Suntory and Toyota International Centres for Economics and Related Disciplines*.

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences*, 107(35):15367–15372.

Huang, C., Z. Li, M. Wang, and R. Martorell. 2010. Early Life Exposure to the 1959–1961 Chinese Famine Has Long-term Health Consequences. *The Journal of nutrition 140*(10): 1874–1878.

Kuijs, L. 2006. How will China's Saving-Investment Balance Evolve? *World Bank Policy Research Working Paper* (3958).

Kung, J. K. and S. Chen. 2011. The Tragedy of the Nomenklatura: Career Incentives and Political Radicalism During China's Great Leap Famine. *American Political Science Review 105*(01): 27–45.

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

Libois, F. 2015. *Households in Times of War: Adaptation Strategies During the Nepal Civil War*. Ph. D. thesis, University of Namur.

Lin, J. Y. 1990. Collectivization and China's Agricultural Crisis in 1959-1961. *Journal of Political Economy*: 1228–1252.

Lin, J. Y. and D. T. Yang. 2000. Food Availability, Entitlements and the Chinese Famine of 1959-61. *Economic Journal*: 136–158.

Malmendier, U. and S. Nagel. 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics 126*: 373–416.

Meng, X. and N. Qian. 2009. The Long-term Consequences of Famine on Survivors: Evidence from a Unique Natural Experiment Using China's Great Famine. Working Paper.

Meng, X., N. Qian, and P. Yared. 2015. The Institutional Causes of China's Great Famine, 1959–1961. *The Review of Economic Studies*: 1568–1611.

Modigliani, F. and S. L. Cao. 2004. The Chinese Saving Puzzle and the Life-cycle Hypothesis. *Journal of Economic Literature*: 145–170.

Moosa, I. 2012. *The US-China Trade Dispute: Facts, Figures and Myths*. Edward Elgar Publishing.

Peng, X. 1987. Demographic Consequences of the Great Leap Forward in China's Provinces. *Population and Development Review*: 639–670.

Ravallion, M. 1997. Famines and Economics. *Journal of Economic Literature 35*(3): pp. 1205–1242.

Song, M. and D. T. Yang. 2010. Life Cycle Earnings and the Household Saving Puzzle in a Fast-growing Economy. Working Paper, Chinese University of Hong Kong.

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

Voors, M. J., E. E. Nillesen, P. Verwimp, E. H. Bulte, R. Lensink, and D. P. Van Soest. 2012. Violent Conflict and Behavior: a Field Experiment in Burundi. *The American Economic Review 102*(2): 941–964.

Walker, K. R. 2010. *Food Grain Procurement and Consumption in China*. Cambridge University Press.

Wei, S.-J. and X. Zhang. 2011. The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China. *Journal of Political Economy 119*(3): 511–564.

Yang, D. T., J. Zhang, and S. Zhou. 2012. Why Are Saving Rates So High in China? In *Capitalizing China*, pp. 249–278. University of Chicago Press.

Table 1. *Savings ratio and famine - OLS regressions - Dependent variable : log of the savings ratio (income/consumption).*

	(1)	(2)	(3)	(4)	(5)	(6)
Famine survival index	-0.277***	-0.272***	-0.211**			
	(0.09)	(0.09)	(0.09)			
famine X 1 inc. quart.				-0.414***	-0.311***	
				(0.12)	(0.09)	
famine X 2 inc. quart.				-0.191**	-0.092	
				(0.09)	(0.07)	
famine X 3 inc. quart.				-0.162	-0.080	
				(0.11)	(0.06)	
famine X 4 inc. quart.				-0.081		
				(0.11)		
Excess death rate						0.012***
						(0.00)
Hh characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Pre-famine conditions	Yes	Yes	Yes	Yes	No	Yes
Post famine rain shocks	Yes	Yes	Yes	Yes	No	Yes
Growth	No	Yes	Yes	Yes	No	Yes
Land, income and migration	No	No	Yes	Yes	Yes	Yes
Observations	8025	8025	8025	8025	8025	8025
Number of villages (FE)					804	
Adjusted R-squared	0.0909	0.0918	0.325	0.327	0.295	0.327

Significance level : * 10%, ** 5%, *** 1%. Standard errors (in parentheses) are clustered at the county level except in (5) (at the village level). (5) includes village fixed effects.

Table 2. *Savings ratio and famine - 2SLS regressions - The dependent variable is the log of the savings ratio (income/consumption).*

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Famine survival index	-0.179*	-0.340**	-0.212**	-0.372**	-0.153	-0.330**		
	(0.10)	(0.16)	(0.10)	(0.16)	(0.10)	(0.16)		
famine X 1 inc. quart.							-0.246**	-0.314*
							(0.10)	(0.19)
famine X 2 inc. quart.							-0.027	-0.006
							(0.08)	(0.14)
famine X 3 inc. quart.							-0.001	0.130
							(0.07)	(0.13)
Hh characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-famine conditions	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Post famine rain shocks	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Growth	No	No	Yes	Yes	Yes	Yes	No	No
Land, income and migration	No	No	No	No	Yes	Yes	Yes	Yes
Observations	6775	6775	6775	6775	6775	6775	6775	6775
Number of villages (FE)							679	679
Adjusted R-squared	0.0803	0.0854	0.0951	0.0928	0.326	0.323	0.297	0.218
KP F-stat		18.32		17.55		18.93		24.04
Hansen J p-value		0.694		0.592		0.449		0.152

Standard errors (in parentheses) are clustered at the county level. See the Table 7 for the description of controls. **Excluded instruments** : there are 8 instruments in specifications (2)(4) and (6). In col. (8) the set of 15 instruments (for 3 endogenous) includes only interactions with income quartiles.

Table 3. *Savings ratio and famine - 2SLS regressions - The dependent variable is the log of the savings ratio (income/consumption).*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Famine survival index	-0.317**	-0.392**	-0.324**	-0.311**	-0.311**	-0.320**	-0.346**
	(0.16)	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)	(0.15)
One child policy index		-0.306**					
		(0.12)					
Hh head is CCP member			-0.084***				
			(0.02)				
Head/spouse in bad health				-0.065***	-0.064***		
				(0.02)	(0.02)		
Head parents in bad health					-0.267		
					(0.31)		
Kids in bad health					-0.037		
					(0.04)		
Age of eldest child						-0.004*	
						(0.00)	
Age of eldest child squared						0.000**	
						(0.00)	
Eldest child is a boy						0.014	
						(0.07)	
County level sex ratio							-0.772*
							(0.45)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6775	6645	6775	6775	6775	6775	6775
Adjusted R-squared	0.323	0.329	0.326	0.324	0.324	0.323	0.324
KP F-stat	18.30	19.17	18.31	18.32	18.34	18.38	17.18
Hansen J p-value	0.457	0.852	0.460	0.483	0.485	0.457	0.486

Significance level : * 10%, ** 5%, *** 1%. Standard errors (in parentheses) are clustered at the county level. Controls include hh characteristics, pre-famine conditions, rain shocks, growth, land, income, and migration.

Table 4. *Dependent variable : the log of the savings ratio (income/consumption).*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Subsample : eldest born after 1964			Subsample : eldest born before 1959			Without top/bottom pct	
	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS
Famine survival index	-0.231**	-0.243**	-0.306+	-0.276***	-0.204**	-0.347**	-0.223**	-0.164*
	(0.11)	(0.11)	(0.19)	(0.09)	(0.09)	(0.16)	(0.10)	(0.09)
Survival Index X eldest born 1951-1954							-0.174**	-0.126*
							(0.08)	(0.07)
Survival Index X eldest born 1955-1958							-0.038	-0.061
							(0.09)	(0.08)
Survival Index X eldest born 1959-1962							-0.029	-0.059
							(0.11)	(0.08)
Survival Index X eldest born >=1963							-0.084**	-0.079**
							(0.04)	(0.03)
Eldest born in 1951-1954							0.093*	0.035
							(0.05)	(0.05)
Eldest born in 1955-1958							0.077	0.055
							(0.06)	(0.06)
Eldest born in 1959-1962							0.072	0.065
							(0.08)	(0.07)
Eldest born >=1963							0.085**	0.034
							(0.03)	(0.04)
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Long run conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rain shocks history (80s, 90s)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recent rain shocks (1999-2001)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Growth	No	Yes	Yes	No	Yes	Yes	No	Yes
Land, income, migration	No	Yes	Yes	No	Yes	Yes	No	Yes
Observations	1219	1219	1027	5702	5702	4825	8032	7942
Adjusted R-squared	0.0857	0.339	0.343	0.0868	0.324	0.319	0.0857	0.326
KP F-stat			14.59			18.65		
Hansen J p-value			0.181			0.565		

Significance level : $^{+}$ 11%, * 10%, ** 5%, *** 1%. Standard errors (in parentheses) are clustered at the county level. Controls include hh characteristics, pre-famine conditions, rain shocks, growth, land, income, and migration.

Table 5. *Determinants of log deposit per capita - OLS estimation with county and year fixed effect, standard errors corrected for spatial and serial correlation*

	(1)	(2)	(3)	(4)	(5)
Log of GDP per capita (t-1)	0.085***	0.085***	0.086***	0.086***	0.101***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)
Log GCP pc (t-1)X 1st quartile survival	0.112***	0.139*	0.113***	0.139*	0.134*
	(0.010)	(0.072)	(0.010)	(0.072)	(0.072)
Log of population (lag 1)	-0.259***	-0.259***	-0.259***	-0.259***	-0.255***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Yearly rain anomaly (z-score)	-0.007***	-0.007***	-0.007***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log GDPpc (t-1) X 1st quartile GDP		-0.027		-0.026	-0.023
		(0.071)		(0.071)	(0.071)
Log GDPpc X agr pop above97%			-0.005	-0.005	-0.006
			(0.008)	(0.008)	(0.008)
Log (GDPpc t-1/GDPpc t-2)					-0.047
					(0.031)
N	22340	22340	22340	22340	20648

Sample : rural counties from 1997 to 2010.

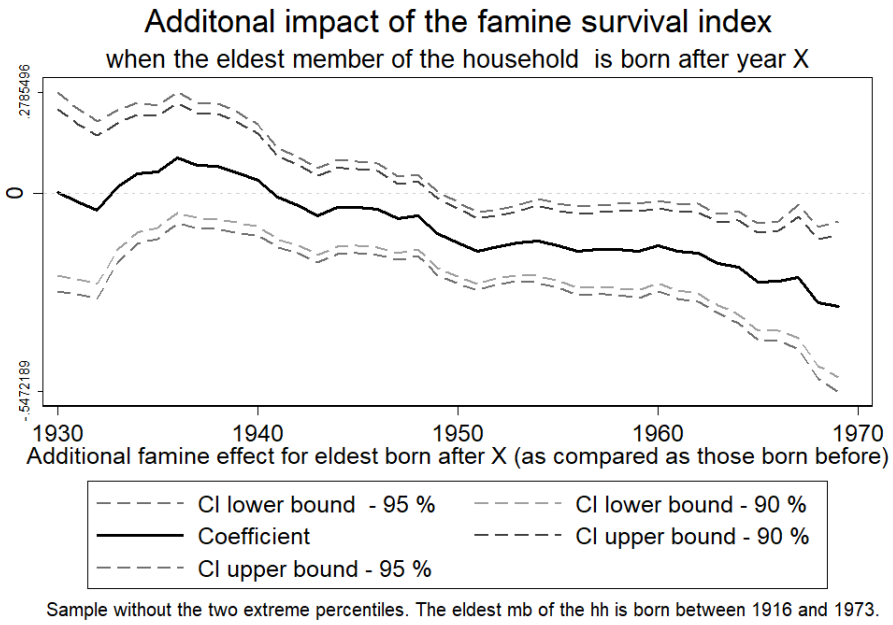

Figure 1. *Cohort effects. The graph plots the estimated coefficient of the interaction term between the famine survival index, and a dummy equal to one if the eldest household member is born after year X.*

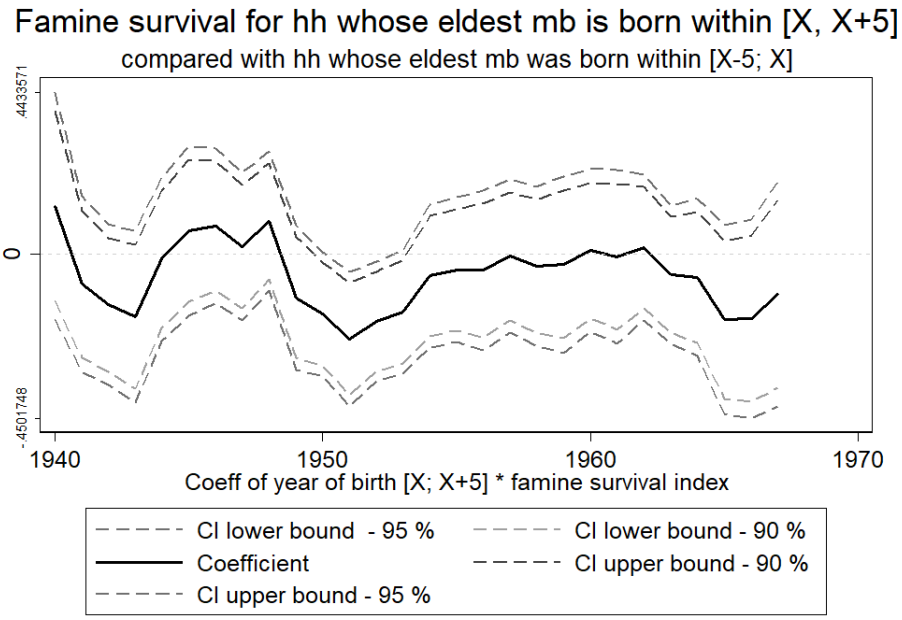

Figure 2. *Cohort effects in ten-year cohorts samples, including households whose eldest member is born during the years [X-5; X+5]. The graph plots the estimated coefficient of the interaction term between the famine survival effect and a dummy equal to one if the eldest member of the household is born after X. OLS regressions.*

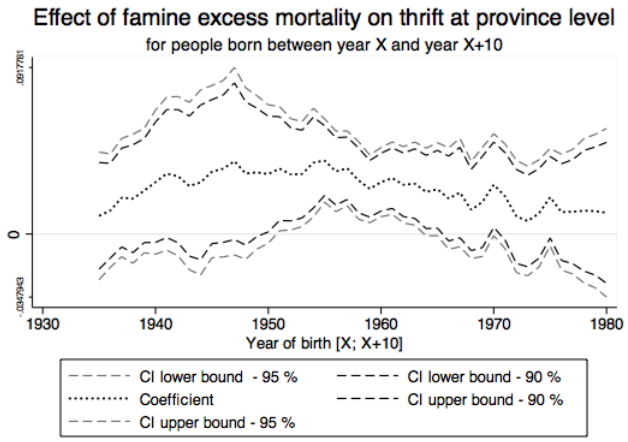

Figure 3. *Cohort Effects.*

Table 6. *Thrift - Linear probability model - Excess death rate, instrumented by province area affected by natural disasters.*

	(1) thrift	(2) hard work	(3) imagination	(4) determination	(5) unselfishness	(6) religion	(7) tolerance	(8) responsibilities	(9) obedience	(10) independence
Excess death rate	0.021**	0.011	-0.004	0.006	0.015	-0.004	0.004	-0.002	0.001	0.001
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
Age, age squared, gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Party membership	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Has a child	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City size	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Per capita income (log)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-famine climatic cond.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-famine economic cond.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial leader	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5554	5554	5554	5554	5554	5554	5554	5554	5554	5554
Adjusted R-squared	0.0301	0.0439	0.0359	0.0412	-0.0157	-0.00913	0.0268	0.00749	0.0204	0.0377
KP F-stat	10.36	10.36	10.36	10.36	10.36	10.36	10.36	10.36	10.36	10.36

Appendix

Table 7. *Variables sources*

Variables	Sources	Computation
Dependent variable		
Savings ratio	CHIP	Log(declared net income/consumption)
Famine measures		
Famine survival index	1990 census	Av. birth cohort size in 1959-1961 / av. birth cohort size in 1954-1957
Excess death rate	Lin and Yang (2000)	
Household characteristics		
Prop. of dependents in hh	CHIP	
Household size	CHIP	
Age of household head	CHIP	
Ethnic minority	CHIP	
Prop. of women in hh	CHIP	
Average education	CHIP	
Water from tap (ref cat. : well)	CHIP	
Water from pump (ref cat. : well)	CHIP	
Energy : coal (ref : fuel)	CHIP	
Energy : firewood (ref : fuel)	CHIP	
Age of the eldest hh member	CHIP	
Pre-famine conditions		
Rainfalls long term annual average	University of Delaware	Average of yearly precipitations during the Xxth century
Rainfalls long term st deviation	University of Delaware	Standard deviation of yearly precipitations during the XXth century
Share of plain land in the county	SOTER	Authors computation, following the USDA soil texture triangle
Share of sloping land in the county	SOTER	Authors computation, following the USDA soil texture triangle
Share of sandy soil in the county	SOTER	Authors computation, following the USDA soil texture triangle
Share of clayey soil in the county	SOTER	Authors computation, following the USDA soil texture triangle
Average altitude in county	SOTER	
Coastal provinces	Beijing, Hebei Jiangsu, Zhejiang, Shandong Guangdong	
Western provinces	Guangxi, Sichuan, Chongqing, Guizhou, Yunnan, Shaanxi, Gansu, Xinjiang	
Provinces in the reference category	Shanxi, Liaoning, Jilin, Anhui, Jiangxi, Henan, Hubei, Hunan	
Village terrain is plain	CHIP	
Village terrain is hilly	CHIP	
Distance to the county seat	CHIP	
Suburban area	CHIP	
Village was a CCP base	CHIP	
1957 province leader is an alternate CCP member	Kung and Chen (2011)	
Irrigation projects in 56-58 signaling radicalism	County gazetteers	
Share of agricultural income in province income in 1956	CNBS	
Post famine rain shocks		
Average rain anomalies in the 1980s	University of Delaware	
Average rain anomalies in the 1990s	University of Delaware	
Rain anomaly in village in 1999	TRMM satellite data and University of Delaware	
Rain anomaly in village in 2000	TRMM satellite data and University of Delaware	
Rain anomaly in village in 2001	TRMM satellite data and University of Delaware	
Land, income and migration		
Village land per hh in 1998	CHIP	
Number of migrants	CHIP	
2nd income quartile	CHIP	Income quartiles defined at the county level
3rd income quartile	CHIP	
4th income quartile	CHIP	
Village implemented tax reform	CHIP	
Log of cadres salary	CHIP	
Growth		
Village cadres salary growth 1998-2002	CHIP	
Village population growth	CHIP	
Nighttime light intensity of the county main city in 2001	DMSP/OLS F-15 satellite	
Growth in the nightime light intensity of the main city 2000-2001	DMSP/OLS F-15 satellite	
Excluded instruments		
County rain shock in 1958 (z-score)	University of Delaware	Z-score of county precipitations in 1958
Province area hit by natural disasters in 1959	Kung and Chen (2011)	
1959 county shock X prov. area hit by disasters	Kung and Chen (2011)	
1958 county shock X prov. area hit by disasters	Kung and Chen (2011)	
Number of disasters in 1956 and 1957	County gazetteers	Sum of the number of disasters (drought, flood, drought, insects, hailstone) experienced by counties in 1956 and 1957
Number of disasters in 1958 and 1959	County gazetteers	Sum of the number of disasters (drought, flood, drought, insects, hailstone) experienced by counties in 1958 and 1959
A 1956 project was a waste	County gazetteers	
A 1957 project was a waste	County gazetteers	
Additional variables		
One Child Policy Index	1990 census	Average birth cohort size in 1978-1981 divided by the average birth cohort size in 1975-1977
Hh head is CCP member	CHIP	
Head or spouse in bad health	CHIP	
Head's parents in bad health	CHIP	
Kids are in bad health	CHIP	
Age of eldest child	CHIP	
Eldest child is a boy	CHIP	
County level sex ratio	1990 census	

Table 8. *Descriptive statistics*

	OLS sample Mean	IV sample Mean
Ratio log(netincome declared / consumption)	.46	.46
	(.57)	(.58)
Famine survival index	.66	.68
	(.22)	(.22)
Household characteristics		
Prop. of dependents in hh	1.16	1.16
	(1)	(1)
Household size	4.11	4.14
	(1.26)	(1.26)
Age of household head	46.15	46.18
	(10.27)	(10.2)
Ethnic minority	.11	.11
	(.31)	(.32)
Prop. of women in hh	.53	.53
	(.15)	(.15)
Age of head squared	2235.6	2236.7
	(986)	(977)
Average education	6.98	7.01
	(2.22)	(2.21)
Water from tap (ref cat. : well)	.3	.3
	(.46)	(.46)
Water from pump (ref cat. : well)	.24	.25
	(.43)	(.43)
Energy : coal (ref : fuel)	.32	.34
	(.47)	(.47)
Energy : firewood (ref : fuel)	.6	.57
	(.49)	(.49)
Age of the eldest hh member	51.38	51.33
	(13.4)	(13.34)
Pre-famine conditions		
Radicalism		
Village was a CCP base	.2	.21
	(.4)	(.41)
1956-1958 irrigation projects signaling radicalism	.44	.34
	(.5)	(.47)
1957 province leader is an alternate CCP mb	.53	.53
	(.5)	(.5)
Environment		
1901-2002 county average of yearly precipitations	1015.1	1035.61
	(416.73)	(420.73)
1901-2002 county st. dev. of yearly precipitations	176.41	179.84
	(56.86)	(56.7)
Share of plain land in the county	.59	.6
	(.38)	(.37)
Share of sloping land in the county	.32	.32
	(.33)	(.33)
Share of sandy soil in the county	.05	.05
	(.08)	(.08)
Share of clayey soil in the county	.07	.06
	(.12)	(.12)
Average altitude in county	477.09	441.42
	(536.2)	(510.65)
Coastal provinces	.26	.29
	(.44)	(.45)
Western provinces	.3	.27
	(.46)	(.44)
Observations	8025	6775

Descriptive statistics for the samples used in the paper.
Standard deviation in parentheses

Table 9. *Descriptive statistics (continued)*

	OLS sample Mean	IV sample Mean
Village terrain is plain	.44	.46
	(.5)	(.5)
Village terrain is hilly	.34	.33
	(.47)	(.47)
Village distance to the county seat	24.15	23.75
	(20.37)	(19.89)
Village located in a suburban area	.05	.06
	(.23)	(.23)
Pre-famine agricultural characteristics		
Proportion of agricultural income in 1956	55.92	55.95
	(11.08)	(10.74)
Post famine rain shocks		
Average rain anomalies in the 1980s	-.06	-.09
	(.29)	(.28)
Average rain anomalies in the 1990s	.26	.27
	(.22)	(.23)
Rain anomaly in village in 1999	.16	.09
	(1.15)	(1.07)
Rain anomaly in village in 2000	.28	.26
	(.98)	(1)
Rain anomaly in village in 2001	-.16	-.12
	(1.19)	(1.23)
Land and migration		
Village land per hh in 1998 (mu)	8.54	8.79
	(9)	(9.4)
Number of migrants in household	.57	.56
	(.77)	(.77)
Income		
2nd income quartile	.25	.25
	(.43)	(.43)
3rd income quartile	.25	.25
	(.43)	(.43)
4th income quartile	.25	.25
	(.43)	(.43)
Village has implemented tax reform	.79	.76
	(.41)	(.43)
Log of cadres salary in the village	2.2	2.2
	(.69)	(.7)
Village cadres salary growth 1998-2002	-.06	-.07
	(.4)	(.29)
Village population growth	.02	.02
	(.1)	(.09)
Nighttime light intensity of the county main city in 2001	11.33	12.04
	(14.08)	(14.94)
Increase in main city's nightime light intensity 2000-2001	.17	.11
	(1.85)	(1.93)
Observations	8025	6775

Descriptive statistics for the samples used in the paper.
Standard deviation in parentheses

Table 10. *Dependent variable : famine survival index. OLS regressions at the county level.*

	(1)
There was a flood or a drought in 1957	0.124**
	(0.06)
There was a flood or a drought in 1958	-0.025
	(0.06)
There was a flood or a drought in 1959	-0.106*
	(0.06)
Total number of disasters in 1956 and 1957	-0.055**
	(0.02)
Total number of disasters in 1958 and 1959	0.026
	(0.02)
Country rain shock in 1957 (z-score)	0.009
	(0.03)
County rain shock in 1958 (z-score)	0.100***
	(0.03)
Country rain shock in 1959 (z-score)	0.035
	(0.03)
Province area hit by natural disasters in 1957	3.426*
	(1.81)
Province area hit by natural disasters in 1958	1.446
	(1.45)
Province area hit by natural disasters in 1959	-5.622***
	(1.09)
Irrigation projects diverting resources in 1956-1958	-0.021
	(0.05)
A 1956 project was a waste	0.211*
	(0.11)
A 1957 project was a waste	-0.092
	(0.16)
Average altitude in county	-0.000
	(0.00)
Rainfalls long term standard deviation	-0.001
	(0.00)
Rainfalls long term average	0.000
	(0.00)
Observations	89
Adjusted R-squared	0.267

Significance level : * 10%, ** 5%, *** 1%. Standard errors (in parentheses) clustered at the county level.

Table 11. *Summary statistics - Set of instrumental variables*

	Mean	**Std. Dev.**	**N**
County rain shock in 1958 (z-score)	-0.252	1.003	6775
1959 county shock X prov. area hit by disasters	0.003	0.041	6775
1958 county shock X prov. area hit by disasters	-0.012	0.037	6775
Province area hit by natural disasters in 1959	0.045	0.027	6775
Number of disasters in 1956 and 1957	1.795	1.528	6775
Number of disasters in 1958 and 1959	1.658	1.602	6775
A 1956 project was a waste	0.047	0.212	6775
A 1957 project was a waste	0.019	0.137	6775

Table 12. *Savings ratio and famine - first stage regressions - The dependent variable is the famine survival index.*

	(1)	(2)	(3)	(4)
County rain shock in 1958 (z-score)	0.133***	0.146***	0.143***	0.105***
	(0.03)	(0.03)	(0.03)	(0.03)
1959 county shock X prov. area hit by disasters	-0.721	-0.803	-0.801	-0.898
	(0.62)	(0.62)	(0.62)	(0.55)
1958 county shock X prov. area hit by disasters	-0.467	-0.739	-0.644	0.363
	(0.90)	(0.90)	(0.92)	(0.82)
Province area hit by natural disasters in 1959	-4.689***	-4.718***	-4.638***	-0.307
	(0.86)	(0.87)	(0.86)	(1.69)
Number of disasters in 1956 and 1957	-0.043***	-0.041***	-0.041***	-0.037***
	(0.01)	(0.01)	(0.01)	(0.01)
Number of disasters in 1958 and 1959	0.024**	0.022**	0.022*	0.009
	(0.01)	(0.01)	(0.01)	(0.01)
A 1956 project was a waste	0.217***	0.212***	0.221***	0.224***
	(0.05)	(0.05)	(0.06)	(0.05)
A 1957 project was a waste	-0.074	-0.128	-0.134	-0.166*
	(0.09)	(0.09)	(0.09)	(0.08)
Controls	Yes	Yes	Yes	Yes
Growth	No	Yes	Yes	Yes
Land, income and migration	No	No	Yes	Yes
Regions FE	No	No	No	Yes
Observations	6775	6775	6775	6775
Adjusted R-squared	0.620	0.635	0.639	0.703

Standard errors (in parentheses) are clustered at the county level. Significance level : * 10%, ** 5%, *** 1%.

Table 13. *Savings ratio and famine - 2SLS regressions - The dependent variable is the log of the savings ratio (income/consumption). The famine survival index is computed using cohorts with at least 10 members.*

	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS
Famine survival index	-0.232**	-0.400**	-0.258**	-0.428**	-0.197*	-0.377**		
	(0.11)	(0.17)	(0.10)	(0.17)	(0.10)	(0.17)		
famine X 1 inc. quart.							-0.297***	-0.388*
							(0.10)	(0.20)
famine X 2 inc. quart.							-0.033	-0.010
							(0.08)	(0.15)
famine X 3 inc. quart.							0.008	0.151
							(0.07)	(0.14)
Hh characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-famine conditions	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Post famine rain shocks	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Growth	No	No	Yes	Yes	Yes	Yes	No	No
Land, income, migration	No	No	No	No	Yes	Yes	Yes	Yes
Observations	6645	6645	6645	6645	6645	6645	6645	6645
Number of villages (FE)							666	666
Adjusted R-squared	0.0814	0.0862	0.0962	0.0937	0.327	0.324	0.299	0.219
KP F-stat		20.26		19.15		19.58		22.48
Hansen J p-value		0.868		0.765		0.663		0.181

Standard errors (in parentheses) are clustered at the county level. See the Table 7 for the description of controls. **Excluded instruments** : there are 8 instruments in specifications (2)(4) and (6). In col. (8) the set of 15 instruments (for 3 endogenous) includes only interactions with income quartiles.