NBER WORKING PAPER SERIES

THE INSTITUTIONAL CAUSES OF CHINA'S GREAT FAMINE, 1959-61

Xin Meng
Nancy Qian
Pierre Yared

Working Paper 16361
http://www.nber.org/papers/w16361

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2010

We thank Abhijit Banerjee, Daron Acemoglu, Robin Burgess, Sylvain Chassang, Andrew Foster, Claudia Goldin, Mikhail Golosov, Avner Greif, Rick Hornbeck, Dean Karlan, Michael Kremer, Roger Myerson, Emi Nakamura, Suresh Naidu, Nathan Nunn, Gerard Padro-i-Miquel, Torsten Persson, Canice Prendergast, Andrei Shleifer, Chris Udry and Shang-Jin Wei for their comments; the partcipants at the Berkeley Development Seminar, UCLA Applied Economics Seminar, Stanford Economic History Seminar, University of Texas at Austin Applied Micro Seminar, University of Toronto Applied Micro Seminar, Tsinghua University Economics Seminar, the Applied Workshop at CCER PKU, Yale University Macro Lunch and Harvard History Tea, and the participants at BREAD-CIPREE, CEPR-DE, the NBER Summer Institute for Economic Growth and the University of Houston Health and Development Conference for useful comments; and Sara Lowes, Ang Sun, David Liu and Jaya Wen for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Institutional Causes of China's Great Famine, 1959-61
Xin Meng, Nancy Qian, and Pierre Yared
NBER Working Paper No. 16361
September 2010
JEL No. N45,O43,P16,P21

ABSTRACT

This paper investigates the institutional causes of China's Great Famine. It presents two empirical findings: 1) in 1959, when the famine began, food production was almost three times more than population subsistence needs; and 2) regions with higher per capita food production that year suffered higher famine mortality rates, a surprising reversal of a typically negative correlation. A simple model based on historical institutional details shows that these patterns are consistent with the policy outcomes in a centrally planned economy in which the government is unable to easily collect and respond to new information in the presence of an aggregate shock to production.

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

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

Pierre Yared
Columbia University
Graduate School of Business
Uris Hall, 3022 Broadway
New York, NY 10027
pyared@columbia.edu

1 Introduction

Over the course of the twentieth century, over 100 million people perished from famines (Sen, 1981). Famines are not only damaging in the short run, but they reduce the quality of life of survivors for decades afterwards.[1] In this paper, we study the causes of China's Great Famine, which began in the winter of 1959-60 and lasted until 1961. In less than three years, the Great Famine claimed the lives of between seventeen and thirty million people, the highest number of fatalities of any single historical event.[2] Although the famine has received growing attention from scholars and the general public in recent years, our understanding of the spatial patterns of famine severity across China and the mechanisms behind it are still very limited.

This study attempts to fill this gap in knowledge with three exercises. First, we present two empirical findings that suggest that government policy in food distribution was a main contributor to the famine. Our results show novel spatial patterns of famine severity and food production that are difficult to reconcile with conventional explanations of famine. Second, to gain insight into the nature of the policy outcome that caused the famine, we document the institutional details of China's centrally planned food procurement regime and China's political climate of the 1950s. The qualitative evidence suggests that government policy was inflexible, a term we use to refer to the government's difficulty in collecting and responding to new information. Finally, we combine the evidence from the first two exercises to formulate a theory of famine under central planning. Our model demonstrates how an inflexible procurement system combined with a drop in aggregate food production can cause a famine with the surprising spatial patterns observed in the data.

The first exercise establishes two empirical findings. The first finding is that the 1959 aggregate per capita production of grain, the main component of the Chinese diet, was well above subsistence needs even though per capita production had fallen by 15% relative to 1958.[3] Our estimates show that aggregate food production in 1959 was at least 205% above the amount required to avoid mortality. We also find that the fall in production, measured as a proportion to past production, was broadly similar across provinces. More importantly, all provinces produced more than what was needed to avoid mortality. The second finding is that regions with higher per capita food production that year suffered *higher* famine mortality rates, a surprising reversal of a typically negative correlation in non-famine years. This surprising cross-sectional relationship is true across all areas and also for a subsample restricted to rural areas.

The main empirical challenge for this study – the same as what all studies of the Chinese

[1] There are many studies on the effects of the Irish Famine (e.g., O'Rourke, 1994). Also, see O'Grada (2007b) and Meng and Qian (2009) for overviews of the literature on the effects of famine.

[2] See e.g., Ashton et al. (1984), Banister (1987), Coale (1981), Peng (1987), and Yao (1999) for estimates of the total mortality from the famine.

[3] Grain makes up over 95% of rural diets (Walker, 1984). Urban workers in China during this period also consumed a grain heavy diet (as they do today). For example, in 1957, an average urban worker in Shanghai, one of the richest cities at the time, consumed approximate 270 kg of grains but only 15 kg of meat in one year (Reynolds, 1981).

famine face – is the questionable quality of the historical data. For political reasons, the government has historically understated the severity of mortality from the famine and overstated food production. We address this issue in several ways. Our historical analysis uses the most recently corrected data available and all of our measures are constructed to conservatively bias against our findings. In addition, in investigating the spatial patterns of famine, we conduct a supplementary analysis with retrospectively created proxies for regional famine severity and grain production. We use regional birth cohort size from the *1990 Population Census* as a measure of famine severity, using the assumption that higher famine severity resulted in lower fertility, higher infant mortality, and thus, smaller birth cohorts. We use two measures to proxy for grain production: a time-invariant measure of suitability for grain cultivation based on agro-climatic conditions and a time-varying proxy measured as annual spring rainfall. Since the retrospective data is not subject to government mis-reporting and can be disaggregated to the much finer county-level, our analysis can address the measurement issues and facilitate the interpretation of the historical province-level analysis.

These empirical results provide several important insights. The fact that enough food was produced to avoid mortality in 1959 means that the famine could not have been solely caused by a drop in food production or the factors behind it. The distribution of food consumption must have played a role in causing the famine. This motivates the central question of our analysis: What transformed a fall in production into the largest famine in history?

Our second finding of a positive relationship between regional mortality rates and per capita food production is surprising in the context of previous studies of famine. Famine studies in market economies typically predict that regional food production is negatively correlated with famine severity, the opposite of the pattern we find.[4] Past studies of the Chinese Famine have similarly not proposed theories that can adequately explain the geographic distribution of the famine that we uncover.[5] In order to explain this second finding, we propose a new theory.

Because market mechanisms for food distribution were not operational in centrally-planned China, government policy in food distribution must have caused the inequality in food consumption which resulted in famine. To understand the factors behind this policy , our second exercise focuses on documenting qualitative historical evidence on the Chinese grain procurement policy and the political regime of the time. The Chinese government experienced a number of practical difficulties faced by all centrally-planned governments of poor countries. Neither local leaders nor peasants had incentives to truthfully report grain production. Limited bureaucratic capacity and China's geographic size and poor transportation and communication infrastructure added to these difficulties by making the central collection of new information extremely costly. To

[4]For example, one of the arguments made by Sen (1981) in his studies of famines in market economies is that local weather shocks can reduce both local food production and income since incomes in developing countries often depend on agriculture so that those in affected areas cannot purchase food from unaffected regions. Therefore, a famine can occur even if aggregate food production is high. However, regional food production will be negatively correlated with famine.

[5]These studies are reviewed in detail later in the paper.

adapt to these challenges, the government's food procurement policy based regional procurement on past regional production, creating a system that could not perfectly adjust to aggregate food production shocks. Thus, a key feature of the centrally planned procurement system was inflexibility – the government could not easily collect and respond to new information. This inflexibility was likely to have been exacerbated by political tensions within China at that time, which caused bureaucrats to be especially reluctant to disobey the prescribed rules.

Our final exercise investigates the extent to which institutional inflexibility can explain the patterns of the 1959 famine in China. We propose a simple model of food distribution where food production varies across regions and the government can reallocate food through procurements and subsidies. We assume that mortality is decreasing in food consumption (i.e., regional mortality rates are decreasing in regional food retention). The key constraint the government faces is inflexibility. Because the government cannot perfectly observe or respond to realized production, procurement policy cannot perfectly adjust to aggregate shocks. For illustrative purposes, we focus on the extreme case in which no adjustment is possible. In this case, the government assigns an inflexible region-specific level of procurement based on expectations of regional production formed from observations of regional endowments for agricultural production or past production. The model predicts that, under this policy, the government over-procures from agricultural areas, amplifying the mortality caused by a fall in production. Moreover, with a production fall that is broadly proportional across regions (which was the case in China in 1959), the procurement policy causes higher over-procurement in more productive rural regions. Through this mechanism, the model shows that policy inflexibility can generate the surprising pattern of more severe famine in higher grain producing regions even if the central government has utilitarian preferences and does not explicitly favor particular regions over others.

To illustrate the model, we provide a stylized example in Table 1. There are three regions: two rural regions (A and B) and a city. Each region has a similar population and therefore similar subsistence needs. For simplicity, we assume the latter to be 100 tons of food. However, agricultural endowments differ across regions. Region A is better endowed than region B and hence produces more food per capita. The city produces no food. As was the case in China, the government restricts population movements so that individuals from low production regions cannot move to high production regions. There are two states of the world. In the normal state which occurs with 80% probability, region A produces 225 tons and region B produces 150 tons. The second state is caused by an aggregate shock which occurs with 20% probability, production is 20% lower for each region, reducing production in regions A and B to 180 and 120 tons. The government expects the aggregate shock to occur with 20% probability. It follows that its expected production of regions A and B are 216 ($225 \times .8 + 180 \times .2$) and 144 ($150 \times .8 + 120 \times .2$) tons, which sum to 360 tons of total production. To illustrate the effect of inflexibility as starkly as possible, we suppose that the government is utilitarian and equalizes expected consumption across its citizens, giving each of the three regions 120 tons of expected consumption. Due to inflexibility, the government consistently procures the difference between expected production and expected consumption in all states of the world , taking away 96 (216-120) and 24 (144-120) tons from regions A and B and giving the city a subsidy of 120 tons. Actual consumption is not constant across states of the world since it equals the difference between actual production and procurement. Therefore, in the good state, regions A and B consume 129 (225-96) and 126 (150-24) tons of grain; in the bad state, they consume 84 (180-96) and 96 (120-24) tons. The city always consumes 120 tons. This example illustrates how inflexible policy can reverse the positive correlation between food production and consumption during an aggregate production shock that is broadly proportional across regions. This surprising reversal is not only true between rural and urban regions, but also between the two rural regions.

The main purpose of our simple model is to explain the two empirical facts that we observe in the data, and by doing so, we provide a novel theory for the causes of famine. The model provides two additional benefits. First, the model offers a framework for understanding the interaction of different factors in causing famine. It illustrates clearly how the Great Famine was partly the outcome of an inflexible and ambitious food procurement policy combined with a large fall in aggregate production. Within this framework, additional factors proposed by other studies (e.g., transport cost, regional favoritism, and local mis-reporting of production) can all contribute further to famine by making it more severe. Second, the model allows us to assess the merits of the Chinese procurement policy of fixing quantities relative to an alternative policy of fixing prices. In an exercise that is similar in spirit to the well-known study by Weitzman (1974), we show that quantity controls, as those used by the Chinese government during the 1950s, are better than price controls if the rural population is sufficiently large in size relative to the urban population and if there is little heterogeneity in the magnitude of productivity shocks across rural regions.

This paper makes several contributions. First, we build on the seminal work of Sen (1981) in exploring the causes of famine. Our results are consistent with his thesis showing that unequal food distribution can cause famine even when aggregate food production levels are sufficient for subsistence. In exploring the causes of famine beyond a drop in aggregate food production, our work is closely related to a growing number of recent studies (e.g., Shiue, 2004, 2005; Burgess and Donaldson, 2010).[6] We also show that the geographic patterns of famine severity and food production can be an important clue for understanding the causes of particular famines. Furthermore, we extend the existing literature by focusing on a non-market economy, illustrating a precise mechanism through which government policy can generate a famine. This is especially important since three of the most devastating and controversial famines in history, China's Great

[6]Shiue (2004, 2005) explores the role of government policy in determining famine relief during the famines in Nineteenth Century China. Burgess and Donaldson (2010) study the role of trade and market institutions in mitigating famines in India. In addition, Shiue (2002) explores the roles of transport costs in China more generally.

There are many studies on the causes of famine beyond the ones cited in this paper. Most of these focus on the reduction in food supply as the primary driver of famine – an argument due to the original ideas of Malthus (1798) – and many have argued that famine was worsened by institutional factors. See O'Grada (2007a), Dreze (1999), and Ravallion (1997) for overviews of recent economic studies on famines.

Famine (1959-61), the Ukrainian Famine (1932-33), and the more recent North Korean Famine (1992-95) have all occurred within non-market economies.[7]

It is important to point out that, while the nature of the fall in production in 1959 and many of the institutional details are specific to the context of the Chinese Famine, neither the occurrence of large imperfectly anticipated aggregate production shocks nor the constraints that drive the inflexibility of China's procurement policy are unique to the context of our study. Because no government can perfectly collect and respond to new information, this constraint of inflexibility can lead to policy failures which amplify the impact of aggregate shocks. This insight is generalizable beyond the Chinese context, especially to other centrally-planned economies.

Our second contribution is to add to studies on the causes of China's Great Famine by providing evidence for two new empirical facts and by developing a theory to explain them.[8] Our theory complements existing explanations of the famine by illustrating how previously documented forces can interact with inflexibility to contribute to famine. Like past studies, our results point to a central planning as the primary cause of the famine. The key difference between our study and past studies is our focus on explaining the transformation of the production drop in 1959 into a devastating famine. Previous studies, which we review in Section 2, have typically focused on the causes of the production drop rather than the distribution of food consumption. Li and Yang (2005) expand on these studies by arguing that the weakened rural labor supply and loss of organic inputs (e.g., seeds and organic fertilizers) to human consumption during the first and most severe winter of the famine caused the famine to persist by reducing production in the subsequent year. Our study abstracts from the production process by taking the production fall in 1959 as given. We also take as given that the labor supply will be weakened and production in subsequent years will be reduced once a devastating famine has occurred. Instead, we focus on the intermediary process and explain how the production fall in 1959 transformed into a famine that winter, when an apparently plentiful production had just been harvested. In this sense, our investigation is most closely related to a study by Lin and Yang (2000), which argues that the famine was caused by the combination of the fall in production and the urban bias in food distribution. Our empirical analysis expands on theirs by demonstrating extensive inequality in distribution, not only between urban and rural areas, but also within rural areas. We then provide a comprehensive theory that explains the observed distributional differences across all regions.

[7]Demographers estimate that approximately 3.2 million died during the Ukrainian famine. The cause of this famine is a subject of intense scholarly and political debate (e.g., see Vallin et al., 2002 for an overview). In North Korea, it is commonly believed that 2-3 million individuals, approximately 10% of the total population, died during this famine (e.g., see Haggard and Noland, 2005; and Demick, 2009). There are very few academic studies or reliable accounts of details related to this famine.

[8]The finding that aggregate production in 1959 should have been sufficient for the subsistence of the population is consistent with many anecdotal accounts and arguably not surprising. However, to the best of our knowledge, past studies have not attempted to systematically quantify the production "surplus" in 1959. The finding that famine was more severe in more productive regions is novel. To the best of our knowledge, it has only been mentioned in a companion paper on the consequences of famine by Meng and Qian (2009), and as a casual observation in Jasper Becker's (1996) book on the Chinese famine.

Our third contribution is to add to studies on the efficiency of central planning and the trade-off between quantity and price controls. Our model generally builds off of arguments made during the historic Socialist Calculation Debate, when Austrian economists such as Von Mises (1935) and Hayek (1945) argued that it was practically impossible for central planners to aggregate necessary information in a timely fashion. Our mechanism for policy outcome due to inflexibility captures this notion.[9] More precisely, our theoretical model and counterfactual exercise are extensions of Weitzman (1974). Our model differs from his in allowing for multiple producers of varying productivity (i.e., multiple rural regions).[10] In this sense, our model is the first to use Weitzman's framework to understand the impact of fixing quantities on the spatial distribution of an aggregate economic shock. Our focus on inflexibility is also closely related to recent research on bureaucratic capacity such as studies by Besley and Persson (2009) and Greif (2008).[11] Finally, our study is related to recent work on the role of state capacity in a government's response to aggregate shocks (e.g., Cohen and Werker, 2008; Kahn, 2005; and Zeckhauser, 1996).

The paper is organized as follows. Section 2 describes the historical background for the famine. Section 3 describes the data. Section 4 presents the empirical evidence. Section 5 uses the empirical evidence to conclude that institutions were greatly responsible for the famine and then documents key features of relevant Chinese institutions. Section 6 introduces a model of procurement policy which is consistent with the evidence. Section 7 concludes. A lengthy Appendix includes additional material that is omitted from the main text for purposes of brevity.

2 Historical Background

The early years of the "New" China government, which came to power in 1949, have been the subject of many scholarly works. It is beyond the scope of this paper to fully describe this period of history; we therefore discuss only issues that are directly related to our study.[12] First, we discuss how the collectivization of farming resulted in the government becoming the only provider of insurance in case of a shock. Second, we describe the 1959 fall in grain production and discuss some of its potential causes. Finally, we examine the timeline of the famine – most of the famine deaths reportedly occurred in the winter of 1959-60, only a few months after the

[9]For example, also see Heal (1969), Malinvaud (1967) and Weitzman (1970) for more studies on the efficiency of central planning.

[10]The result that fixing quantities dominates fixing prices if the rural population exceeds the urban population is a direct application of the more general results in his framework. The result that fixing quantities dominates fixing prices if the heterogeneity in productivity shocks across the rural population is low follows from our extension.

[11]Besley and Persson (2009) analyze the implications of administrative capacity on public policy and Greif (2008) examines government's dependence on administrators to implement policy choices in a historical context.

[12]For more detailed historical accounts of the political organization of China, please refer to the scholarly works of historians such as Fairbank (1986) and Spence (1991). Becker (1996) provides detailed descriptions and a rich collection of anecdotal accounts of the Great Famine from survivors. Also, a two-volume Chinese publication commissioned by the Ministry of Agriculture (1989) entitled *Villages for Thirty Years* documents the details of the social and economic histories of select Chinese villages during the famine era.

Chinese government procured the fall harvest from rural regions. The specifics of the grain procurement policy are described in detail later in Section 5 before we introduce the model.

2.1 New China Reforms 1949-59

The New Communist government of China led by, amongst others, Party Chairman Mao Zedong (in power 1949-76) designed a centrally-planned economy similar to that of the Soviets. Some of the goals of the new government were to equalize land access between tenant farmers and landlords, rapidly industrialize and improve military defense. Historians have not yet fully agreed on why the Chinese government chose to model its economy based on the Soviets. In our study, we take the central planning environment of China as given.

In this economy, where approximately 80% of the population worked in agriculture, grain procurement was seen by the government as the key for development. Most of the grain was used to fund industrialization, which accounted for 43% of government spending during the 1950s (Eckstein, 1977: pp. 186). This included providing grain to urban populations that worked in industry and exporting grain (mostly to the U.S.S.R.) in exchange for equipment and expertise.[13] To a lesser extent, grain was also stored in government reserves as insurance against natural disasters.

Land reforms, which ultimately led to full collectivization by the late 1950s, enabled the government to control and improve agricultural production and distribution (Twitchett and Fairbank, 1985; Spence, 1991: pp. 544). They occurred in three phases. The first began in 1952 and encouraged farmers to form mutual aid teams of six to nine households. The households pooled their assets and land. The second phase began around 1954 and was later called "low level collectivization". This often required all households within a village to pool together their land and assets. However, the return that each household was entitled to depended on the amount of land and assets it contributed to the pool as well as the amount of labor it provided. During this time, agricultural production increased due to the usage of land strips that were formerly used to separate private plots and to increased mechanization, which became more productive due to the pooling of land. During low collectivization, peasants were forced to sell a quota amount of grain to the government at a set low price and allowed to sell their remaining production in markets. Approximately 5% of land was left to peasants as private plots from which they retained all of the production. Many observers have noted with interest that farmers had much more incentive to work on these private plots such that a disproportionately large amount of agricultural production came from them. For example, in 1957, these private plots produced 83% of China's pig and poultry (Spence, 1991: pp. 531-50).

Full collectivization, the third phase that is also often referred to as "high level collectivization" was phased in after low level collectivization. The main change was that although the farmers in each village had contributed land and capital assets for production, they now only

[13] For example, in 1959, approximately 4.3 million tons were exported to the U.S.S.R., which was approximately 2.3% of total production.

received food in return for their labor input. Furthermore, while labor was a requirement in order to receive food and other subsidies, farmers were typically not rewarded for their marginal labor input. In other words, farmers who contributed to the collective received food and other subsidies for their own consumption. But there was no system for rewarding farmers for production beyond their subsistence level (Johnson, 1998). This effectively erased private property rights to land and assets. Private plots were abolished. By 1959, 93% of agricultural land was under high level collectivization (Spence, 1991: pp. 549-50). At this time, mutual aid teams had ceased to exist. Markets for private transactions were also banned (Fairbank, 1986: pp. 281-85).

The central government faced two main problems as it increased the scope of collectivization. The first problem was that farmers were not incentivized to produce more than what was needed for their own consumption, which was guaranteed by the New Communist government. The collective system addressed this by forcing farmers to work under threats of severe punishment, constant monitoring and peer pressure. The second problem was that farmers were incentivized to under-report true production or to hide production. The government attempted to address this by collectivizing the harvest and storage of grains so that harvest went directly from the field to communal storage depots. Communal kitchens were established so that the collective also controlled food preparation and consumption.[14] There are also accounts that the government attempted to collect the little grain that farmers could take in their pockets with virulent anti-hiding campaigns, where fields and even the floors of homes were dug up to expose hidden grain, and where the culprit would typically be publicly humiliated and punished (Becker, 1996: pp. 109).

Chinese peasants, like those in the U.S.S.R. before full collectivization, slaughtered and ate enormous quantities of meat in anticipation of losing the property rights to their animals, reducing China's livestock by half between 1957 and 1958.[15] In response to this, the Chinese government declared that slaughtering animals without permission would be considered a crime against the state and offenders were threatened with severe punishment. By 1959, the remaining livestock and draught animals were typically under-nourished and badly tended as peasants no longer had much interest in caring for them.

For the purposes of our study, these phenomena are important because they mean that by 1959, the state had effectively destroyed private savings and become the only provider of insurance against shocks.

2.2 Grain Production

During the late 1950s, most people in China believed that food production would continue to grow as it had since 1949. From 1949 to 1958, food production grew almost monotonically at

[14] Collectivizing food preparation was also meant to free female labor from household production so that it could be shifted into agricultural production.

[15] See Becker (1996) for comparisons of historical accounts. See Yang (2008) for an economic comparison of the famines in China and the U.S.S.R.

approximately 4% per year. This steady growth was partly due to renewed political stability after decades of conflict and efficiency gains from early phases of collectivization. New farming methods introduced during the collectivization period may also have improved output.[16] In general, there was a belief that China was awash with food. For example, in the fall of 1958, villagers were encouraged to eat as much as they wanted from communal kitchens (Becker, 1996, pp. 80).

The trends in food growth and the attitude of people towards food are important for understanding the broader context of the famine. General optimism undoubtedly played a role in the mindset of the central government when it set high procurement levels and when local leaders delivered large quantities of grain for central procurement. We revisit this issue later in Section 5.

In 1959, grain production fell by 15% from the previous year. After harvest, in mid-October and November, approximately 38% of total production was procured by the central government (see Li and Yang, 2005).[17] Had production in 1959 grown at the same 4% per annum rate as the previous years (on average), procurement would have been 32% of production, a less severe increase from the 26% in 1958, and close to the 1954 procurement rate. The majority of famine deaths occurred in January and February of 1960, two to three months after the grain was procured (Becker, 1996: pp. 94).

The Chinese government has claimed that the famine was a result of bad weather and low production in 1959 (Coale, 1981; Yao, 1999; Peng, 1987; Ashton et al., 1984; and Banister, 1987). However, over time, scholars have estimated the effects of weather on the famine and find that bad can weather can explain at most 14-50% of the famine (Li and Yang, 2005; Kueh, 1995). Studies conducted during the post-Mao era have also hypothesized that the fall in production was associated with Great Leap Forward (GLF, 1958-1960) era policies such as labor and acreage reductions in grain production (e.g., Peng, 1987; Yao, 1999), implementation of radical programs such as communal dining (e.g., Yang, 1996; Chang and Wen, 1997), reduced work incentives due to the formation of the people's communes (Perkins and Yusuf, 1984), and the denial of peasants' rights to exit from the commune (Lin, 1990).[18] Li and Yang (2005) compile province-level panel

[16]However, some of these methods such as multiple cropping may have been unsustainable in the long run and ultimately contributed to the fall in production in 1959 (Spence, 1991: pp. 183). For example, collectivized farms increased the number of crops planted per year. They also followed the Soviets in implementing methods such as the Lysenko method which prescribed deep sowing of seeds for better production. The results of such methods are still debated today.

[17]In our data, which includes production of 27 provinces in 1959 (all except Tibet, Hainan and Sichuan), we observe a 13% fall in aggregate production. In the Li and Yang (2005) data, which includes 24 provinces, there is a 15% fall. In this paper, when we refer to production data, we always refer to our data. However, when we refer to procurement data, we refer to Li and Yang (2005).

[18]Chang and Wen (1997) also argue that the famine actually began in 1958, which resulted in high mortality rates beginning in 1958 and then escalated in subsequent years. However, in this paper, we follow convention and define the famine to have begun following the harvest of 1959. The fall harvest of 1959 was associated with a dramatic increase in mortality rates in 1960, which was almost 22 per 1,000, twice as high as the low point for mortality rates during the 1950s (~10 per 1,000). In contrast, the mortality rates in 1958-59 were moderate at approximately 11-13 per 1,000 (see Figure 1), and they were only slightly higher than the lowest mortality rates during the 1950s. Moreover, the data show that the higher mortality rates before 1960 were driven by one or two

data on grain production and attempt to quantify the impact of each potential factor. They find that in addition to weather, the relevant factors were over-procurement and the diversion of labor away from agriculture for projects such as rural industrialization during the Great Leap Forward. In particular, they argue that over-procurement weakened the agricultural labor force and caused further drops in food production from 1960-61.[19]

Our study takes both the fall in grain production in 1959 and the lasting effects of famine as given. Instead, we focus on how the first drop in production in 1959 caused the high famine mortality rates in the subsequent winter.

2.3 The Famine 1959-61

Most accounts agree that the highest rates of mortality occurred in January and February of 1960, two to three months after the 1959 harvest was procured.[20] Half of the deaths are believed to have been of children under ten years of age (Spence, 1991: pp. 583).[21] The number of deaths is staggering, particularly when one considers that relatively little food is needed to stay alive in the absence of disease and the presence of clean water, two conditions true of rural China after the massive public health improvements undertaken during the 1950s (Fairbank, 1986: pp. 279).[22] The food deficit in rural areas, where most of the mortality occurred, must have been enormous. Many believe that communal food storage exacerbated the shortage of food in the countryside because peasants, not realizing the extent of food scarcity after procurement, did not fight the collective for control over the food supplies so that they could optimize their consumption of food stores through the winter. Most peasants had little cause for suspicion as most remember a plentiful harvest and plentiful meals from communal kitchens. To many, it was a surprise when the food suddenly and completely ran out during the winter months (e.g., Yang, 1996; Chang and Wen, 1997).

It seems that the Chinese government began to respond to the famine as early as the spring

provinces (e.g., Gansu and Yunnan) (see Appendix Figure FA2).

[19]See Yang (2008) for a recent review of the studies on the causes of China's famine. Li and Yang (2005) use a dynamic model to argue that erroneous expectations of production caused over-procurement in 1959, which in turn reduced inputs for agricultural production (e.g., labor was weakened, and seeds were consumed by hungry peasants) in 1960, leading to a further decline in production. They calculate grain retention after procurement in 1959 to be 223 kg per person and in 1960 to be 212 kg per person. Our study builds on the spirit of theirs by making the additional argument that estimates for grain retention such as theirs are not actually low enough to cause mortality. Therefore, for so many to have died, there must have been distributional problems which caused some to have much less than the national average. See section 4.1.

[20]Mortality data from the famine is not available at a monthly frequency. Historians and survivors provide consistent accounts that almost all of the mortality happened during the first winter. For a detailed description see Becker (1996) and Fairbank (1986).

[21]Younger children may have been more vulnerable to famine for biological or food allocation reasons. They may have been physically more vulnerable to nutritional deprivation, which for infants could reflect a decrease in the supply or quality of breastmilk from mothers. Alternatively, households may decide to allocate more food towards adult members who can use their labor to convert these calories into more food consumption for the household. For similar reasons, the elderly are considered to be more vulnerable in times of shocks.

[22]The Chinese Famine is similar to the Leningrad and Dutch famines where mortality is mostly due to starvation rather than succumbing to infectious disease (O'Grada, 2007b).

of 1960. The government implicitly acknowledged the famine by reducing food rations for urban areas, but nonetheless, urban areas never experienced extreme famine or mass starvation. The government also reassigned labor that had once been transferred to cities back to the agricutural sector in order to supplement the greatly weakened rural labor supply and prevent further falls in production. However, the number of returned workers was small relative to the demands of the agriculture sector and organic inputs to production such as seeds and organic fertilizer had been consumed, severley limiting the productivity of labor. Production in 1960 declined dramatically from 1959. After the bad harvest of 1960, the government delivered large amounts of grain to famine stricken areas. There is no detailed or systematic historical account of where exactly the grain replenishment came from, but presumably, some of the grain came from government reserves. By observing the large decreases in the government's investment in industry and military expenditures which decreased by 30% from 1959 to 1960 (Gittings, 1967: pp. 309), we can speculate that much of the grain replenishments also came from what was otherwise designated for military and industrial purposes. Production slowly recovered in the subsequent years.[23]

There are two important facts to keep in mind in understanding the chronology of the famine. First, the communist party rose to power promising to end famines and the Communist party's primary stronghold was the rural population. The Chinese Communist Party (CCP) membership of approximately 5.2 million in 1957 was approximately 70% rural.[24] Therefore, we have little reason to believe that the Chinese government would have had political motivations for desiring a famine, especially one that struck rural areas. Throughout this period, the party leadership in China was aware that they could not politically afford to implement a policy that would cause a famine like what occurred in the U.S.S.R. (Spence, 1991; pp. 575-76). While one may suspect some government officials of being callous, the selfish desire of remaining in power would make it hard to believe that the government intended for the famine to occur.[25] The unintentional aspect of causing the famine is important to keep in mind when we specify the objective function of the central planner in Section 6.1.

The second fact to keep in mind is that unlike any other famine in history, the Chinese famine did not result in mass migration. The central and local governments together prevented the population of famine stricken regions from migrating to other regions in search for food. This undoubtedly exaggerated mortality rates that resulted from the fall in food production. For our study, it also means that regional variation in the size of the birth cohort during the famine can be used as a proxy for regional variation in famine intensity.

[23]It is difficult to retrospectively account for grain allocation in a rigorous way. Part of the reason is the lack of data. But the difficulty also comes from the fact that historical accounts typically report grain in units of tons and government expenditures in units of RMB (yuan), and we have not been able to uncover how grain was valued.

[24]Approximately, six out of every one thousand rural residents were CCP members.

[25]This is slightly different from the situation of the Soviet Ukrainian famine. Soviet Communist party membership was approximately 70% urban and in the case of the Ukrainian Famine (1932-33), the famine occurred in a region that was historically not aligned with the Stalin-led government.

3 Data

The empirical analysis attempts to establish the geographic and temporal patterns of food production and famine severity. The main empirical challenge is to address issues regarding measurement error in the historical data on production and mortality. As in previous studies on the Chinese Famine, we face the difficulty that the Chinese government systematically over-reported production and under-reported mortality to understate the failures of government policy and famine severity. We employ a variety of methods to address this problem. The most important one is the use of retrospectively constructed proxies of famine severity and grain production in a supplementary analysis to check the robustness of the findings regarding the spatial distribution of famine severity. The results from using these proxies, which are not subject to systematic government mis-reporting, indicates whether our main results from using the historical data are driven by measurement error. The following discussion provides a detailed description of each data source and highlights important facts from the descriptive data.

3.1 Historical Data

In the historical analysis, we use historical data on production, mortality rates, and population. These data have been retrospectively corrected by the National Bureau of Statistics (NBS) to account for mis-reporting or missing data from past contemporaneous reports. We use the most conservative esimates of production and mortality of which we are aware. This data reports similar or lower production and similar or higher mortality than previous comparable studies. More discussion on the corrections and comparisons with other data is included in the Data Appendix.

The annual province-level historical data on grain production, population and mortality rates are from the *Comprehensive Statistical Data and Materials on 50 years of New China* (CSDM50) published by the China Statistical Press in 1999. In our study, we use data from all available years, 1949-98. There are two main issues that arise in our analysis of this historical series. The first issue regards accuracy: for political reasons, the government has historically overstated production. Such a bias would lead us to incorrectly overestimate available food per capita in 1959.[26] The second issue regards the completeness of the province-level data: several provinces report missing values for historical production data. NBS officials suggest that these are typically the provinces for which they could not make accurate revisions, so our province level data includes 27 provinces.[27]

[26] This can cause our calculations of aggregate food needs and production in 1959 to show that there was enough food even if actual production was below subsistence needs. Therefore, when we interpret the estimated production surplus we will not interpret the level of surplus literally. Rather, we will argue that the magnitude of the surplus is so large, that it is highly implausible that it is generated only by over-reporting. See Section 4.1 for more discussion.

[27] Our province analysis excludes Sichuan, Tibet and Hainan. Each of these provinces were missing either mortality or production data for 1959. Another issue is that mortality data is not available for many provinces before 1954. For example, in 1949, data on mortality is only available for 15 of the 27 provinces. This is not

The corrected population data takes into account corrected mortality data and fertility changes observed in famine year birth cohort sizes from later population censuses. From discussions with official statisticians, we believe that the population data may still understate population losses from the famine, but this inaccuracy should not affect our estimates for food requirements in 1959, since famine deaths mostly occurred in 1960.[28] Since only total population is reported, we use the sex and age distribution from Coale (1981), which is based on the *1954 Population Census*, to calculate caloric needs. We assume that in each year, the population for each province had the same sex and age distribution as the national level data in 1954. This should be a reasonable estimate for the years close before and after 1954, which is the focus of our study.[29]

We plot average mortality rates over time in Figure 1. It shows that over the fifty years of the New Communist regime, there was a strong secular trend of declining mortality such that average mortality rates fell from approximately fifteen per 1,000 to approximately five per 1,000. The data show that this decline was not strictly monotonic – there were occasional mortality increases of up to 10% relative to the previous year's rates (e.g., in 1964, 1972, 1990). However, none of these increases are close in magnitude to what occurred in 1960, when mortality rates almost doubled from approximately 11 per 1,000 in 1958 to approximately 22 per 1,000 in 1960. These data are consistent with historical accounts which assign the highest mortality rates to January and February of 1960. Mortality rates return to trend in 1962.

Of the 27 provinces in our sample, each has approximately 29 million people in a given year, approximately 86% of provinces experienced higher mortality rates in 1960 than in previous years (1954-57). When one compares 1954-57 and 1959, one sees that average mortality essentially doubled in the latter year from 9 to 22 deaths per 1,000 people. At this same time, agricultural production fell. In 1959, average per capita production was approximately 254 kg, only 83% of the sample average during 1954-57. Additionally, 96% of the provinces experienced a fall in production in 1959 relative to 1954-57. Xinjiang, an autonomous province in the northwestern corner of China, was the only province that did not experience such a fall in production. See Appendix Table A1 Panel A for descriptive statistics of the province-level data.

The descriptive statistics provide several additional facts about the fall in production in 1959

surprising as the NBS was being constructed at the time and the amount of statistics it was able to gather was increasing over these early years. We are particularly careful to check that the selection of provinces do not affect our results by repeating all of our estimates on a restricted sample where the years 1949-1953 are omitted for all provinces.

[28]For the years 1960 and 1961, such understatement would cause us to overestimate population food need and overestimate any aggregate food shortages. Therefore, it would bias against our finding that there was no shortage.

[29]For years after the famine began, this becomes increasingly inaccurate, especially since the famine killed disproportionately more young children and elderly. See Coale (1981) for details on the quality, collection, and subsequent corrections of this data. Assuming that each province has the same age and sex distribution as the 1954 national average may cause us to over-state caloric needs for provinces that were experiencing positive growth in fertility or life expectancy as these provinces will have a larger share of young children and elderly individuals, who require less calories. By the same logic, this assumption may cause us to understate population food needs that are experiencing declines in fertility and life expectancy.

that are important for developing our model later in the paper. First, per capita production in 1959 was greater for regions that produced more in the past; the correlation coefficient between log production per capita in 1959 and log production per capita in 1958 is 0.97. Second, the fall in production was broadly uniform across regions when measured as a proportion to past regional production. This means that the spatial patterns of famine cannot be due to larger percentage drops in production in regions that historically produced more.[30] If anything, the data suggest that regions with higher past production experienced slightly smaller drops in production.[31] Nevertheless, the percentage drops in production across regions were similar enough that historically higher producing regions experienced larger absolute falls in per capita production in 1959.

3.2 Retrospective Data

In the supplementary analysis, we use retrospectively calculated regional birth cohort size as a proxy for famine severity and two alternative proxies for production: a time-invariant measure of suitability for grain production based on geography and climate and a time-varying measure of historical weather conditions.

For famine severity, we use birth cohort size from the *1990 Population Census*. A smaller birth cohort size during famine years reflects a more intense famine. This can be seen from the high degree of correlation between birth cohort size and mortality rates across provinces. For 1959-60, this correlation is -0.65.[32] The key advantage of this alternative measure is that it is not subject to systematic mis-reporting by the government. Additionally, this proxy allows us to disaggregate the data to the county-level and capture much more of the variation in famine. Furthermore, we can split the data into agricultural and non-agricultural households. While non-agricultural households typically live in cities and towns, they are sometimes in the same county as rural households, which typically live in surrounding villages. Distinguishing

[30] One could suspect that regions that produced more on average were under more pressure to increase production and thereby were more likely to resort to cultivation methods that increased production in the short run but hurt production in the long run. Alternatively, one could wonder whether GLF policies diverted more labor away from agriculture in areas that were historically more productive and therefore caused larger percentage falls in these regions. The data show that this is not the case.

[31] The correlation between the percentage change in per capita grain production between 1958 and 1959 and log per capita production in 1958 is positive, but small in magnitude and statistically insignificant (i.e., the correlation is 0.13).

[32] To visually illustrate this correlation, we aggregate cohort size to the province level and plot province level mortality rates (Figure A2A) and birth cohort sizes (Figure A2B) over time for each province. A comparison shows that for each province that exhibits a spike in mortality in 1960, there is a corresponding drop in cohort size. Moreover, there are many provinces which do not show an increase in mortality but which show a drop in fertility (e.g., Guangdong). This is consistent with the belief that cohort size is a more sensitive measure than mortality (e.g., when a famine occurs, victims will forego fertility before allowing a living person to starve to death).

Note that mortality and survival rates are both proxies for famine severity. One cannot retrospectively back out true mortality rates from the cohort size data. This is because we do not know the fertility rates in the years leading up to the famine, and more importantly, we cannot observe those who were elderly at the time of the famine in the retrospective data.

these two household types is interesting because they faced very different policies in the Chinese grain distribution system, which procured grain from agricultural households and provided non-agricultural households with food subsidies. These categories were first assigned during the 1950s and transitioning from one to the other is difficult. A third advantage of our proxy is that birth cohort size is likely a more sensitive measure of famine than mortality, since famine victims presumably forego fertility before allowing living family members to die of starvation. Finally, famine birth cohort size data exists for all thirty provinces, allowing us to expand the geographic scope of our analysis to cover the entire country. See the Appendix for more discussion on the birth cohort size data.

Our main proxy for historical production at the county-level is local suitability for grain cultivation as assessed by an agro-climatic model constructed by the Food and Agricultural Organization (FAO). This measure of suitability is based purely on the biophysical environment of a region and is not influenced by which crops are actually adopted in an area. Factors that are easily affected by human actions, such as soil pH, are not parameters in this model. The suitability measure at the county-level is the fraction of grids within a county that is suitable, a measure we use for the sake of computational ease. See the Appendix for more details.

After matching cohort size data and suitability data at the county level, we form two county-level samples, which are each a balanced panel of birth cohorts for 18 birth years (1949-66). One sample is comprised of agricultural households and the other is comprised of non-agricultural households. We set the end date of both samples to 1966 to avoid confounding factors that could have influenced fertility when the Cultural Revolution began.[33]

Agricultural households have approximately 5,266 individuals per cohort on average in each county. Famine cohorts are smaller on average, comprising of approximately 3,508 individuals, 71% of the size of cohorts born prior to the famine (1954-57). If we use this ratio being below one to indicate whether there was a famine, then 89% of the counties in our sample experienced the famine to some extent. On average, 13% of the land in counties with agricultural households is suitable for cultivation of rice and wheat according to our definition. The patterns for non-agricultural households are similar. However, the difference between famine cohort sizes and pre-famine cohort sizes is smaller in non-agricultural households than for agricultural households. Also, non-agricultural households typically lived in agriculturally richer areas, where 20% of the land is suitable for cultivation. (See Appendix Table A1 Panel B for descriptive statistics in the retrospective county-level sample).

To observe aggregate cohort size over time, we aggregate county-level data to the national level and plot the number of people living from each birth year for agricultural and non-

[33]Fertility is a much more sensitive measure than mortality because parents are more likely to forgo having children before allowing living members to die. This is relevant during the late 1960s and 1970s in China as there were many events such as the Cultural Revolution or the beginning of family planning policies (in the early 1970s) which may have delayed fertility but not have had large affects on mortality. Our results are robust to the inclusion of later birth cohorts. These results are not presented for the sake of brevity. They are available upon request.

agricultural households in Figure 2. Both agricultural and non-agricultural households experienced a decrease in cohort size close to the famine years, though the drop is much more dramatic for the agricultural population. The drop in cohort size for those born immediately before the famine reflects the mortality of young children during the crisis. The more severe drop for the famine birth cohort (1959-61) reflects an increase in infant mortality and a dramatic decrease in fertility. In the figure, we plot a projected linear trend for agricultural households and show that there is a positive linear trend in cohort sizes for those born in the mid-1950s and before. For the years right before and during the famine, a time period indicated by the vertical lines, cohort sizes are well below trend, though they return to trend after 1961 when the famine ends.[34]

Using suitability for grain cultivation to proxy for grain production has both advantages and disadvantages. The main advantage is that the meausure is not subject to government reporting error. The key disadvantage is that suitability may not be positively correlated with actual production every year. However, the high autocorrelation in province level production suggests that this problem is unlikely to be a big concern.

To address this issue quantitatively, we use historical weather conditions as an alternative time-varying proxy for historical production. The historical weather data is reported by scientific weather stations of the People's Republic of China and includes monthly precipitation and temperature for each station. These data have only recently become available to researchers, and to the best of our knowledge, it was collected and solely used by government meteorologists. Accordingly, there are no accounts of manipulation for these data. Because the sample for weather data is restricted in size, we focus on the results using suitability for grain cultivation and refer to the results from using weather data mainly as a robustness check. See the Appendix for more details on these data.

4 Empirical Results

This section presents two empirical findings. First, we find that food production in 1959 was significantly above what was needed to avert famine-induced mortality. Second, we find that rural regions that produced more food per capita in 1959 suffered higher mortality during the famine, reversing the normally negative correlation between per capita food production and mortality.

4.1 National Grain Production and Subsistence in 1959

In this section, we discuss our first finding that food production in 1959 exceeded per capita subsistence needs by comparing the historical estimates of national food production to two benchmarks for caloric needs. The two benchmarks distinguish between the caloric requirements for preventing a decrease in labor productivity from the requirements for preventing mortality. Be-

[34] Using the ratio of famine cohort size to pre- or post-famine cohort size as a proxy for famine severity, we find that there is substantial cross-sectional variation in famine intensity. This can be seen in a histogram of the ratio of famine cohort size to pre-famine cohort size for agricultural populations in Appendix Figure A3.

cause the majority of famine mortality occurred in the winter following the harvest of 1959, we focus on the level of production in 1959. We do not consider the additional fall in production in 1960 which occurred after a large proportion of the rural workers were already starving or dead. While we examine these two benchmarks for interest, the lower benchmark is more relevant for our study because we are interested in why people died in the winter of 1959-60, when the production from 1959 had already been harvested (i.e., people did not need to work to produce more food to stay alive). Note that *a priori*, it would not be surprising to find that production in 1959 did not fall below the needs for preventing mortality, since aggregate production per capita in 1959 only dropped to the same level as in 1949-1951, a period in which there was no famine.[35]

To calculate population caloric needs, we use China's 1954 population age and sex distribution as reported by Coale (1981). Caloric requirements for working and healthy child development are calculated from a model published by the United States Department of Agriculture (USDA). Our calculation shows that the average per capita population caloric need for productive agricultural laborers and normal child development is 1,871 calories per day and that the average need to stay alive is approximately 804 calories per day. The details behind this calculation are described in the Appendix.

To calculate the total population's caloric needs in terms of grain, we use the calculated per capita caloric needs together with the annual sum of historical population for the 27 provinces of our sample. The Ministry of Health and Hygiene of China estimates that one kilogram of grain in the mix and form as consumed by the average Chinese worker provides 3,587 calories. We assume that individuals subsist solely on grain and that each individual consumes the same amount every day of the year. This leads to the conclusion that, with a diet of 1,870 calories per day, per capita grain needs are 190 ($365 \times \frac{1870}{3587}$) kg per year. Using this figure, we determine the needs of the population and calculate the deficit in production as the difference between production and need. We repeat the same exercise for the lower benchmark of caloric need and find that 82 ($365 \times \frac{804}{3587}$) kg of grain per capita would have prevented mortality.

The estimated population food need and food production over time from our sample of 27 provinces are plotted in Figure 3 (and shown in Appendix Table A3). In 1959, production was far greater than the baseline subsistence level. Specifically, there was a 99 million ton surplus above the 48 million tons required. Production was in surplus even relative to the higher benchmark of food required for high labor productivity; production exceeded this level by 36 million tons.

Figure 3 illustrates several interesting facts. First, the fall in production by itself could not have caused the famine in 1959. Production was almost three times as much as what was needed to keep the entire population alive. While we do not interpret these calculated surplus levels literally, it is important to note that the magnitude of the surpluses suggests that it is highly unlikely for over-reporting of production to undermine our argument that there was no

[35]When comparing the caloric needs between our two benchmarks, one should also note that the relationship between calories consumed and work capacity is potentially highly non-linear, as suggested by the nutritonal poverty trap theory (Das Gupta and Ray, 1986). This does not play an important role for our study since we will argue that production in 1959 was in excess of both benchmarks.

aggregate deficit in food. One would need to be believe that our data, the most conservative data available, still exaggerates production by 205% ($\frac{99}{48} \times 100$) to think that aggregate production was too little to prevent mortality in 1959. This exercise also shows that the fall in production in 1959 was the largest drop experienced by the new government since coming to power in 1949. This observation is consistent with historical accounts of the general expectation of continued growth in food production and the widely held belief that the production fall in 1959 was largely unanticipated by central and local leaders.

To be conservative, we intentionally construct our calculations for caloric needs to bias us against finding that food production was sufficient, and we do so in three ways. First, we assume that the entire adult population worked as laborers in agriculture, whereas in reality, approximately 20% of the population in 1959 worked in less physically intensive non-agricultural jobs. Second, we proxy for the number of calories needed to stay alive with the conventional benchmark for the number of calories needed to do"some" work (e.g., see Das Gupta and Ray, 1986). We use this measure because there is little consensus on the minimum caloric consumption necessary for staying alive. Finally, post-World War II fertility rates were very high, so a larger fraction of the population consisted of children in 1959 relative to 1954. Since children need fewer calories than adults, we will overstate caloric needs. See the Appendix for more details on the calculation.[36] The fact that we are overstating caloric needs is consistent with the fact that Figure 3 shows that 1949-50 food production as slightly below population needs according to the higher benchmark, but there are no accounts of famine or starvation in those years.

In interpreting these results, several concerns should be considered. First, one could raise the concern of whether the national level is the appropriate level of aggregation. China is a large country in terms of geographic size. If food is typically redistributed regionally, it may be more reasonable to think of the province as the appropriate level of aggregation. Therefore, we repeat this exercise for each province. We find that the result holds for the province level. All provinces produced more than what was needed to avoid mortality, and only the three province-level municipalities produced less than what was needed for a healthy labor force and normal child development. Ironically, the mortality rates for these regions were amongst the lowest during the famine. (See the Appendix for more discussion and Appendix Table A4 for province level production and caloric needs in 1959).[37] Second, one may worry that our estimates of average

[36]In addition, the aggregate statistics could also overstate food deficit in 1959 because Sichuan, Hainan and Tibet are omitted from our sample of 27 provinces. The omission of Sichuan is likely to cause us to understate per capita production and overstate the food deficit. On average, Sichuan is the largest grain producer in China and the 6th largest in per capita terms. On average, this one province accounts for more than 9% of national production. (Tibet and Hainan typically produce just enough food for their own subsistence. Therefore, their omission should not have significant effects on our estimates).

[37]For this result, it is important to recall that by assuming that all provinces have the same demographic composition, we overestimate food needs in provinces with positive growth in fertility and life expectancy; the higher the growth, the more we overestimate. We can reasonably infer that all provinces in our sample experienced positive growth in fertility and life expectancy from the fact that our data shows that all provinces experienced positive growth in population before 1959. This means that we are overestimating food needs for the national population.

caloric needs understate the caloric requirement during winter conditions, when the human body expends more energy to combat exposure to cold temperatures. However, peasants typically work less during the winter, in between fall harvest and spring sowing, and related evidence from nutritional studies on caloric requirements during winter conditions (e.g., Milan and Rodahl, 1961) all suggest that our result that enough food was produced to prevent mortality in 1959 is unlikely to be undermined by the additional caloric requirements necessary to survive winter conditions.[38]

In summary, the estimates presented in this section show that China produced more than enough food to avoid famine deaths in 1959 at the national and at the provincial level.

4.2 Regional Grain Production and Famine in 1959

4.2.1 Historical Data

To investigate the cross-sectional relationship between per capita production and mortality rates more systematically, we estimate the cross-sectional correlation between the per capita production in province p and year t and the mortality rate in that province the following year. To allow the effect to differ for the famine years, we introduce the interaction term of grain production and a dummy variable for 1959, when aggregate production fell. Note that because mortality is recorded annually and the majority of deaths are believed to have occurred during the early months of 1960, we estimate the relationship between production in a given year and mortality in the following year. For this exercise, we transform the reported mortality rate data into the number of deaths in province p and year t by multiplying it by total provincial population in province p and year t.[39] The estimated equation is:

[38] There is no systematic evidence on caloric requirements for staying alive, not to mention evidence on caloric requirements across different physical environments. However, based on related evidence, we believe that winter conditions are unlikely to have caused caloric requirements to be so high as to create an aggregate food deficit in 1959. For example, in the Minnesota Starvation Experiments, where adult males were subject to mild starvation and given diets of approximately 1,560 calories per day for six months and subject to harsh temperature drops (to simulate war conditions), there were no deaths due to the temperature drops (Keys, 1950). Our benchmark for caloric need for heavy labor is actually 17% above the caloric provision in this experiment, which suggests that our estimates are probably conservative even taking winter conditions into account. In a study of caloric needs for scientists working in Antarctica (Milan and Rodahl, 1961), it was found that a 30% increase in caloric intake was more than enough for the winter months which dropped to -70 degrees Fahrenheit (i.e., the scientists performed the same amount of labor in the winter and gained weight with this diet). From our estimates, one can clearly see that a 30% increase from the lower benchmark for caloric need for staying alive is still far from grain production levels in 1959. Moreover, winters in China are mild relative to Antarctica. For example, in Anhui, a province that experienced one of the highest famine mortality rates, average winter temperatures are around 59 degrees Fahrenheit.

[39] Official documentation of the CSDM50 states that mortality rate is calculated as the total number of deaths in a year measured at the end of the year divided by the population as measured during the middle of the year. This suggests that the correct measure of death is $Death_t = MortalityRate_t \times TotPop_t$. However, one may reasonably doubt the Chinese government's ability to collect population data during the middle of 1960, and think that instead they may have used the population from 1959 as the denominator for calculating mortality rates such that $Death_t = MortalityRate_t \times TotPop_{t-1}$. The results presented in this paper are from using the latter method. However, the results are similar between the two different methods of normalization. Those from using the first method are reported in the Appendix Table A5.

$$\ln Deaths_{p,t+1} = \alpha(\ln Grain_{p,t} \times FamineDummy_t) + \beta \ln Grain_{p,t} + \gamma \ln TotPop_{p,t} + \delta_t + \varepsilon_{p,t} \quad (1)$$

The log of the number of deaths in province p and year $t+1$ is a function of: the interaction between log grain production in province p in year t, $\ln Grain_{pt}$, and a dummy variable for the year 1959, $FamineDummy_t$; the main effect for log grain production; the log total population of a province p in year t, $\ln TotPop_{pt}$; and, year fixed effects, δ_t. More populous provinces will naturally have more deaths and more grain production, and controlling for total population reflects the fact that we care about per capita grain production and deaths. This is a more flexible way of normalizing than dividing each measure with the total population.[40] In this estimation, β measures the average correlation between per capita grain production and mortality rates for 1949-1958 and 1960-1997. If grain production is negatively correlated with death, then $\hat{\beta} < 0$. α reflects the marginal correlation in 1959. The average relationship between grain production and mortality in 1959 is the sum of the two coefficients, $\alpha + \beta$.

The estimated results are presented in Table 2. For each coefficient, we present the unadjusted and robust standard errors. The small number of provinces means that we cannot cluster the standard errors at the province level without being vulnerable to small sample bias. Column (1) shows the estimates using all 27 provinces of our sample. It shows that the correlation between grain production and mortality is negative, but not statistically significant. However, the interaction term for grain production and the 1959 dummy variable is positive and statistically significant at the 1% level. In column (2), we omit the autonomous regions of Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai, as these regions had a much larger share of ethnic minorities and were often subject to different policies from the provinces which contained almost exclusively Han Chinese.[41] When we exclude these provinces, the results become more precisely estimated. The interaction term between grain production and the 1959 dummy remains large, positive, and statistically significant at the 1% level. The main effect of grain production is negative and statistically significant at the 1% level. The coefficients show that the average elasticity between per capita grain production and mortality rate is approximately -0.06 in normal years, but in 1959, the elasticity is increased by 0.256. To see the total effect in 1959 and its statistical significance, we sum these two coefficients. The sum and the p-value using the two different methods of computing standard errors are reported at the bottom of the table. Here, we see that

[40]Different ways of normalization do not change the results. For brevity, we only report the more flexible method in this paper. The alternative results are available upon request.

[41]For example, in Tibet, there was a large uprising against the Chinese government in 1958. Subsequently, the Chinese government imposed a particular set of policies for the province of Tibet and all Tibetan autonomous regions in the neighboring provinces of Qinghai, Sichuan, and Gansu. Generally, autonomous regions face different fiscal policies (e.g., food and revenue redistribution) and are organized differently from the central government. In later years, ethnic minorities also faced different fertility control policies from the Han Chinese population.

the correlation is reversed in 1959. A 1% increase in grain production in 1959 is correlated with a 0.194% increase in mortality. The estimated p-values show that this joint statistic is highly statistically significant.

In column (3), we add a control for government expenditure on health and education, which we use to proxy for public goods expenditure.[42] This controls for potential confounding factors such as the way in which regional governments responded to the famine. For example, governments in regions that produce less food on average may be more prepared for famine and therefore respond to famine with higher spending on public health. If this is the case, then the positive correlation between production and mortality in 1959 would reflect differences in public expenditures as opposed to differences in food consumption. The results in column (3) suggest that this is probably not the case, as our estimates are very robust to this additional control.

Next, we control for province-specific time trends, which addresses the concern that more productive regions may have experienced different mortality trends from regions that produce less on average. For example, more productive regions may have invested more in public health, which could lower mortality rates on average and exaggerate the famine versus non-famine difference in mortality rates. Column (4) shows that our estimates are robust to this rigorous control.

Finally, we address the problem that we lack mortality data for many provinces before 1954. Our sample is only a balanced panel of 27 provinces from 1954-98. To check that the selection of the provinces that reported data in the previous years do not drive our results, we re-estimate the same equation on a restricted sample to the years 1954-98. The estimates are very robust.

One concern about the estimates from equation (1) is that we are not capturing differences that are specific to the famine years. For example, one could worry that the relationship between grain production and mortality changes over time for completely spurious reasons. Alternatively, the positive correlation we detect in 1959 could simply be capturing GLF policies, which were in place during 1957-60. While we cannot completely rule out the possibility that spurious changes occurred, we can examine the data more carefully to see whether the relationship between grain production and mortality is positive in other years. We estimate the correlation between grain production and mortality for each year using the following equation:

$$\ln Deaths_{p,t+1} = \sum_{\tau=1949}^{1997} \alpha_\tau (\ln Grain_{p,\tau} \times YearDummy_{\tau=t}) + \gamma \ln TotPop_{p,t} + \delta_t + \varepsilon_{p,t}, \quad (2)$$

where $YearDummy_{\tau=t}$ equals 1 if $\tau = t$ and equals 0 otherwise. This specification is similar to equation (1). The difference is that we now interact grain production with a dummy variable for each year, and because we have an interaction term for each year in the sample, we drop the main effect of grain production. Therefore, the average correlation between grain produced in a province and mortality is α_τ for each year τ. As before, we normalize by total population by controlling for that term on the right hand side. To increase precision, we use the restricted sample where we omit autonomous regions. The vector of estimated $\hat{\alpha}$'s and their 95% confidence

[42] These data are reported in the CSDM50.

intervals are plotted in Figure 4A. They are shown in Appendix Table A6. As with the earlier estimates, we present both unadjusted and robust standard errors. The figure shows clearly that relative to other years, the relationship between grain production and mortality spikes upwards in 1959. The time path of the correlation between grain production and mortality looks extremely similar to the time path of the raw data on mortality rates plotted in Figure 1. This means that the finding from Table 2 that the relationship between grain production and mortality changes in 1959 is specific to that year and not confounded by spurious changes unrelated to the production fall in 1959. Similarly, it is equally unlikely that the sharp positive relationship between production and mortality we observe in 1959 is simply an outcome of general GLF (1958-61) policies, which were in place for at least three years other than 1959.

In summary, the historical data illustrate a stark cross-sectional pattern between famine severity and food production in 1959: the normally negative correlation reverses to be positive in 1959. There are several caveats to interpreting this correlation. First, there are the concerns over data quality which we discussed in Section 3. Second, the results in this section do not identify whether the patterns are solely driven by urban-rural differences or if they also exist across rural households. Third, the province-level results could reflect differences in regional responses to the production shortfall in 1959. We address these issues in the next section.

4.2.2 Retrospective Data

In this section, we conduct a similar empirical exercise with retrospective measures of famine severity, using regional birth cohort size from the 1990 Population Census and two proxies for grain production: the regional suitability for cultivating rice or wheat as predicted by time-invariant natural conditions, and annual regional spring precipitation. This supplementary analysis is a robustness check for the historic province-level analysis in that it allows us to address the caveats associated with interpreting the results from the previous section.

We estimate the cross-sectional correlation between log birth cohort size and suitability for grain production using the following equation:

$$\ln CohortSize_{ct} = \alpha(Suitability_{ct} \times FamineDummy_t) + \beta Suitability_{ct} + \quad (3)$$

$$\gamma Avg \ln CohortSize_c + \delta_t + \varepsilon_{ct} \quad (4)$$

Log birth cohort size for birth county c in birth year t is a function of: the interaction term between the predicted suitability for county c for cultivating grain, $Suitability_c$, and a dummy variable indicating whether a cohort was born during 1959-61, $FamineDummy_t$; the main effect of suitability for grain cultivation; the average of the log of cohort size for each county c; and birth year fixed effects, δ_t. This equation is conceptually similar to equation (1), except that we use log birth cohort size as the dependent variable, have a time-invariant proxy for production, and control for the average of the log of each county's birth cohort size on the right hand side. Note that there is no historical data for county-level population for most counties in China. Hence, controlling for average log county birth cohort size is a proxy for controlling for county population, which allows us to interpret the estimated coefficients in per capita terms.[43] The standard errors are clustered at the county level.

The average correlation between suitability and birth cohort size is reflected by β. One would expect that more suitable regions would typically have larger birth cohort sizes (from higher fertility rates or child survival rates) such that $\beta > 0$. α is the marginal correlation between suitability and cohort size for those born during the famine such that the average correlation between suitability and cohort size for the famine cohorts is reflected by the sum of the two coefficients, $\alpha + \beta$.

Table 3 columns (1)-(4) present the estimates for agricultural households. Column (1) uses a full sample of all provinces. It shows that grain suitability is on average positively correlated with cohort size. The elasticity between per capita land suitable for grain production and per capita birth cohort size is approximately 0.048 for birth cohorts on average. The estimate for the interaction term shows the marginal correlation for those born during famine years; it is negative. Both of the estimates are statistically significant at the 1% level. As with the historical estimates, the average correlation between grain suitability and cohort size for those born during the famine is the sum of the two coefficients presented at the bottom of the table. It is negative and statistically significant. For those born during the famine, the usual positive relationship between suitability for grain cultivation and cohort size is reversed such that the elasticity between per capita land suitable for grain production and per capita birth cohort size is -0.21. In columns (2)-(4), we omit the autonomous provinces Tibet, Xinjiang, Neimeng, Ningxia, Guangxi and Qinghai. Column (2) shows that this changes our estimates very little. In column (3), we add province-specific time trends as additional controls to address the possibility that fertility or child mortality trends may have differed between provinces. The estimates are again very robust. In column (4), we add the most rigorous set of controls, province-year-specific fixed effects. This controls for any changes across provinces and over time in a fully flexible manner. For example, if certain provincial governments implemented particular policies in response to the famine, this regression controls for these policies where we compare the differences across counties within years. As before, the estimates change little in magnitude. All of the estimates presented in columns (2)-(4) are statistically significant at the 1% level.

Next, we repeat the same estimation on a sample of non-agricultural households. The non-agricultural sample is smaller because there are many counties that contain only agricultural households. The estimated coefficients are much smaller in magnitude and statistically insignificant. Towards the bottom of the table, we see that the sum of the coefficients are all very small in magnitude and statistically insignificant. These estimates suggest that on average, the cohort

[43] The fact that the control is averaged over birth years and time invariant for each county means that the coefficients of interest are identical to those in a specification which controls for county fixed effects.

sizes of urban households are not correlated with local grain production.

For consistency with the historical analysis, we estimate the correlation between grain suitability and cohort size for each birth year. This exercise is analogous to the yearly estimates using the historical data in equation (2). It can be written as:

$$\ln CohortSize_{ct} = \sum_{\tau=1949}^{1966} \alpha_\tau \left(Suitability_\tau \times YearDummy_{\tau=t}\right) + \gamma Avg \ln CohortSize_c + \delta_t + \varepsilon_{ct}, \quad (5)$$

where $YearDummy_{\tau=t}$ equals 1 if $\tau = t$ and equals 0 otherwise. The log cohort size of individuals born in year t in county c is a function of: the interaction terms between the fraction of land that is suitable for grain production in county c and dummy variables for each year; the average of log birth cohort size in a county and birth year fixed effects. All standard errors are clustered at the county level. The inclusion of year fixed effects controls for secular changes in fertility and mortality that may affect cohort sizes. The inclusion of average log cohort size normalizes the outcome variable and suitability by population and is identical to controlling for county fixed effects.

For brevity, we focus our discussion on the estimates for agricultural households, which are presented in Appendix Table A7 column (1). They and their 95% confidence intervals are plotted in Figure 4B. This figure demonstrates that relative to cohorts born far before the famine, cohort sizes for those born right before the famine (and were very young when the famine began) and those born during the famine are negatively correlated with grain suitability. Interestingly, the pattern in the correlation between suitability and cohort size across birth years is very similar to the pattern of birth cohort sizes across birth years that we observe in the raw data in Figure 2.

This retrospective analysis supports the earlier historical analysis. Under the assumption that regions which produce more grain on average also produced more in 1959, the findings show that there is a reversal in the relationship between per capita production and survival rate for the famine birth cohort. However, if 1959 grain production was actually lower for counties that produce more on average relative to those that produce less, then the retrospective analysis using the grain suitability proxy becomes difficult to interpret. Since the historical province level data show a very high degree of autocorrelation between per capita production in 1958 and 1959, our prior is that there is probably also high autocorrelation in county level production in those years. However, without any historical county level production data, we need an alternative method of verifying this assumption.

We address this difficulty by re-estimating equation (3) using an alternative time-varying proxy for food production: the log of average precipitation during the spring months. Higher precipitation during the spring typically results in higher production.

The estimated effects for the agricultural population sample are presented in Table 4. First, we show the estimated effects using suitability as a benchmark for comparison. Because weather data is only available for a subsample of counties, we re-estimate the effect of the proxy for grain suitability on the subsample for which we have historical weather information. The estimates shown in column (1) are less precise than those for the full sample, but still show the same patterns. In column (2), we present the estimates of (3) which replace grain suitability with log average spring precipitation. The estimates show that this time-varying proxy for local production produces the same pattern. During normal years, spring precipitation is positively correlated with birth cohort size: the elasticity between rainfall and birth cohort size is 0.019. During famine years, it is negatively correlated. The elasticity, which is the sum of the main effect of precipitation and the interaction effect, is -0.077. Both coefficients and the sum of the two are statistically significant at the 1% level. In column (3), we control for mean spring temperature and its interaction with the famine birth years dummy. The estimates are virtually unchanged. Note that the historical weather data does not form a balanced panel because the number of weather stations increased rapidly during this period. To check that this does not bias our estimates, we restrict the sample to counties for which weather data exist for at least twelve of the eighteen years of our sample. Although less precisely estimated, the coefficients remain similar.

For consistency, we also estimate the yearly correlation between spring precipitation and birth cohort size by estimating equation (5) with the interaction terms of spring precipitation and year dummies as explanatory variables replacing the interaction terms of grain suitability and year dummies. For this estimate, we exclude the years 1949 and 1950 because there were very few counties with weather stations in those early years. The coefficients and their 95% confidence intervals are plotted in Figure 4C and reported in Appendix Table A7 column (2). The patterns we observe are similar to the grain suitability estimates but are less precisely estimated.[44] For the sake of brevity, we only report the results for agricultural households.[45]

In summary, the results from the retrospective analysis support the main findings using historical data and show that the relationship between per capita food production and birth cohort size reversed during the famine. The retrospective analysis also provides several additional insights. First, it shows that the cross-sectional patterns in famine severity are primarily driven by agricultural households. Second, it shows that the reversal in the correlation between production and survival is not affected by controls for time-varying differences across provinces. This fact is important because it means that the reversal exists even allowing for the possibility that each provincial government enacted different policies in response to the aggregate production fall in 1959. Third, the results incorporating precipitation data show that the retrospective estimates from using a time-invariant proxy for production are not driven by the possibility that counties that are more productive on average were less productive in 1959. Finally, and most

[44]In theory, we can also aggregate the weather data to the province level and repeat the estimates with mortality rates. However, this will result in extremely imprecise estimates in practice since for many years, there are only one or two weather stations in a province.

[45]When we repeat these estimates on the sample of non-agricultural population, we find that weather does not systematically affect birth cohort sizes. The estimates are less than half the magnitude as than those for agricultural households and statistically insignificant. As with the estimates for agricultural households, these results are consistent with those from our main analysis. They are available upon request.

importantly, the fact that we find the same spatial patterns between historical weather and birth cohort size as between historical production and mortality rates is additional evidence that the latter relationship is not entirely driven by government mis-reporting of the historical production and mortality data.

5 The Grain Procurement System

The previous section provides evidence for two findings that identify grain procurement policy as a driving force of the famine. The first finding that food production in 1959 was sufficient for subsistence implies that the famine could not have been solely caused by the drop in food production.[46] Two important facts follow from this result: 1) the determinants of the fall in production – whatever they were – cannot by themselves explain the famine; and 2) the allocation of food across China must have played a role in causing the famine. This result begs the question: what transformed the fall in production into a famine?

Our second empirical finding is that the correlation between regional well-being and production reversed during the famine such that regional famine severity was increasing with per capita production. Together with the fact that China was a centrally-planned non-market economy, this observation implies that conventional market mechanisms which generate unequal food distribution cannot explain the Chinese famine. Instead, we must consider the role of government food distribution policy in transforming the production fall into a famine.

In the following discussion, we document the details of China's food procurement regime and the political climate in order to better understand the mechanism for this policy and motivate the model in the next section. In particular, we highlight historical evidence that suggests that a key feature of the centrally planned procurement system was inflexibility due to difficulties in aggregating and responding to information. Moreover, the evidence suggests that inflexibility was likely a response to the lack of local incentives to truthfully report production and a consequence of political pressures to follow rules and the limited bureaucratic capacity of the central government.

Grain procurement was planned centrally. The central government decided the production targets each year. These made their way down to regional government officials, who then traveled to collectives each spring to announce the expected production (e.g., production targets) for that collective. Procurement typically took place after the fall harvest around November. The central government's method for determining procurement levels are outlined in initiatives such as the "Three Fix Policy". In 1956, this policy stipulated that to "fix" procurement levels for each collective, expected local production levels in 1956 should be based on production in 1955, and subsistence levels of consumption and seed retention should be based on population and

[46] This is perhaps not surprising since according to the aggregate data reported by Li and Yang (2005), the average worker retained approximately 193 kg of grain *after procurement* in 1959, which is more that what is needed to stay productive and much more than what is needed to stay alive (see Appendix Table A8).

production needs.[47]

The main reason for setting procurement based on expected production is that peasants and local officials were not incentivized to report actual production truthfully. Discussions amongst the top party leadership show that they were well aware that peasants had an incentive to under-report production in order to retain a larger amount of grain and to prevent the government from demanding greater production in the future. Local leaders were similarly unreliable as they could share the incentives of the peasants or over-report production for political advantage (Fairbank, 1986: pp. 305-8). Given the lack of reliability in the reported information, it was logical for the government to condition procurement on historic information.

To a large extent, bureaucrats seem to have followed the prescribed procurement rules, especially in 1958 and 1959, when political tensions were intensified between Mao and members of the Politburo who did not support his GLF policies. At the height of tensions, Mao purged a significant proportion of moderate political leaders from the upper and middle levels of government, creating an environment where few were willing to report that production in 1959 was lower than expected before the numbers could be aggregated and presented to Mao in an impersonal manner.[48] In other words, the rules, together with the political pressure to follow them led to a very inflexible policy.[49]

Before moving the discussion forward, it is interesting to note that along with political pres-

[47] See Johnson (1998) for a discussion of the food procurement system. Historical grain policies are outlined in public government archives. See http://2006.panjin.gov.cn/site/gb/pj/pjjz/pjjz_detail.php?column_id=2382.

[48] The political climate in 1959 was extremely tense and most likely caused leaders to follow rules, even those that were likely to prove problematic later. The GLF had been received with cynicism from the very beginning, and its failures and successes were crucial to Mao's political leadership. In December 1958, at a meeting of the Central Committee of the CCP in Wuhan, party leaders refused to fully endorse GLF policies. Following this meeting, Liu Shaoqi replaced Mao, who remained Party Chairman, as the Head of State in early Spring of 1959 (Spence, 1991: pp. 581). Many historians view this as an unwilling step down by Mao. It is therefore not surprising that further challenges of the GLF resulted in a strong response from Mao. In July 1959, Mao famously purged Peng Dehuai, a field marshal of extremely high political standing, for criticizing collectivization and other GLF policies and expressing forebodings of famine. These problems of the collective system mandated by the GLF were a source of contention between communist party moderates and hard-liners who backed Mao. However, with the exception of Peng Dehuai who did a tour of the countryside during the spring of 1959, there is no evidence that any top leader ever obtained an accurate picture of the problems of collectivization and the danger of famine. Peng discretely reported these problems to Communist Party Chairman Mao Zedong in a personal letter. The problems he mentioned included reduced incentives to work, a diversion of labor away from agriculture, and over-procurement of grain by mid-level party leaders who were under-pressure to fulfill grain target quotas that had been set too high. Fearing a political revolt against his leadership based on perceived failures of the GLF, Mao used the contents of this letter to purge Peng as a rightist at the historic Lushan conference in July of 1959. Peng was put under house arrest and later executed during the Cultural Revolution. At this conference, the top party leaders made clear that the first year of the GLF was a success and that collectivization was increasing grain harvest more than ever (Becker, 1996, pp. 87-92). The Lushan conference had important consequences. The removal of Peng was accompanied by a purge of his supporters amongst top party members as well as moderate mid-level party leaders who had expressed concerns about collectivization and the dangers of famine (Fairbank, 1986: pp. 303-335; Becker, 1996, pp. 93). It put remaining leaders under enormous pressure to deliver the high targets for grain quotas for the harvest of 1959 in order to not be grouped with the critics of Mao (Spence, 1991: pp. 574-583).

[49] The rigidity of rules and how it caused officials to sacrifice efficiency can be observed in food delivery even in the post-Mao era. Oi (1989) documents that local leaders punctually put harvests by the roadside for pick up even in bad weather causing huge losses sometimes. Presumably, these leaders knew that they would be punished for the lack of punctual delivery but not for bad weather induced losses.

sure, an additional reason for why bureaucrats were willing to deliver large quantities of grain to central procurers is the implicit promise the new government made that no one would ever be allowed to starve. The party often referred to this promise as "the iron rice bowl". This implicit social contract may have contributed to the belief on the part of local leaders that once people began to starve, they would be given grain replenishments by the central government. This attitude is consistent with accounts of collective kitchens providing large quantities of food even after procurement.[50] This optimism was probably also encouraged by the fact that in 1959 production was remembered as plentiful and well above subsistence needs. In this case, a reasonable response for local leaders is to give the government the planned amount of procurement and postpone their plea for grain (and potential punishment) from the central government.[51]

An additional source of inflexibility that is entirely independent of incentive issues is the combination of limited bureaucratic capacity and political centralization. The Standing Committee of around seven Politburo members was the only government organ with the power to make major policy decisions. It attempted to directly control 21 provinces, five autonomous regions, and three municipalities, which in turn governed approximately 2,300 county-level governments that supervised over one million branch offices of the Chinese Communist Party in towns, villages, army units, factories, mines, and schools. Policies were established at the top levels and implemented by lower level governments. Information on the effectiveness of policies was collected locally, aggregated by the regional government, and then eventually reported upwards to the Standing Committee (Fairbank, 1986: pp. 297-341; Spence, 1991: pp. 542).

China's size proved to be a difficulty for the centrally planned regime. China is the world's third largest country in terms of geographic size.[52] Therefore, conditions which determine agricultural production, amongst other concerns of the central government, varied widely across regions. China's poor transportation and communications infrastructure also greatly added to the central government's difficulties in obtaining and aggregating information.[53]

In the late 1950s, three factors significantly exacerbated these structural difficulties in admin-

[50]The role of communal kitchens are studied by works such as Yang (1996) and Chang and Wen (1997).

[51]To the best of our knowledge, there is unfortunately, no systematic documentation on why bureaucrats followed the rules. Therefore, our views here should be conservatively interpreted as reasonable speculations based on the historical facts. See Becker (1996) for overview of the promises of the new government.

[52]More precisely, it is the third or fourth largest country, and the exact ranking depends on boundary definitions for certain territories of China and the United States.

[53]Thousands of officials were sent from urban areas to collectives for procurement and information gathering. When they returned to cities, information from each was collected and cumulatively reported to the provincial capital, which aggregated information from across the province and then, in turn, reported it to Beijing. Only then could Beijing have information for the entire country. Traveling between cities, where information was accumulated and policies made, and rural areas, where the food was produced, was very time consuming. Transportation networks were almost completely destroyed by decades of civil unrest (e.g., the civil war between the Sun-Yat Sen led Guomingtang (KMT) and warlords, 1911-1935; the war with Japan, 1936-1945; the civil war between Communist CCP and Chang Kai-Shek led KMT, 1945-49) and reparations had just begun (Fairbank, 1986: pp. 278). The most common method of transportation for officials who traveled to rural areas was a combination of government conveyance vehicles, bicycles and beasts of burden. In a country as geographically vast as China, where urban centers were relatively few and geographically concentrated, it could take many weeks to reach an outlying collective. Moreover, rural areas were typically not connected by telecommunications infrastructure. This meant that the central government learned about production figures from rural areas rather slowly.

istration. First, in order to reduce the budget deficit, the government severely cut expenditures on administration, which declined from 19.3% of total government budget expenditure in 1950 to only 7.8% in 1957 (Eckstein, 1977: pp. 186). Since both China's economy and government expenditures were increasing during this period, these figures suggest that government administration did not grow even though the economy and thus the scope of central planning had increased substantially. Second, the government lost much of its able personnel from the bureaucracy when approximately 700,000 of its most educated bureaucrats were purged in 1957 after the *Hundred Flowers Movement*.[54] Moreover, in 1958, Mao abolished the State Statistical Bureau, which meant that there were no statisticians or demographers in 1959 to project national production figures before all of the harvests were physically procured and aggregated across regions (Fairbank, 1986: pp. 300; Spence, 1991: pp. 580). Third, for political reasons, Mao implemented measures which further decreased the structural flexibility of the system. For example, after the Lushan meeting in 1959, in order to solidify his power, Mao banned the twice-weekly meetings of the Standing Committee and further removed decision-making powers from regional governments, two institutions which helped the leadership address unexpected shocks. By the end of 1959, the Standing Committee met only once every two months, and the regional leadership had little power for independent decision making (Fairbank, 1986: pp. 303).

To summarize, many factors hampered the Chinese government's ability to aggregate and respond to new information. The inflexibility of the Chinese government is similar to the inherent inflexibility in centrally planned economies as discussed in the historic works of Von Mises (1935) and Hayek (1945) and in the theoretical work of Weitzman (1974).

6 Model of Procurement

In this section, we develop a model of procurement policy which shows that policy inflexibility, as described in the previous section, could have generated a famine with the spatial patterns in our empirical results. The model is also useful for understanding the contribution of additional factors (e.g., transport cost, regional favoritism, and local mis-reporting of production) to the famine. Moreover, it allows us to assess the merits of the Chinese procurement policy of fixing quantities relative to an alternative central planning policy of fixing prices.

6.1 Model

We consider procurement policy in an environment in which different regions produce different quantities of food. The government can procure food from some regions and redistribute it to

[54] In 1957, in order to fight off criticism from intellectuals during the *Hundred Flowers Movement*, the leadership promoted the anti-rightist campaign, where as many as 700,000 intellectuals (e.g., high school graduates and above) were removed from government positions. Being branded as a rightist effectively ended the career of the individual. Many were demoted to manual labor jobs for re-education. In extreme cases, individuals were sent to labor camps. This did not directly affect agricultural production, which did not require the labor of intellectuals, but it crippled the bureaucracy.

other regions. A key feature of the environment is that all regions are subject to an aggregate shock that reduces food production. Given our discussion in Section 5, the key constraint faced by the government is that procurement policy cannot respond perfectly to this shock. For simplicity, we focus on the extreme case in which no adjustment to the shock is possible, capturing the notion that the government is either unaware of the shock or cannot respond to the shock. Given this constraint on policy, our model is therefore in the spirit of Weitzman (1974), who studies the optimal choice of quantities in a centrally planned economy in which quantities cannot change with aggregate shocks.

More formally, the economy consists of M rural regions labeled by $i = \{1, ..., M\}$ and N urban regions labeled by $i = \{M+1, ..., M+N\}$. Every region is populated by a mass p_i of identical households with a stochastic per-capita agricultural endowment $e_i(s) \geq 0$. This endowment depends on the aggregate shock $s = \{H, L\}$ which can be high (H) or low (L). Let $\Pr\{s = H\} = 1 - \Pr\{s = L\} = 1 - \mu \in (0, 1)$, the probability that a food reducing aggregate shock is avoided. Let $e_i(H) = \widehat{e_i}$ and $e_i(L) = \widehat{e_i} - \sigma_i$. $\widehat{e_i}$ parameterizes the productivity of a region since a higher $\widehat{e_i}$ corresponds to a higher level of food production per capita. σ_i captures the volatility of production in region i. Urban regions do not produce any food, so that $e_i(s) = 0$ for $s = \{H, L\}$ if $i \in \{M+1, M+N\}$. We consider economies subject to the following two assumptions regarding the food production process:

Assumption 1 $e_i(s)$ is strictly increasing in $\widehat{e_i}$ for $s = \{H, L\}$.

Assumption 2 σ_i is strictly increasing in $\widehat{e_i}$.

Assumption 1 states that more productive regions produce more food per capita during both the high and the low shock. Assumption 2 states that more productive regions experience a higher variance in production (i.e., a sharper drop during the aggregate food downturn). These two assumptions are satisfied for instance if the percentage drop in production is the same across regions. Recall from Section 3.1 that both of these assumptions hold for 1959.

Every household in region i produces food $e_i(s)$ and is subject to a level of food procurement τ_i, where a negative value of τ_i corresponds to a food subsidy. A household's level of food consumption $c_i(s)$ therefore satisfies

$$c_i(s) = e_i(s) - \tau_i \text{ for } s = H, L. \tag{6}$$

Note that while food consumption and production depend on the aggregate shock s, the level of procurement τ_i *does not depend on the aggregate shock*. This assumption is motivated by our discussion in Section 5 where we argue that a central feature of the Chinese procurement is its inflexibility. Though we focus on the extreme situation of complete inflexibility, all of our results also apply in more general settings in which the government can observe an imperfect noisy signal about the state of the economy or adjust planned policies at a cost.[55] We take this inflexibility as

[55]The details are omitted for brevities but are available upon request.

given and do not microfound it since our discussion in Section 5 indicates that there are multiple historical factors behind it.[56] Note that since the shock σ_i equals 0 for urban regions (since they produce no food), the level of consumption $c_i(s)$ is independent of the aggregate shock for these households.

We also allow the government to have an exogenous aggregate procurement target $\theta \geq 0$ (net of interregional transfers), which represents the use of grain for activities other than immediate food consumption, such as storage or export. As such, the government's budget constraint can be written as the following condition:[57]

$$\sum_{i=1}^{M+N} p_i \tau_i = \theta. \tag{7}$$

The government is utilitarian. We make this assumption to highlight the point that, even if the government has the best intentions and weighs individuals equally, a famine with the empirically observed spatial patterns can occur.[58] Therefore, the government chooses a procurement system to maximize the following object:

$$\sum_{i=1}^{M+N} p_i \left((1-\mu)\,\pi(c_i(H))\,\chi + \mu\pi(c_i(L))\,\chi \right) . \tag{8}$$

$\pi(c_i(s))$ corresponds to the probability of survival as a function of consumption $c_i(s)$, and χ corresponds to the value of life.[59] We assume that $\pi(\cdot)$ is continuously differentiable, strictly increasing and strictly concave, so that the probability of survival rises with food consumption but is subject to diminishing returns.[60]

The government knows the productivity $\widehat{e}_i$ and the volatility σ_i of each region and the probability of the aggregate shocks. It is clear in this environment that if the government could

[56]One natural microfoundation takes into account incentives to mis-report production, either for farmers wishing to under-report production or for career-motivated region leaders wishing to over-report production. In such an environment, if the government must confine policies in region i to depend only on region i's report of production (because of limited communication infrastructure, for instance), then the only incentive compatible system assigns a fixed level of procurement for each region which does not depend on that region's report.

[57]As we discussed in the section on background, the government was using grain to fund investment in industry. The government may have also wished to procure increasing quantities of grain from rural regions as a means of providing incentives for farmers to raise production levels since it was procuring more and more over time. Note that an indirect effect of this policy would be to also reduce the total amount produced in each region $e_i(s)$ by weakening the labor supply, an argument made by Li and Yang (2005).

[58]More generally, our results hold if the government is approximately utilitarian since the government then chooses policies which make expected marginal utility approximately equal across households. See Section 6.2 for a discussion of the effect of assuming alternative preferences for the government.

[59]One can alternatively interpret $\pi(c_i(s))$ as corresponding to the fraction of the population in region i which dies from starvation in an environment in which a minimal level of consumption $\underline{c}$ is required to stay alive. Specifically, one can imagine that the food consumption of a given individual z in region i equals $c_i(s) + \varepsilon_{iz}$ for ε_{iz} which represents an idiosyncratic shock. Thus $\pi(c_i(s)) = \Pr\{c_i(s) + \varepsilon_{iz} \geq \underline{c}\}$.

[60]One could alternatively let the procurement target θ be endogenous in our framework by adding an additional term in the government objective: $V(\theta)$, for $V(\cdot)$ which is increasing and concave. This refinement does not affect any of our results.

Note that one could easily incorporate the government's potential bias towards the urban elite without changing any of our results since this would correspond to assigning a higher weight to urban regions in the social objective.

condition procurement τ_i on the shock s, then it would provide all households with the same level of food consumption conditional on the shock. In such an environment, there would be no cross-regional variation in mortality in response to an agricultural shock.[61]

In our environment, such a redistributive policy is not possible because procurement cannot respond to the shock. More specifically, consider a hypothetical procurement policy $\tau = \{\tau_i\}_{i\in\{1,M+N\}}$. Given Assumption 2 and equation (6), it is clear that the variance in consumption for a given region is increasing in $\widehat{e_i}$ (i.e., more productive regions experience a higher variance in production). Thus, under an inflexible procurement policy, it is not possible for the government to equalize consumption across regions in all states of the world. More specifically, the government, in choosing the optimal policy, solves the following program:

$$\max_{\tau}(8) \text{ s.t. } (6) \text{ and } (7).$$

Letting ψ correspond to the Lagrange multiplier on constraint (7), the first order conditions to the government's program yield:

$$(1-\mu)\pi'(c_i(H)) + \mu\pi'(c_i(L)) \;= \psi \;\forall i. \tag{9}$$

Therefore, the government equates the expected marginal utility of food consumption of all households, taking into account that more productive households will inevitably experience a higher variance in food consumption. Equation (9) has some important implications which are summarized in the below proposition. All of the proofs are shown in the Appendix.

Proposition 1 *The policy of the government has the following features:*

1. *Aggregate survival $\sum_{i=1}^{M+N} p_i\pi(c_i(s))$ conditional on $s = \{H,L\}$ under an inflexible policy is below that which occurs under a fully flexible policy,*

2. *Procurement τ_i is increasing in productivity $\widehat{e_i}$, and*

3. *Regional survival $\pi(c_i(s))$ is increasing in production $e_i(s)$ if $s = H$, and regional survival $\pi(c_i(s))$ is decreasing in production $e_i(s)$ if $s = L$.*

Corollary 1 *The variance of mortality $Var(\pi(c_i(s)))$ is increasing in productivity $\widehat{e_i}$.*

The first part of Proposition 1 states that famine intensity is higher under an inflexible government policy relative to famine intensity under a fully flexible policy which equalizes the distribution of food consumption (and minimizes mortality). This result follows from Assumption 2 which implies that the variance in food consumption must differ across regions under an

[61] In principle, a famine could still occur if the procurement target θ is very high. However, as a reminder, recall from the discussion in the background section that if one takes the aggregate procurement data used by Li and Yang (2005) at face value, average per capita grain retention in 1959 was not low enough to cause mortality.

inflexible policy. This result suggests that even under the best procurement policy, it is possible for some individuals to die of famine.

The second part of Proposition 1 states that procurement is increasing in productivity $\widehat{e_i}$, so that more productive regions experience a higher procurement tax relative to less productive regions. This conclusion follows from Assumption 1. More productive regions produce more in all states of the world, so that a government seeking to equalize food consumption will redistribute towards the less productive regions and procure more from more productive regions. This implication is consistent with data we collected on province-level procurement targets and production for the years 1980-88, which, to the best of our knowledge (based on interviews with government officials), used a similar formula for setting procurement as in the 1950s and 1960s.[62] Appendix Figure A4 plots procurement targets as a function of a moving average of per capita production in the 1980s. It shows a strong positive correlation between procurement targets and average per capita production.

The third part of Proposition 1 states that, during a food production boom, mortality and production are negatively correlated across regions. In contrast, during a food production downturn, mortality and production are positively correlated across regions. This prediction is consistent with our empirical findings regarding the spatial distribution of famine intensity. To understand this phenomenon intuitively, recall that more productive regions have more volatile production (Assumption 2), though all regions are subject to an inflexible and non-volatile procurement policy. Thus, more productive regions experience more volatile consumption, a result stated formally in Corollary 1. Since the government cares about all households equally, it follows that households subject to more volatile consumption experience relatively higher consumption during the food production boom and relatively lower consumption during the food production downturn, leading to the spatial patterns of mortality.[63] In other words, in the presence of a large shock to production, the government over-procures from the more productive regions (relative to the fully flexible policy), and this over-procurement amplifies the mortality effect of the downturn in production. For an illustration of the mechanics of this model, recall the simple stylized example presented in Table 1.

It is important to note that this third result on the correlation between food production and mortality refers to food production booms and downturns in a relative sense and not in an absolute sense. In other words, food production and mortality are positively correlated if food production is lower than *anticipated*, and vice versa. This can explain why mortality and regional food production were positively correlated during the famine but not in other years when aggregate food production per capita may have been comparable in levels to 1959. We present a dynamic version of the model in the Appendix to illustrate this formally. All of our

[62]The main difference between the famine era and subsequent decades is that the government aimed for a lower level of procurement, but the method for determining the differences in amount by region did not change. We were unable to obtain procurement target data from the famine era.

[63]If the government is biased towards the urban elite, this prediction will hold for the sample of rural regions but not for the comparison of urban versus rural regions.

results are robust to this extension.

Notice that the quantitative magnitude of such a mechanism can be potentially large. In practice, it is difficult to estimate this model given the challenge associated with determining the exact functional form for $\pi(\cdot)$ (see the discussion in Section 4.1). In principle, the combination of an inflexible policy, a large unexpected drop in production, and a high aggregate procurement target could significantly amplify the mortality consequence of a drop in per capita production. See the Appendix for a stylized example that formally illustrates this point.

In addition to capturing patterns associated with the famine, our model makes several general predictions regarding mortality outcomes under an inflexible procurement system. For instance, consider a generalization of our setting in which every single region's stochastic food endowment $e_i(s)$ is subject to a set of aggregate and idiosyncratic shocks, where the level of procurement is independent of the realized shock because of inflexibility. In such an environment, the following two patterns emerge: (1) mortality is negatively correlated with per capita food production *within* a region, and (2) variance in per capita food production is positively correlated with variance in mortality *across* regions. These two patterns emerge precisely because of the imperfect insurance resulting from the inflexibility of the food procurement system; regions must bear a portion of the risk associated with their own stochastic production. We find that these two patterns are present in the data. More specifically, if we re-estimate (1) with province fixed effects, we find that the coefficient on grain production is negative and that the sum of this coefficient plus the interaction effect is also negative. Moreover, we find that the correlation between the within-province standard deviation in log mortality and the within-province standard deviation in log production per capita is positive.[64] These patterns which point to the presence of imperfect insurance provide additional evidence regarding the inflexibility of the procurement system.

6.2 Additional Factors

Our model shows how a procurement policy that cannot adjust to aggregate shocks can amplify the mortality increase from an aggregate food production downturn. Moreover, when the percentage drop in production is broadly the same across regions, as was the case in China in 1959, an inflexible procurement policy can lead to the spatial patterns of production and mortality that appear in the historical and retrospective data. Within this framework, many factors can further amplify the famine. In this section, we discuss the role of factors that are commonly believed to have contributed to the Chinese famine: transport costs, government bias towards certain regions and mis-reporting of production. In each case, we discuss how the factor can amplify the magnitude of the famine predicted by an inflexible procurement policy.

First, we consider the role of transport cost, an important mechanism in the general context of famines and the focus of several recent studies on the causes of famine (e.g., Shiue, 2004,

[64]These correlations are statistically significant at the 1% level. For brevity, we do not report them in the paper. They are available upon request. In addition, we find that the within-county standard deviation in log cohort size is positively correlated with the within-county standard deviation in rainfall.

2005; Burgess and Donaldson, 2010). If transport costs are high, they could generate a famine by making it difficult for the government to transfer food from high food productivity regions to low food productivity regions. It is straightforward to introduce transport costs into our model without changing our results. For instance, one can subtract $\delta\tau_i^2/2$ from the right hand side of equation (6), where $\delta \geq 0$ reflects the transport cost. Adding such a transport cost will make the government's planning problem even more difficult and further amplify the effect of a downturn in food production by causing a significant portion of food to be lost in transit.

Second, we consider the possibility of governmental regional favoritism. This phenomenon could arise from political self-interest or malevolence, as leaders may prioritize certain regions that are more important for ensuring political success or maintaining political stability.[65] In particular, Lin and Yang (2000) argue that the famine era government strongly favored urban areas. In our model, favoritism could be introduced by assigning more weight to certain regions in the social welfare function. This would lead to higher famine intensity in the less favored regions and to further inequality in food consumption across regions.

Finally, we consider the effects of local mis-reporting of production. Numerous survivors and past officials recall that local officials over-reported grain production in order to appease party leaders.[66] It is important to note that our model already incorporates this possibility of mis-reporting – it is one of the reasons that the central government cannot trust local reports and resorts to an inflexible policy. However, we can further incorporate mis-reporting in a dynamic extension of our environment which is explore more formally in the Appendix. One can imagine that the government forms its regional production expectations based on past production reports. In such an environment, it is clear that over-reporting in previous years will cause the government to over-procure during an aggregate downturn, which will amplify the ensuing mortality rates. More generally, government incompetence can also cause the miscalculation of parameters and amplify famine. For instance, if the government under-estimates region-specific food production volatility, σ_i, or the probability of a food production downturn, μ, the government will over-procure by even more, a behavior that reinforces our result.

While the factors discussed above contribute to famine, they cannot alone easily explain the spatial patterns of famine that we observe in the data absent an inflexible procurement policy. First, transport costs make it difficult for the government to transport food from productive regions to unproductive regions, causing famine severity to *decrease* with regional production. In such a scenario, cities, which do not produce food, will always suffer the worst famine. Both implications run contrary to the findings in the data.[67] Second, a government regional bias on its own would cause some regions to consistently experience higher mortality both during food

[65]Recall from Section 2 that there is little evidence that the government intended for the famine to occur in general. Therefore, we do not consider the obvious consequences of a malevolent government that wants famine.

[66]Becker (1996) provides many examples in his book.

[67]In addition, transports costs probably also contributed to general inflexibility in the centrally planned procurement regime by making communication among government bureaucrats difficult and by making it difficult for the government to reverse its procurement policy and return the grain back from depots to rural areas.

production booms and downturns. It cannot explain the reversal in the correlation between food production and mortality between famine and non-famine years unless if favoritism somehow also reversed in 1959. Moreover, recall that the retrospective analysis using county-level data shows that the reversal of the relationship between food production and well-being exists even when we control for province-year fixed effects. This means that for favoritism to generate the reversal, favoritism would have to be reversed not only at the province-level, but also at the county-level within provinces.[68] Finally, over-reporting of production cannot, absent inflexibility, generate the reversal in the correlation between per capita production and mortality rate during the famine unless the correlation between production and over-reporting is also reversed during the famine. That this scenario took place precisely in 1959, and in 1959 alone, seems highly implausible.[69]

More generally, additional factors which may have contributed to famine cannot on their own yield theoretical predictions consistent with the striking positive correlation between regional food production and mortality that exists only in 1959 and is reversed in non-famine years. This is particularly the case since there are no known major policy changes that are specific to 1959.[70] This suggests that the combination of inflexibility and an unexpected drop in production in 1959 are critical for understanding the geographic distribution of famine in China.

[68]We can explore the existence of an urban bias directly by including the log of urban population and its interaction term with the famine year dummy as controls into equation (1). The results are reported in Appendix Table A9. They show that having a larger urban population results in higher provincial mortality rates on average. This supports Lin and Yang's (2000) argument that there was an urban bias. However, the urban population does not affect mortality differently during the famine (the difference between the main effect and interaction effect is insignificant). More importantly, our main results, the reversal in the correlation between grain production and mortality during the famine years is robust to these controls. In fact, it becomes more prominent in magnitude and is statistically significant at the 1% level. These results suggest that there was an urban bias in the procurement policy, but that this bias does not affect the reversal in the relationship between production and mortality caused by the inflexibility in the grain procurement system.

[69]For mis-reporting alone (i.e., with a hypothetically fully flexible procurement policy) to account for the spatial distribution of famine, one would have to make the following two assumptions: (i) For some behavioral reasons, regional leaders who over-report during the food production downturn come from more productive regions and also under-report during the food production boom; and (ii) the government is unaware of this mis-reporting bias. This would generate the positive correlation between productivity and mortality in famine years and the reversal of this correlation in non-famine years. Our prior is that this is unlikely. However, we cannot be conclusive on this point because data for measuring regional over-reporting do not exist. In the interest of completeness, we searched through 1958-59 provincial newspapers stored in the National Library Archives in Beijing for contemporaneous official reports of production. We were only able to find reported production numbers for a few provinces in 1959. They show no apparent correlation between production and over-reporting, which we measured as the difference between the newspaper reports and either 1958-59 production or average production. These data should be interpreted very cautiously due to their poor quality. They are not reported in the paper for brevity.

[70]Recall from Section 2 that communism began in 1949, collectivization began in the early 1950s, the grain procurement system had been in place since the mid-1950s, communal kitchens had been in place at least since the Great Leap Forward began in 1957, which also implies that other GLF policies such as the diversion of labor from agriculture also began in 1957. Procurement levels had also been high since the mid 1950s. Appendix Table A8 column (6) shows that procurement had been approximately 30% since 1954. Had production in 1959 continued to grow at the same rate as previous years, procurement in 1959 would have been very similar.

6.3 Counterfactual Exercise: Fixing Quantities vs. Prices

Thus far, we have provided qualitative and quantitative evidence that suggests the Chinese government's commitment to central planning led to an inflexible procurement policy, which in 1959, caused severe over-procurement and famine. A natural question that follows is whether the Chinese government, as central-planners constrained by inflexibility, could have chosen a better policy. For an inflexible central planner, the obvious alternative is to purchase food from rural areas at a fixed price to redistribute to urban areas. This style of policy was used in other central planning regimes like the Soviet Union during the New Economic Policy. In China, price fixing was used in the early 1950s and again in more recent years.

In this section, we extend the model developed in Section 6.1 to discuss the trade-offs between quantity versus price controls. These two policies have different implications for which segment of the population bears the burden of aggregate shocks to food production. In highlighting these trade-offs, we show that a government is better off pursuing quantity controls if a large fraction of the population is rural and if the magnitude of productivity shocks across the rural populations is relatively homogenous.

For this exercise, we introduce a second consumption good to the model to serve as the numeraire for the price of food P. Our exercise is similar in spirit to Weitzman (1974) who studies the use of price and quantity controls in an inflexible policy setting. As in Weitzman (1974), one can gain insight by assuming linear preferences over the non-food good and the quadratic functional form for $\pi(\cdot)$. More specifically, letting $x_i(s)$ represent household i's non-food consumption as a function of the state s, the social welfare function is

$$\sum_{i=1}^{M+N} p_i \chi\left((1-\mu)\left(\pi\left(c_i(H)\right)\chi + x_i(H)\right) + \mu\left(\pi\left(c_i(L)\right)\chi + x_i(L)\right)\right)$$

$$\text{for } \pi\left(c_i(s)\right) = \begin{cases} 1-\alpha\left(\bar{c}-c_i(s)\right)^2 & \text{if } c_i(s) \leq \bar{c} \\ 1 & \text{if } c_i(s) > \bar{c} \end{cases}$$

for some parameter $\alpha > 0$.[71] The total supply of the non-food good is normalized to some positive number X, so that $\sum_{i=1}^{M+N} x_i(s) = X$. For simplicity, we set the aggregate procurement target θ to zero.

The government fixes the price of food as follows. It commits to purchasing any quantity of food from rural households at price P and redistributes this food to urban households. To finance these purchases, the government taxes the non-food endowment of households. Note that because preferences over non-food consumption are linear, the government does not care about inequality in non-food consumption from this taxation, so we can effectively ignore non-food consumption in the government's optimization program. For simplicity, suppose that the

[71] Given our assumption on preferences in this extension, the utilitarian optimum could be achieved with perfectly competitive markets. For this exercise, we rule out this possibility to examine the less dramatic measure of fixed supplier prices which the Chinese government may have been able to pursue during this time period.

government chooses an interior price $P \geq \max_{i \in \{1,M\}} \pi'(e_i(L))$ so that it is sufficiently high that all rural households would choose to sell food to the government in all states of the world.[72] Moreover, to facilitate interpretation, suppose that the implied level of consumption under the optimal policy always satisfies $\bar{c} \geq c_i(s) \geq \bar{c} - \alpha^{-1/2}$ so that the value of $\pi(\cdot)$ is always between 0 and 1. In this circumstance, the first order conditions for rural households would imply that

$$\pi'(c_i(s)) = P \text{ for } s = \{H, L\} \ \forall i. \tag{10}$$

so that all rural households have a level of food consumption that is independent of the aggregate shock and that sets the marginal utility of food consumption equal to the price of food. This setup means that during the food production boom, farmers sell more food to the government, and during the food production downturn, they sell less food to the government. Consequently, urban households all have a volatile consumption and endure the entire risk associated with the aggregate production shock. Interestingly, this situation is the exact *opposite* to the environment with fixed quantities in which the entire burden of the aggregate production shock is borne by rural households.

By analogous reasoning as in the environment with fixed quantities, optimal policy implies the first order condition in equation (9), where ψ must be interpreted as the Lagrange multiplier for the resource constraint of the entire economy. More specifically, the government equates the expected marginal utility of food consumption across households, taking into account that this level of consumption is deterministic for rural households and stochastic for urban households. In addition to treating all rural households symmetrically, the government treats urban households symmetrically, so that they all equally bear the burden of the aggregate shocks.

Proposition 2 ***(fixing quantities dominate fixing prices)*** *Expected mortality is lower under a fixed quantities policy relative to fixed prices policy if and only if the following condition holds:*

$$\left[\sum_{i=1}^{M} p_i / \sum_{i=M+1}^{M+N} p_i\right] \left[\left(\sum_{i=1}^{M} \frac{p_i}{\sum_{i=1}^{M} p_i} \sigma_i\right)^2 / \sum_{i=1}^{M} \frac{p_i}{\sum_{i=1}^{M} p_i} \sigma_i^2\right] > 1 \tag{11}$$

Proposition 2 states that a policy of controlling quantities dominates a policy of controlling prices if the size of the rural population is significantly higher than the size of the urban population (i.e., $\sum_{i=1}^{M} p_i$ is significantly higher than $\sum_{i=M+1}^{M+N} p_i$) and if the cross-sectional variance in the magnitude of shocks σ_i across rural regions is sufficiently low (i.e., $\sum_{i=1}^{M} p_i \sigma_i^2 / \sum_{i=1}^{M} p_i$ is sufficiently low relative to $\left(\sum_{i=1}^{M} p_i \sigma_i / \sum_{i=1}^{M} p_i\right)^2$).

The intuition for this proposition is as follows. Imagine for simplicity that all rural regions are identical so that condition (11) collapses to $\sum_{i=1}^{M} p_i > \sum_{i=M+1}^{M+N} p_i$. This means that quantity controls dominate price controls if the urban population is in the minority. To understand this

[72] If the government could choose a region-specific price, it would choose the same price for all regions since it is utilitarian.

result, note that if the rural households are a majority, then the government faces a choice between having a majority of the population experiencing small consumption fluctuations under fixed quantities versus having a minority of the population facing large consumption fluctuations under fixed prices. The government prefers to let a majority experience the shock because large volatilities in consumption are extremely costly to the government from a welfare perspective and it is better to pool this risk across rural households.[73]

To understand why quantity controls dominate price controls only if the cross-sectional variance in the magnitude of shocks is low, imagine for simplicity that rural and urban regions have the same population size so that (11) collapses to $0 > Var(\sigma_i)$. Thus, if rural households are homogeneous, then quantity and price controls are equivalent from a welfare perspective for reasons previously discussed. However, if there is any heterogeneity in the productivity shocks across rural households, then price controls dominate quantity controls. The reason is because price controls make it possible for the urban population to pool all of the differential risk faced by the rural population. For example, if there were two rural regions of equal size, one with a higher value of the shock σ_i than the other, then the government would prefer to let the urban population experience an intermediate level of consumption volatility under price controls versus having one half of the farmers experiencing very high volatility and one half of the farmers experiencing very low volatility under quantity controls.[74]

In conclusion, a retrospective evaluation of the merits of the Chinese procurement policy of fixing quantities versus the alternative policy of fixing prices relies on two factors. On the one hand, the fact that the urban population was a small minority of the total Chinese population points to the advantages of the quantity-fixing policy practiced by the Chinese government over an alternative policy of fixing prices. On the other hand, the fact that rural regions were far from identical (e.g., more productive regions experienced a larger reduction in total production during the famine) points to the advantages of fixing prices over fixing quantities. In principle, the government can pool the risk associated with this heterogeneity by fixing a price at which farmers will sell their food to the cities, but this will come at the cost of increasing the volatility of mortality outcomes in cities. The extent to which the procurement policy dominated price controls is an important quantitative question for future research.[75]

[73]This insight is related to Weitzman's (1974) result that quantity controls dominate price controls if the absolute value of the second derivative of the benefit function with respect to quantity exceeds the second derivative of the cost function with respect to quantity. This is also true in our setting if one interprets the benefit function as the portion of social welfare attributable to the urban population and the cost function as the negative of the portion of social welfare attributable to the rural population. In this light, the relative curvature of each function depends on the relative size of the urban and rural population.

[74]Note that this second effect regarding the distribution of productivity is not present in Weitzman (1974) since he assumes only one producer for each good.

[75]Our analysis also ignores the fact that the government may have also put a lot of weight on the accumulation of grain either for storage purposes or for exportation. To the extent that this was a major concern, this would bias the government away from choosing a policy of fixing prices and towards a policy of fixing quantities.

7 Conclusion

Our study points to inflexible government policy for food distribution as an important factor in causing the largest famine in history. We show that even if a government is not obviously malevolent or incompetent, an ideological commitment to central planning, together with practical constraints for gathering and responding to information, result in an inflexible policy that can cause a famine when aggregate production falls. If the production falls are broadly proportional across regions, as was the case in China in 1959, then an inflexible policy will reverse the normally positive correlation between per capita food production and well-being during the shock, causing the counter-intuitive pattern that famine is more severe in regions that produce more food per capita.

As with all historical events, the total underlying forces driving the famine are inevitably more complex than the insights highlighted by one study. While our results show that understanding the inflexible nature of central planning and how it relates to food distribution is crucial for understanding famine in non-market economies, it would be naive to assert that inflexibility was the sole driver of the deaths of thirty million people or that central planning made their mortality inevitable. Our theoretical framework, quantitative analyses, and qualitative historical evidence show that the famine must have in part been the outcome of the unfortunate combination of an inflexible and ambitious procurement policy together with a surprisingly large fall in aggregate production in 1959. This crisis was further exacerbated by political tensions of the time, which pressured bureaucrats to follow the prescribed rules. It is interesting to note that the government did not respond to the famine by abandoning the inflexible procurement system altogether. Instead, it simply reduced procurement rates from approximately 30% in the 1950s to 18% in the 1970s (see Appendix Table A8). This adjustment, together with the absence of another production drop of a similar proportions, perhaps explains why China did not experience famine in subsequent years.[76]

This study opens several interesting avenues of future research. A natural one is to study the geographic patterns and institutional details of other famines that have occurred in non-market economies. Such studies could provide much insight into the determinants of famine and more generally, shed light on the relevant constraints for central planning. This is important as all governments engage in some form of central planning. Other interesting topics for future work include understanding the determinants of the flexibility of government policy and quantifying the trade-off between flexible and inflexible policies.

[76] In our sample of 27 provinces, per capita grain production fell by 15% in 1959, which is more than two standard deviations below the mean growth rate in per capita grain production. There were several other drops in aggregate production. But the size of these other falls were smaller, typically ranging between 5-10%.

References

Ashton, Basil, Kenneth Hill, Alan Piazza, and Robin Zeitz (1984) "Famine in China, 1958-1961," *Population and Development Review*, 10, 613-645.

Banister, Judith (1987) *China's Changing Population*, Stanford University Press, Stanford.

Becker, Jasper (1996) *Hungry Ghosts: China's Secret Famine*, John Murray Publishers, London.

Besley, Timothy and Torsten Persson (2009) "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 99, 1218-1244.

Burgess, Robin and David Donaldson (2010) "Can Openness Mitigate the Effects of Weather Shocks? Evidence from India's Famine Era," *American Economic Review*, 100, 449-453.

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

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

Cohen, Charles and Eric D. Werker (2008) "The Political Economy of "Natural" Disasters," *Journal of Conflict Resolution*, 52, 795-819.

Dasgupta, Partha and Ray, Debraj (1986) "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *Economic Journal*, 96, 1011-1034.

Demick, Barbara (2009) "The Good Cook: A Battle Against Famine in North Korea," *The New Yorker*, November 2nd 2009: 58-64.

Dreze, Jean (1999) *The Economics of Famine*, Edward Elgar Publishing, London.

Eckstein, Alexander (1977) *China's Economic Revolution*, Cambridge University Press, Cambridge.

Fairbank, John King (1986) *The Great Chinese Revolution: 1800-1985*, Harper & Row, New York.

Gittings, John (1967) *The Role of the Chinese Army*, Oxford University Press, New York.

Greif, Avner (2008) "Towards Political Economy of Implementation: The Impact of Administrative Power on Institutional and Economic Developments," forthcoming in Elhanan Helpman (ed.), *Institutions and Growth*, Harvard University Press, Cambridge

Haggard, Stephan and Marcus Noland (2005) "Hunger and Human Rights: The Politics of Famine in North Korea," *U.S. Committee for Human Rights in North Korea*, Washington D.C.

Hayek, Friedrich A. (1945) "The Use of Knowledge in Society," *American Economic Review*, 35, 519-530

Heal, Geoffrey M. (1969) "Planning Without Prices," *Review of Economic Studies*, 36, 347-362.

Johnson, D. Gale (1998) “China’s Great Famine: Introductory Remarks” *China Economic Review,* 9, 103-109.

Kahn, Matthew (2005) "The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions," *Review of Economics and Statistics*, 87, 271-284.

Keys, Ancel (1950) *The Biology of Human Starvation: Volume I*, University of Minnesota Press, Minnesota.

Kueh, Y.Y. (1995) *Agricultural Instability in China, 1931-1990: Weather, Technology and Institutions*, Oxford University Press, United Kingdom.

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

Lin, Justin Yifu (1990) "Collectivization and China’s Agricultural Crisis in 1959-1961," *Journal of Political Economy*, 98, 1228-1252.

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

Malinvaud, Edmond (1967) "Decentralized Procedures for Planning ", ch. 7 in Malinvaud (ed.), *Activity Analysis in the Theory of Growth and Planning*, Macmillan, New York.

Malthus, Thomas (1798) *An Essay on the Principle of Population*, Penguin Classics, New York.

Meng, Xin and Nancy Qian (2009) “The Long Run Impact of Exposure to Famine on Survivors: Evidence from China’s Great Famine 1959-61," NBER Working Papers 14917.

Milan, F.A. and Rodahl, K. (1961) “Caloric Requirements of Man in the Antarctic” *Journal of Nutrition*, 75, 152-156.

Ministry of Agriculture (1989) *Chinese Village for Thirty Years (Xiang Cun San Shi Nian) 1949-83*, Agriculture Press of China, Beijing.

Nunn, Nathan and Nancy Qian (2009) "The Impact of Potatoes on Old World Population and Urbanization," Brown University Working Paper.

O’Grada, Cormac (2007a) "Famine and Markets," UCD Centre for Economic Research Working Paper WP07/20.

O’Grada, Cormac (2007b) "Famine Demography," UCD Centre for Economic Research Working Paper WP07/21.

Oi, Jean (1989) *State and Peasant in Contemporary China: The Political Economy of Village Government*, University of California Press, Berkeley.

O’Rourke, Kevin (1994) "The Economic Impact of the Famine in the Short and Long Run," *American Economic Review*, 84, 309-313.

Peng, Xizhe (1987) "Demographic Consequences of the Great Leap Forward in China’s Provinces," *Population and Development Review*, 13, 639-670.

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

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

Reynolds, Bruce L. (1981) "Changes in the Standard of Living of Shanghai Industrial Workers, 1930-73," in Christopher Howe (e.d.), *Shanghai: Revolution and Development in an Asian Metropolis*, Cambridge University Press, Cambridge.

Sen, Amartya (1981) *Poverty and Famines: An Essay on Entitlement and Deprivation*, Clarendon Press, New York.

Shiue, Carol H. (2002) "Transport Costs and the Geography of Arbitrage in Eighteenth Century China," *American Economic Review*, 92, 1406-1419.

Shiue, Carol H. (2004) "Local Granaries and Central Government Disaster Relief: Moral Hazard and Intergovernmental Finance in Eighteenth and Nineteenth Century China," *Journal of Economic History*, 64, 100-124.

Shiue, Carol H. (2005) "The Political Economy of Famine Relief in China, 1740-1820," *Journal of Interdisciplinary History*, 36, 33-55.

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

Twitchett, Dennis and John King Fairbank (1985) *Cambridge History of China*, 14, 378-386, Cambridge University Press, New York.

Vallin, Jacques, France Mesle, Serguei Adamets and Serhii Pyrozhkov (2002) "A New Estimate of Ukrainian Population Losses during the Crises of the 1930s and 1940s," *Population Studies,* 56, 249-64.

Von Mises, Ludwig (1935) "Economic Calculation in the Socialist Commonwealth," in in F.A. Hayek (ed.), *Collectivist Economic Planning*, Routledge and Sons, London.

Walker, Kenneth R. (1984) *Food Grain Procurement and Consumption in China,* Cambridge University Press, Cambridge.

Weitzman, Martin (1974) "Prices vs. Quantities," *Review of Economic Studies*, 41, 477-491.

Weitzman, Martin (1970) "Iterative Multilevel Planning with Production Targets," *Econometrica*, 38, 50-65.

West, Lorraine and Yaohui Zhao (eds.) (2000) *Rural Labor Flows in China*, Institute for East Asian Studies, University of California, Berkeley.

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

Yang, Dennis Tao (2008) "China's Agricultural Crisis and Famine of 1959-61: A Survey and Comparison to Soviet Famines," *Comparative Economic Studies*, 50, 1-29.

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

Zeckhauser, Richard (1996) "The Economics of Catastrophes," *Journal of Risk and Uncertainty*, 12, 113-140.

Data Appendix

Historical Data

Much of the recent revisions have been made possible by the uncovering of contemporaneous reports of production that were not exaggerated and not published in the past. An example of some of these “re-discovered” collective reports can be found in the multi-volume government publication *Villages for Thirty Years* (Ministry of Agriculture, 1989). In theory, the NBS could have used such reports to estimate the amount of exaggeration and to make projections across similar regions. As far as we are aware, the details of how the revisions are made in practice have not been documented. However, a comparison of contemporaneous reports of grain production and the reconstructed data suggests that the production numbers have been drastically revised downwards. For example, *The People's Daily* (August 1, 1958), claimed that “Rice production exceeded 7500 kg per mu (0.067 hectare)” for a county in Hubei province. The revised statistics report that actual grain output in that province was closer to being 120 kg per mu.

If we aggregate production across provinces, our production data for all thirty provinces is approximately 10% lower than the aggregated production from 24 provinces in the 1989 Ministry of Agriculture series used by Li and Yang (2005). This is consistent with the notion that during the ten years between when these two series were published, the NBS made a sincere effort to revise past production numbers.

Famine induced total mortality numbers vary between 16.5 million (Coale, 1981) to 30 million (Banister, 1987) due to different estimation methods (e.g., 18.5 million in Yao, 1999; 23 million in Peng, 1987; and 29 million in Ashton et al., 1984). Our mortality data show that mortality for 27 provinces (excluding Sichuan, Tibet and Hainan) during the years 1959-1961 sum to approximately 21.5 million individuals. Since most believe that mortality numbers were very high in Sichuan, which is not included in our data, this means that our estimates are not far from the higher estimates of mortality.

Birth Cohort Size Data

The relationship between historical mortality data and retrospectively constructed cohort sizes can be illustrated visually by plotting mortality rates and cohort sizes over time for each province. Appendix Figures A2A and A2B show clearly that for every province that exhibits a spike in mortality rate during the famine in Appendix Figure A2A, there is a corresponding drop in famine birth cohort size in Appendix Figure A2B. However, there are several provinces that exhibit a drop in birth cohort size in Appendix Figure A2B for which there is no corresponding spike in mortality rate in Appendix Figure A2A (e.g., Guangdong, Henan, Jiangxi, Shandong). Note that using the survival measure, we are able to proxy for famine intensity for the three provinces for which there is no famine era mortality data (e.g., Sichuan, Hainan, and Tibet).

There are two caveats for interpreting this retrospective measure of famine severity. First, a small number of non-agricultural households in 1990 may have been agricultural households at the time of the famine. We believe that this number is very small because it is very difficult for households to transition one's household type from agriculture to non-agriculture. Second, there may have been some cross-regional migration between the time of the famine and 1990. Recall from Section 2 that famine-driven migration did not occur in mass due to heavy restrictions. However, migration may have occurred for other reasons afterwards. This is unlikely to be a big issue since studies have found that strict migration policies during this period actually made it extremely difficult for rural individuals to move.[77] However, to be cautious, we attempt to address this with the data. The Census does not report region of birth. Therefore, we restrict the sample to households who report as living in the place they are reporting from for at least five years. This excludes very few households from the sample, most of which are amongst non-agricultural households. For the purposes of our study, we then assume that for this restricted sample, the county of residence is the county of birth. As another precautionary measure, we focus our analysis on the agricultural households since most of the migration would have been concentrated among non-agricultural households between urban areas.

Suitability

The GAEZ (2002) data are the result of over twenty years of research and are the product of a joint collaboration between the FAO and the International Institute for Applied Systems Analysis (IIASA). The data on suitability is available at a 50 km×50 km grid cell level, where one can choose the level of agricultural inputs on which to base the calculation. Our chosen level of inputs allows for rain-fed irrigation but no heavy machinery or chemical fertilizers since GLF policies forbade chemical fertilizers and since the use of heavy machinery such as tractors would have been unlikely in this era. We aggregate grid-level data to the county data as follows. The grid-level data reports the predicted amount of output of rice and wheat. If a grid can produce 40% or more of the maximum possible output for any grid (in the world), then we code it as "suitable". The suitability measure at the county-level is the fraction of grids within a county that is suitable. We use this measure for the sake of computational ease.[78] Since procurement targets treated rice and wheat similarly, our measure of suitability is the union of

[77]There is broad consensus that migration was largely controlled until very recently, and most of the migration that did occur was across urban areas, which would not affect this study. In principle, it is possible that some rural regions at the time of the 1990 Census Remuneration may have contained urban youths who were moved from cities to rural areas during the Cultural Revolution (1966-76). However, there have been no accounts to suggest that such movement was correlated with famine intensity. See West and Zhao (2000), which surveys studies on migration in China.

[78]Moderately changing the threshold so that suitable is defined as a grid that produces 20% or 60% of the maximum does not affect the estimates. Using county-level production data from the *1997 Agricultural Census* shows that our measures of suitability are highly correlated with actual production. The correlation across counties is approximately 0.7 and statistically significant at the 1% level. To assess whether our suitability data are good proxies for historical production, we can aggregate the measures to the province-level to show that suitability is also a good predictor of production at that more aggregate level.

land that is suitable for either rice or wheat within each county.[79] Because the GAEZ data uses an administrative map from 2000, changes in administrative boundaries between 1990 and 2000 causes approximately 20% of counties from the 1990 Population Census to not have matches in the GAEZ data. To investigate whether this could bias our estimates, we examine whether being unmatched is correlated with famine severity (as measured by the ratio of the famine birth cohort size to pre-famine birth cohort size). We find no correlation (near zero in magnitude and statistically insignificant) and conclude that this should not affect our analysis.

Historical Weather Data

For each county, we are able to observe the weather variables reported by the nearest station. In our interviews with Chinese farmers, they reported that the most important determinant of a good harvest was precipitation during the Spring months. Using county level production data from the 1997 Agricultural Census (the only year that county-level production data is available nationwide), we verify that spring precipitation is highly positively correlated with production (per hectare of land sown). Therefore, we use average rainfall during the months of February, March and April at the weather station in a county as our proxy for production each year. The sample of counties for which we have historical weather data is small because initially, there were every few stations. During 1949-66, there are 73 stations, only 24 of which existed in 1949.

Calculation of Caloric Needs

We estimate the average body weight of each age-sex group from physical examination data from rural households from the *China Health and Nutritional Survey* 1989. We assume that all adults age 21-50 perform a high level of physical activity, and those age 51-100 perform a medium level of physical activity. Caloric needs for staying alive in Panel B are estimated to be 43% of what is needed for working and healthy child development. This is projected from the assumption that an adult male laborer needs approximately 900 calories to stay alive, which is approximately 43% of the requirement for heavy physical labor. See Dasgupta and Ray (1986) for a discussion of caloric needs. Our calculation shows that the average population caloric need for productive agricultural laborers (or for normal child development) is 1,870.7 calories per day and that the average need to stay alive is approximately 804.4 calories per day (see Appendix Table A2).

Appendix Table A4 lists the provinces in ascending order of mortality rate in 1960 (which captures deaths in the 1960 winter following procurement in the fall of 1959). Columns (1) and (2) show mortality rates in 1960 and production in 1959. In Columns (3) and (4), we calculate grain production that was in surplus of the two benchmarks used in the previous section. We find that all provinces produced more than what was needed to avoid mortality, and only four

[79]Nunn and Qian (2010) provide a detailed description of the construction of this data and how to calculate suitability measures at the regional level from this data. We follow their method. Appendix Figures A1A and A1B present maps that overlay county-level boundaries with the grid-level suitability measures for rice and wheat, respectively.

provinces, three of which are the primarily urban province-level municipalities (Shanghai, Beijing, and Tianjin), produced less than what was needed for a healthy labor force and normal child development.[80]

Theory Appendix

Proofs of Proposition 1 and Corollary 1

An equal distribution of consumption maximizes (8) but cannot be achieved given constraint (6) and (7) by Assumption 2 which proves the first part of the proposition. Consider two regions k and l with $\widehat{e}_k > \widehat{e}_l$. If $\tau_k \leq \tau_l$, then $c_k(s) > c_l(s)$ for $s = \{H, L\}$ by Assumption 1, but given the concavity of $\pi(\cdot)$, this violates (9), which proves the second part of the proposition. If $c_k(L) > c_l(L)$, then by (6) and Assumption 2, this implies that $c_k(H) > c_l(H)$ which violates (9). Therefore, $c_k(L) < c_l(L)$ and satisfaction of (9) implies that $c_k(H) > c_l(H)$, and this proves the third part of the proposition. To prove the corollary, note that $Var(\pi(c_i(s))) = \mu(1-\mu)[\pi(c_i(H)) - \pi(c_i(L))]^2$. Since $c_i(H)$ is rising in $\widehat{e}_i$ and $c_i(L)$ is declining in $\widehat{e}_i$, $\pi(c_i(H)) - \pi(c_i(L))$ is rising in $\widehat{e}_i$, which implies that $Var(\pi(c_i(s)))$ is rising in $\widehat{e}_i$. **Q.E.D.**

Proof of Proposition 2

Under fixed quantities, (9) together with (6) and (7) imply that if $i \in \{1, M\}$, then

$$\begin{aligned} c_i(H) &= \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i + \mu\sigma_i \text{ and} \\ c_i(L) &= \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i - (1-\mu)\sigma_i, \end{aligned}$$

and if $i \in \{M+1, M+N\}$, then

$$c_i(H) = c_i(L) = \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i \tag{12}$$

This implies that government welfare (ignoring non-food consumption) is equal to

$$1 - \alpha \sum_{i=1}^{M+N} p_i \left(\overline{c} - \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i \right)^2 - \alpha\mu(1-\mu) \sum_{i=1}^{M} p_i\sigma_i^2. \tag{13}$$

[80] The fourth is Hainan, for which we do not have data on 1959 mortality rates.

Under fixed prices, (10) and the resource constraint of the economy implied by the substitution of (6) into (7) imply that if $i \in \{1, M\}$, then (12) holds, and if $i \in \{M+1, M+N\}$, then

$$
\begin{aligned}
c_i(H) &= \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i + \mu \sum_{i=1}^{M} p_i\sigma_i / \sum_{i=M+1}^{M+N} p_i \text{ and} \\
c_i(L) &= \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i - (1-\mu) \sum_{i=1}^{M} p_i\sigma_i / \sum_{i=M+1}^{M+N} p_i
\end{aligned}
$$

This implies that government welfare (ignoring non-food consumption) is equal to:

$$
1 - \alpha \sum_{i=1}^{M+N} p_i \left(\overline{c} - \sum_{i=1}^{M} p_i(\widehat{e}_i - \mu\sigma_i) / \sum_{i=1}^{M+N} p_i \right)^2 - \alpha\mu(1-\mu) \left(\sum_{i=1}^{M} p_i\sigma_i \right)^2 / \sum_{i=M+1}^{M+N} p_i. \tag{14}
$$

(13) exceeds (14) if and only if condition (11) holds. **Q.E.D.**

Stylized Example of Model

It is straightforward to see the implications of the model as μ, the probability of a drop in production, approaches zero. This would be the case for instance if the drop in production happens with sufficiently low probability that it is effectively ignored by the government, an approximation which may not be unreasonable in the case of the Great Famine since the drop in per capita production experienced in 1959 was very large by historical standards. In this case, the model predicts that $c_i(H)$ is equalized across households so that the level of procurement satisfies $\tau_i = \widehat{e}_i - \overline{c}$ for $\overline{c} = \overline{e} - \theta / \sum_{j=1}^{M+N} p_j$ where $\overline{e} = \sum_{j=1}^{M+N} p_j\widehat{e}_j / \sum_{j=1}^{M+N} p_j$. In other words, the government procures the difference between expected per capita production of a region and average per capita consumption in the economy, where this consumption level is determined by the difference between total per capita production and the per capita procurement target. Moreover, this implies that the consumption of region i in the event of an unexpected decline in production satisfies $c_i(L) = \overline{c} - \sigma_i$. So for example, suppose per capita production across all regions is 20% lower than expected (i.e., $\sigma_i = .2\widehat{e}_i$) and region i produces twice as much per capita relative to the society as a whole (i.e., $\widehat{e}_i = 2\overline{e}$). Moreover, let $\theta = .2\overline{e}\sum_{j=1}^{M+N} p_j$ so that the government designate 20% of grain production from the economy in normal times for non-food use. In this case, per capita consumption in region i actually declines by 50% and equals only about 50% of national per capita production in the downturn.[81] In sum, the combination of an inflexible policy, a large unexpected drop in production, and a high aggregate procurement target can together significantly amplify the mortality consequence of a drop in per capita production.

[81] Given the procurement target, $\overline{c} = .8\overline{e}$ so that region i's consumption in the downturn equals $(1 - .4/.8)\overline{c} = .4\overline{e}$.

Dynamic Extension of Model

This dynamic extension of the model serves as a robustness check for the results of the static model and it further illustrates how the positive and negative correlations between production and mortality depend on actual production relative to anticipated production.

Let $t=0,....,T$ represent time. Suppose that total per capita food production in region i in period t is e_{it} where

$$e_{it}=\begin{cases}\eta^{H}e_{it-1} & \text{with probability } 1-\mu\\ \eta^{L}e_{it-1} & \text{with probability } \mu\end{cases},$$

where $e_{it}>0$ for $i=\{1,...,M\}$ and $e_{it}=0$ for $i=\{M+1,...,M+N\}$. Every region has a baseline productivity which fluctuates over time, where all regions are subject to an aggregate proportional shock to productivity. Let $\eta_t=\{\eta^L,\eta^H\}$ correspond to the realization of the shock. e_{i-1} is given. Let

$$c_{it}=e_{it}-\tau_{it} \tag{15}$$

the per capita consumption of households in region i at date t. $\tau_{it}\gtreqless 0$ corresponds to procurement and it is inflexible and must be chosen prior to the realization of the shock in period t. Suppose that region i's population at t prior to the realization of the shock is p_{it-1} and following the realization of the shock is

$$p_{it}=N\pi\left(c_{it}\right)p_{it-1}. \tag{16}$$

Population following the shock at t is proportional to population prior to the shock, where the proportion is increasing in consumption at t and in the exogenous population growth rate $N>0$. Intuitively, a larger fraction of the population survives if per capita consumption is higher. N represents the underlying growth rate of the total population due to factors other than food consumption. p_{i-1} is given. The government balances its budget at every date by setting

$$\sum_{i=1}^{M+N}p_{it-1}\tau_{it}=0. \tag{17}$$

Suppose that the government values the life of an individual at t following the shock by $\beta^t\chi$, where $\beta\in(0,1)$ is a discount factor and $\chi>0$ represents the value of life. At every date t, the government can condition its policy on $h_t=\eta_0\times...\times\eta_{t-1}$, which represents the history of shocks experienced by the economy up to date $t-1$ which it has observed. Specifically, $e_{it}=\eta^H h_t e_{i-1}$ with probability $1-\mu$ and $e_{it}=\eta^L h_t e_{i-1}$ with probability μ, which implies that the government can form its expectation of e_{it} based on information it has received from realized production up to $t-1$. Thus, the government chooses a sequence of policies $\tau=\left\{\left\{\tau_i\left(h_t\right)\right\}_{i\in\{1,M+N\}}\right\}_{t=0}^{T}$ so as to maximize the following object:

$$E_0\sum_{t=0}^{T}\sum_{i=1}^{M+N}\beta^t p_{it}\chi \text{ s.t. } (15)-(17).$$

Given the complexity of the government's problem, suppose for simplicity that $\pi(\cdot)$ satisfies the following condition:[82]

$$\pi(c) = 1 - \exp(-\varphi c) \text{ for } \varphi > 0. \tag{18}$$

One can achieve an analogous result as in Proposition 1 for this dynamic economy, where the proof is similar to that of Proposition 1 and relies on arguments which utilize backward induction.

Proposition 3 *The policy of the government has the following features:*

1. *For every t, aggregate survival $\sum_{i=1}^{M+N} p_{it-1}\pi(c_{it})$ conditional on $\{h_t, \eta_t\}$ and $\{p_{it-1}\}_{i=1}^{M+N}$ is below that implied by an equal distribution of consumption at t,*

2. *Procurement τ_{it} is increasing in predicted productivity e_{it-1}, and*

3. *Regional survival $\pi(c_{it})$ is increasing in production e_{it} if $\eta_t = \eta^H$, and regional survival $\pi(c_{it})$ is decreasing in production e_{it} if $\eta_t = \eta^L$.*

Proof. We establish these results by backward induction. The results for period $t = T$ follow from the same arguments as in the proof of Proposition 1. Let $c_{it}(h_t, \eta_t) = e_{it}(h_t, \eta_t) - \tau_{it}(h_t)$, the equilibrium value of consumption, and let

$$V_{iT}(h_T) = (1-\mu)\pi\left(c_{iT}\left(h_T, \eta^H\right)\right)\chi + \mu\pi\left(c_{iT}\left(h_T, \eta^H\right)\right)$$

correspond to the expected continuation value associated with an individual in region i given information prior to the date t shock. From (18), it follows that $\pi'(c) = \exp(-\varphi c) = 1 - \pi(c)$. Given (9) which holds at date T, it follows that $V_{iT}(h_T)$ is the same across all i conditional on h_T. Now consider optimal policy at $T-1$ given that $V_{iT}(h_T)$ is the same for all i. The analogous first order condition to (9) in the dynamic economy is:

$$\begin{aligned}(1-\mu)\pi'\left(c_{iT-1}\left(h_{T-1}, \eta^H\right)\right)\left(\chi + \beta V_{iT}\left(h_{T-1}, \eta^H\right)\right) \\ + \mu\pi'\left(c_{iT-1}\left(h_{T-1}, \eta^L\right)\right)\left(\chi + \beta V_{iT}\left(h_{T-1}, \eta^L\right)\right) = \psi \ \forall i.\end{aligned} \tag{19}$$

It is clear that an equal distribution of consumption cannot be achieved given constraint (15) and the fact that e_{iT-1} varies proportionately across regions, and this establishes the first part of the proposition at $T-1$. Now consider two regions k and l with $e_{kT-2} > e_{lT-2}$. If $\tau_{kT-1} \leq \tau_{lT-1}$, then $c_{kT-1}(h_{T-1}, \eta_{T-1}) > c_{lT-1}(h_{T-1}, \eta_{T-1})$ for $\eta_{T-1} = \left\{\eta^L, \eta^H\right\}$, but given the concavity of $\pi(\cdot)$, this violates (19), which proves the second part of the proposition at $T-1$. If $c_{kT-1}\left(h_{T-1}, \eta^L\right) > c_{lT-1}\left(h_{T-1}, \eta^L\right)$, then by (15) $c_{kT-1}\left(h_{T-1}, \eta^H\right) > c_{lT-1}\left(h_{T-1}, \eta^H\right)$ which violates (19). Therefore, $c_{kT-1}\left(h_{T-1}, \eta^L\right) < c_{lT-1}\left(h_{T-1}, \eta^L\right)$, and satisfaction of (19) implies that

[82] We could alternatively allow for any arbitrary concave $\pi(\cdot)$ but with a myopic social planner and achieve the same results.

$c_{kT-1}\left(h_{T-1},\eta^H\right) > c_{lT-1}\left(h_{T-1},\eta^H\right)$, and this proves the third part of the proposition at $T-1$. Finally, define $V_{iT-1}(h_{T-1})$ analogously to $V_{iT}(h_T)$ so that it represents the continuation value to an individual in region i at $T-1$ starting from history h_{T-1}. It is straightforward to see given (19) and the fact that $\pi'(c) = 1-\pi(c)$ that $V_{iT-1}(h_{T-1})$ is equalized across i, so that backward induction on this argument proves the results for all t. ■

This dynamic extension shows that our results generalize to a dynamic economy in which the social planner values life of survivors for different generations. Moreover, it shows that our results critically depend not on the absolute level of production but on how the realized level of production differs from predicted production. This is highlighted by the fact that the sign of the correlation between regional production and survival does not depend on the level of total production, but on η_t which represents how today's production differs from yesterday's production.

Note that our analysis presumes that the government knows the true value of e_{i-1} for each region so that it is aware of each region's capability. In practice, one can imagine that the government instead has a noisy prior about the value of e_{i-1} and that it receives a noisy signal of the realization of e_{it}. In such an environment, the government knows h_t and tries to predict the value of e_{it} given its expectation of the value of e_{i-1}. As time passes, the government's expectation of e_{it} becomes more and more accurate and the government asymptotically learns the true value of e_{i-1}. What can happen in such an environment is that an extremely optimistic (and inaccurate) prior over the value of e_{i-1} together with extremely positive (and inaccurate) initial signals regarding the values of e_{it} can lead the government over-procure from rural regions in earlier periods relative to later periods. This would clearly amplify the mortality consequences of a drop in aggregate production. Details of such an extension are available upon request.

Table 1: A Stylized Example of Grain Procurement and Consumption

	Region A	Region B	City
Subsistence Needs	100	100	100
Production under High Shock (Probability 80%)	225	150	0
Production under Low Shock (Probability 20%)	180	120	0
Expected Production (0.8 x High + 0.2 x Low)	216	144	0
Expected Consumption	120	120	120
Procurement/Subsidy (Expected Production - Expected Consumption)	96	24	-120
Consumption under High Shock (High Production - Procurement)	129	126	120
Consumption under Low Shock (Low Production - Procurement)	84	96	120

Table 2: The Correlation between Grain Production and Mortality -- Historical Data

		Dependent Variable: Ln Number of Deaths in Year t+1				
		(1)	(2)	(3)	(4)	(5)
		Full Sample	Omit Autonomous	Omit Autonomous	Omit Autonomous	Omit Autonomous & 1949-53
A	Ln Grain Prod x 1959 Dummy	0.119	0.256	0.242	0.237	0.262
		(0.0432)	(0.0482)	(0.0436)	(0.0370)	(0.0485)
	Robust SE	*(0.123)*	*(0.0876)*	*(0.0795)*	*(0.0776)*	*(0.0876)*
B	Ln Grain Prod	-0.0182	-0.0619	-0.0929	-0.0523	-0.0731
		(0.0189)	(0.0192)	(0.0178)	(0.0353)	(0.0205)
	Robust SE	*(0.0205)*	*(0.0237)*	*(0.0174)*	*(0.0192)*	*(0.0412)*
	Ln Total Population	1.050	1.134	1.280	0.978	1.142
		(0.0221)	(0.0256)	(0.0255)	(0.0440)	(0.0277)
	Robust SE	*(0.0318)*	*(0.0224)*	*(0.0299)*	*(0.0544)*	*(0.0243)*
	Controls					
	Gov Exp on Public Goods	N	N	Y	N	N
	Province-Time Trends	N	N	N	Y	N
	Observations	1290	1055	1032	1055	968
	R-squared	0.946	0.931	0.944	0.961	0.927
	Joint A + B	0.101	0.194	0.149	0.185	0.189
	p-value	0.0287	0.0001	0.0011	0.0002	0.0002
	p-value (Robust)	*0.419*	*0.0271*	*0.0641*	*0.0120*	*0.0315*

All regressions control for total province population and year fixed effects. Robust standard errors are presented in italics. All regressions use a sample where Tibet, Hainan and Sichuan are excluded. In Columns (2)-(5), we also exclude all autonomous regions: Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai.

Table 3: The Correlation between Suitability for Grain Production and Birth Cohort Size -- Retrospective Data

		Dependent Variable: Ln Birth Cohort Size							
		Agricultural Households				Non Agricultural Households			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Full Sample	Omit Autonomous Provinces	Omit Autonomous Provinces	Omit Autonomous Provinces	Full Sample	Omit Autonomous Provinces	Omit Autonomous Provinces	Omit Autonomous Provinces
A	Grain Suit x Born 1959-61	-0.289	-0.251	-0.224	-0.126	-0.0560	-0.0383	-0.0104	0.0152
		(0.0495)	(0.0507)	(0.0513)	(0.0458)	(0.0414)	(0.0427)	(0.0432)	(0.0405)
B	Grain Suitability	0.0481	0.0418	0.0373	0.0210	0.00933	0.00638	0.00173	-0.00253
		(0.00825)	(0.00845)	(0.00856)	(0.00763)	(0.00690)	(0.00712)	(0.00721)	(0.00675)
	Controls								
	Province Time Trends	N	N	Y	N	N	N	Y	N
	Province FE * Year FE	N	N	N	Y	N	N	N	Y
	Observations	21420	17622	17622	17622	12006	10368	10368	10368
	Adjusted R-squared	0.904	0.907	0.910	0.925	0.873	0.873	0.877	0.881
	Joint A + B	-0.241	-0.209	-0.186	-0.105	-0.0466	-0.0319	-0.00865	0.0127
	p-value	0.00000	0.00000	0.00001	0.00591	0.17700	0.37000	0.81100	0.70700

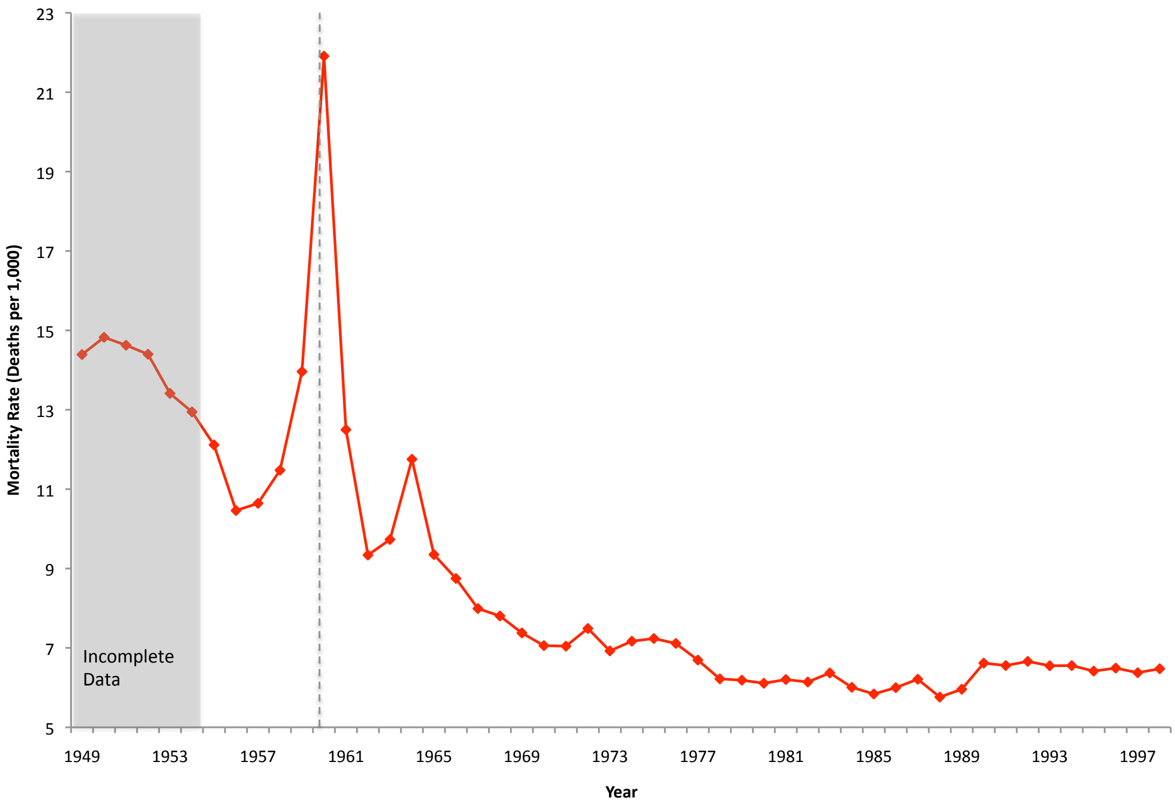
All regressions control for average log county birth cohort size and year fixed effects. Standard errors are clustered at the county level. In Columns (2)-(4) and (5)-(8), we exclude all autonomous regions: Tibet, Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai.

Table 4: The Correlation between Spring Precipitation and Birth Cohort Size

		Dependent Variable: Ln Cohort Size				
		(1)	(2)	(3)	(4)	(5)
		Full Weather Subsample	Full Weather Subsample	Full Weather Subsample	Counties with weather stations for 12+ years	Counties with weather stations for 12+ years
A	Grain Suitability	0.0857				
		(0.0374)				
B	Grain Suit x Born 1959-61	-0.414				
		(0.282)				
C	Ln Mean Spring Precipitation		0.0186	0.0177	0.0170	0.0167
			(0.00676)	(0.00693)	(0.00925)	(0.00941)
D	Ln Mean Spring Precipitation x Born 1959-61		-0.0960	-0.0884	-0.0825	-0.0775
			(0.0299)	(0.0309)	(0.0418)	(0.0412)
	Controls					
	Ln Mean Spring Temperature	N	N	Y	N	Y
	Ln Mean Spring Temperature x Born 1969-61	N	N	Y	N	Y
	Observations	1454	1454	1423	673	652
	Adjusted R-squared	0.909	0.909	0.910	0.894	0.897
	Joint A+B (Column 1); Joint C+D (Columns 2-5)	-0.328	-0.0774	-0.0707	-0.0655	-0.0608
	p-value	0.201	0.00334	0.00983	0.0745	0.0902

All regressions control for average ln cohort size and year fixed effects. Standard errors are clustered at the county level. The regressions in columns (1)-(3) uses the full sample of counties and year for which we have weather data. In Columns (4) and (5), we restrict the sample to the counties that have weather data for at least 12 out of the 18 years of our sample.

Figure 1: Average Mortality Rates Over Time

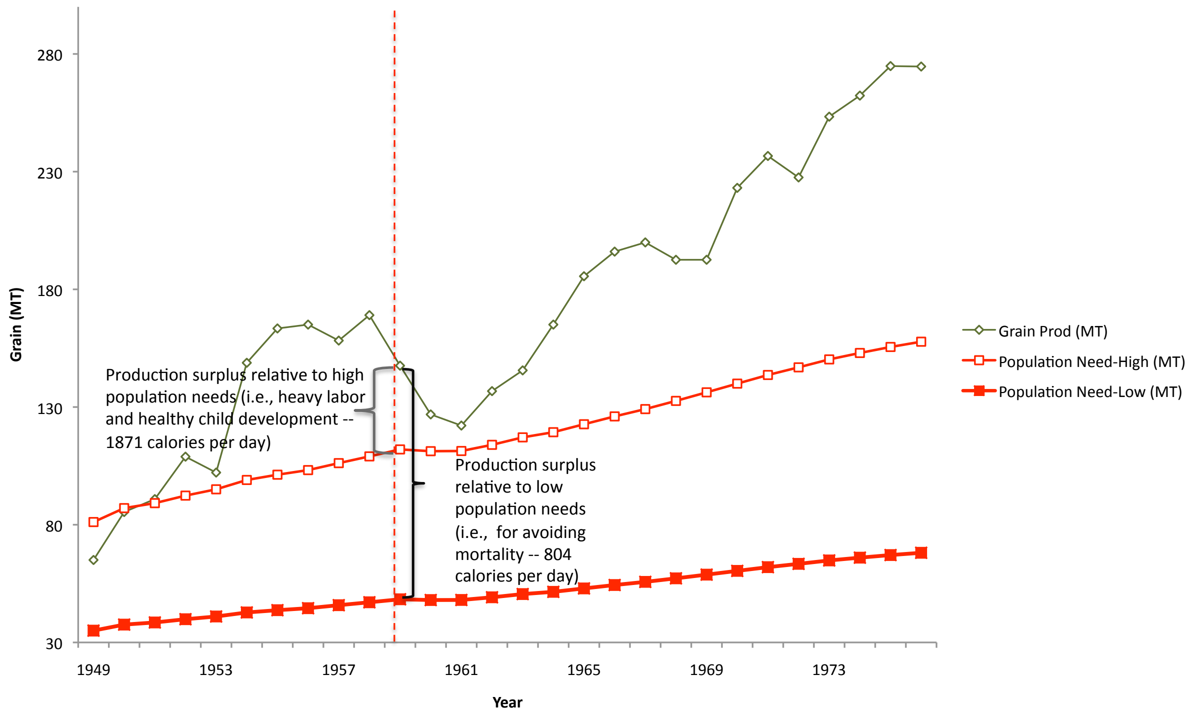


Source: CSDM50

Figure 2: Birth Cohort Size Over Time

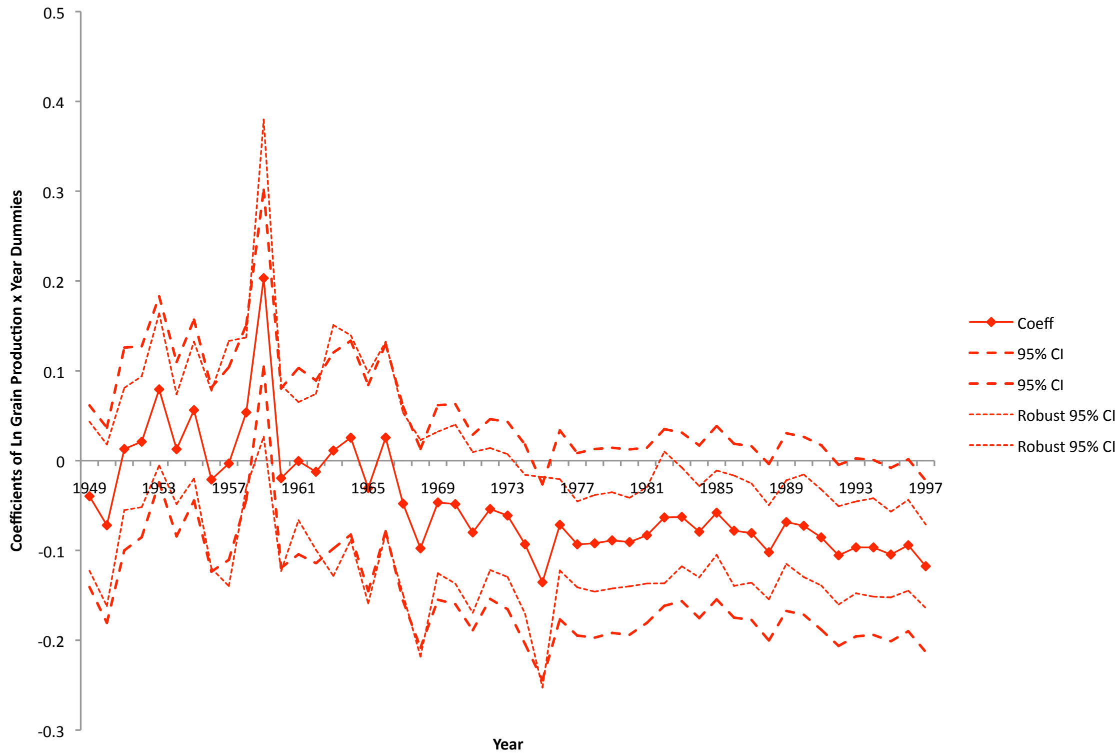
Source: 1990 Population Census

Figure 3: Aggregate Grain Production and Population Food Needs



Source: CSDM50 and Author's Computations

Figure 4A: Yearly Correlations Between Mortality and Grain Production and their 95% Confidence Intervals – The Coefficients of the interaction terms between grain production and year dummies controlling for total provincial population and year fixed effects.



Source: Authors' regression estimates. See Appendix Table A6.

Figure 4B: The Yearly Correlation between Birth Cohort Size and Suitability for Grain Production and their 95% Confidence Intervals – The coefficients of the interaction terms between suitability for grain production and year dummies controlling for average log county birth cohort size and birth year fixed effects.

Source: Authors' regression estimates. See Appendix Table A7 column (1),

Figure 4C: The Yearly Correlations between Birth Cohort Size and Spring Precipitation and their 95% Confidence Intervals – The coefficients of the interaction terms between ln average spring precipitation and year dummies controlling for average log county birth cohort size and birth year fixed effects.

Source: Authors' regression estimates. See Appendix Table A7 column (2).

APPENDIX Table A1: Descriptive Statistics

	A. Historical Province-Year Level Data 1949-98					
	Obs	Mean	Std. Err.			
Population (10,000 People)	1307	2938.39	54.11			
Mortality Rate (per 1,000 People)	1307	8.58	0.12			
Mortality Rate in 1960 (per 1,000 People)	1307	21.97	0.41			
Mortality Rate in 1954-57 (per 1,000 People)	1307	11.55	0.06			
Ratio of Mortality Rate 1960/Mortality Rate 1954-57	1307	1.84	0.03			
Fraction of Provinces where the Ratio of Mortality Rate 1960/ Death Rate 1954-57>1	1307	0.86	0.01			
Grain Production in 1959 (10,000 Tons)	1307	1009.29	23.69			
Annual Per Capital Grain Production (Kg Per Person)	1307	317.79	3.42			
Annual Per Capital Grain Production in 1959 (Kg Per Person)	1307	254.02	2.55			
Annual Per Capital Grain Production in 1954-57 (Kg Per Person)	1307	305.41	2.37			
Ratio of Per Capital Grain Production in 1959/1954-57	1307	0.83	0.00			
Fraction of Provinces where the Ratio in Per Capita Grain Production 1959/1954-57>1	1307	0.96	0.01			
	B. Retrospective County-Birth Year Level Data (1949-66)					
	Agricultural Households			Non-Agricultural Households		
	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.
Average Cohort Size Born in 1949-66 (1%)	21420	52.66	47.86	12006	39.55	45.00
Average Famine Cohort Size born in 1959-61 (1%)	21420	35.08	29.49	12006	32.77	36.29
Average Pre-Famine Cohort Size born in1954-57 (1%)	21420	53.14	44.88	12006	40.92	44.72
Ratio of Famine 1959-61/Pre-Famine 1954-57 Cohort Size	21420	0.71	0.25	12006	0.83	0.30
Fraction of Counties where Ratio of Famine 1959-61/Pre-Famine 1954-57 Cohort Size <1	21420	0.89	0.31	12006	0.76	0.43
Fraction of Land Suitable for Rice or Wheat Cultivation	21420	0.13	0.24	12006	0.20	0.30

In Panel A, the sample contains 27 provinces. Tibet, Sichuan and Hainan are omitted. Each observation is at the province-year level. In Panel B, the sample contains all 30 provinces. Observations are at the birth year - county level.

Sources: Panel A -- CSDM50 (1999), Panel B -- 1990 Population Census, GAEZ (2002)

Table A2: Historic Population Structure and Caloric Requirements

Age Bracket (1)	Population (100) (2)	Daily Caloric Needs (3)	Population Daily Caloric Need (4)	Average Daily Caloric Need (5)
	A. 1954 Caloric Needs for Heavy Agricultural Labor (or Healthy Child Development)			
Female				
0-5	495,641	1,300	64,433,330,000	
6-10	335,192	1,800	60,334,560,000	
11-15	294,474	2,200	64,784,280,000	
16-20	298,419	2,200	65,652,180,000	
21-50	1,055,377	1,800	189,967,860,000	
51-100	432,744	1,300	56,256,720,000	
Male				
0-5	542,455	1,300	70,519,150,000	
6-10	373,404	1,800	67,212,720,000	
11-15	347,053	2,500	86,763,250,000	
16-20	343,704	3,000	103,111,200,000	
21-50	1,165,685	2,100	244,793,850,000	
51-100	387,607	1,600	62,017,120,000	
Total	**6,071,755.00**		**1,135,846,220,000**	**1,870.70**
	B. 1954 Caloric Needs for Avoiding Mortality			
Female				
0-5	495,641	559	27,706,331,900	
6-10	335,192	774	25,943,860,800	
11-15	294,474	946	27,857,240,400	
16-20	298,419	946	28,230,437,400	
21-50	1,055,377	774	81,686,179,800	
51-100	432,744	559	24,190,389,600	
Male				
0-5	542,455	559	30,323,234,500	
6-10	373,404	774	28,901,469,600	
11-15	347,053	1,075	37,308,197,500	
16-20	343,704	1,290	44,337,816,000	
21-50	1,165,685	903	105,261,355,500	
51-100	387,607	688	26,667,361,600	
Total	**6,071,755.00**		**488,413,874,600**	**804.40**

Source: Coale (1981) and authors' computations.

Notes: Caloric requirements are calculated based on model from the USDA. In Panel A., for adults, we assume females 21-50 weigh 120 lbs, females 51-100 weigh 100lbs. Males 21-50 weigh 140 lbs, and 51-100 weigh 120 lbs. We assume that all adults 21-50 perform a high level of physical activity. And those 51-100 perform a medium level of physical activity. Caloric needs for staying alive are estimated to be 43% of those in Panel A. This is projected from the observation that an adult male labor need approximately 900 calories to stay alive, which is approximately 43% of the requirement for heavy physical labor.

Table A3: Grain Production and Population Caloric Requirements Over Time

National Production and Retention Over Time						
			190 kg/person, 1870 Calories		82 kg/person,804 Calories	
	Grain Prod	Population	Needed	Grain Surplus	Needed	Grain Surplus
	(Millions Tons)	(10000)	(Million Tons)	(Million Tons)	(Million Tons)	(Million Tons)
Year	(1)	(2)	(3)	(4)	(5)	(6)
1949	65.03	42708.391	81	-16	35	30
1950	85.41	45820.5	87	-2	38	48
1951	90.89	46933.66	89	2	38	52
1952	108.91	48617.371	92	17	40	69
1953	102.17	50028.059	95	7	41	61
1954	148.80	52116.09	99	50	43	106
1955	163.39	53294.91	101	62	44	120
1956	165.03	54323.5	103	62	45	120
1957	158.24	55881.781	106	52	46	112
1958	169.04	57388.539	109	60	47	122
1959	**147.58**	**58945.941**	**112**	**36**	**48**	**99**
1960	126.87	58551.832	111	16	48	79
1961	122.14	58592.98	111	11	48	74
1962	136.77	59972.609	114	23	49	88
1963	145.58	61641.711	117	28	51	95
1964	165.03	62797.738	119	46	51	114
1965	185.57	64575.449	123	63	53	133
1966	196.03	66330.219	126	70	54	142
1967	199.94	67963.57	129	71	56	144
1968	192.56	69795.711	133	60	57	135
1969	192.56	71711.383	136	56	59	134
1970	223.11	73658.367	140	83	60	163
1971	236.66	75589.594	144	93	62	175
1972	227.55	77297.68	147	81	63	164
1973	253.36	79072.508	150	103	65	189
1974	262.31	80510.727	153	109	66	196
1975	274.85	81853.5	156	119	67	208
1976	274.66	83030.188	158	117	68	207

Source: CDSM50 (1999), CPIRC (2000) and authors' computations.

Notes: Total production reported in column (1) is aggregate from province level production. The sample contains 27 provinces (Sichuan, Hainan and Tibet are omitted). Surplus in Columns (4) and (6) refer to production that is excess of subsistence needs. Average caloric needs in Columns (3) and (5) are computed using the national age distribution of population from the 1954 Census (see Coale, 1981). See Table A2. Based on estimates provided by the Ministry of Health and Hygiene of China, we assume that 1 kg of grain provides 3,587 calories.

Table A4: 1959 Regional Grain Production and Population Caloric Requirements

Famine Mortality and Production by Province				
Province	1960 Death Rate	1959 Grain Prod	1959 "Surplus"	
		Kg/Person	1,870 Calories	804 Calories
	(1)	(2)	(3)	(4)
Shanghai	6.9	107.02	-82.98	36.02
Beijing	9.14	82.01	-107.99	11.01
Neimeng	9.4	412.16	222.16	341.16
Jilin	10.13	401.07	211.07	330.07
Tianjin	10.34	91.42	-98.58	20.42
Heilongjiang	10.52	505.95	315.95	434.95
Shanxi	11.21	244.48	54.48	173.48
Liaoning	11.5	235.91	45.91	164.91
Zhejiang	11.88	382.06	192.06	311.06
Shan'xi	12.27	251.99	61.99	180.99
Ningxia	13.9	303.70	113.70	232.70
Guangdong	15.24	242.70	52.70	171.70
Xinjiang	15.67	304.35	114.35	233.35
Hebei	15.8	195.12	5.12	124.12
Jiangxi	16.06	314.36	124.36	243.36
Jiangshu	18.41	231.42	41.42	160.42
Fujian	20.7	259.23	69.23	188.23
Hubei	21.21	241.07	51.07	170.07
Shandong	23.6	195.24	5.24	124.24
Yunnan	26.26	265.26	75.26	194.26
Hunan	29.42	300.32	110.32	229.32
Guangxi	29.46	246.98	56.98	175.98
Henan	39.56	195.72	5.72	124.72
Qinghai	40.73	200.49	10.49	129.49
Gansu	41.32	223.95	33.95	152.95
Guizhou	52.33	242.67	52.67	171.67
Anhui	68.58	204.55	14.55	133.55
Hainan	*N/A*	*181.51*	*-8.49*	*110.51*
Tibet	*N/A*	*N/A*	*N/A*	*N/A*
Sichuan	*N/A*	*N/A*	*N/A*	*N/A*

Source: CSDM50 (1999) and authors' computations.

Notes: "Surplus" in Columns (3) and (4) refer to production that is excess of what is needed to work (for children, this refers to normal child development), and the excess of what is needed to stay alive. Average caloric needs in Columns (3) and (4) are computed using the national age distribution of population from the 1954 Census (see Coale, 1981). See Table A2. The three italicized provinces (Hainan, Tibet and Sichuan) are excluded from the computation of national production and caloric needs and all regression analsysis.

Table A5: The Correlation between Suitability for Grain Production and Birth Cohort Size -- Alternative Normalization Method

		Dependent Variable: Ln Number of Deaths in Year t+1				
		(1) Full Sample	(2) Omit Autonomous	(3) Omit Autonomous	(4) Omit Autonomous	(5) Omit Autonomous and 1949-53
A	Ln Grain Prod x 1959 Dummy	0.110	0.239	0.226	0.209	0.245
		(0.0458)	(0.0505)	(0.0468)	(0.0391)	(0.0511)
	Robust SE	*(0.113)*	*(0.0838)*	*(0.0762)*	*(0.0721)*	*(0.0838)*
B	Ln Grain Prod	0.0337	-0.0299	-0.0581	-0.0416	-0.0388
		(0.0200)	(0.0202)	(0.0191)	(0.0373)	(0.0217)
	Robust SE	*(0.0432)*	*(0.0301)*	*(0.0369)*	*(0.0434)*	*(0.0343)*
	Ln Total Population	0.982	1.084	1.220	0.748	1.090
		(0.0234)	(0.0269)	(0.0274)	(0.0465)	(0.0292)
	Robust SE	*(0.0556)*	*(0.0472)*	*(0.0667)*	*(0.169)*	*(0.0545)*
	Controls					
	Gov Exp on Public Goods	N	N	Y	N	N
	Province-Time Trends	N	N	N	Y	N
	Observations	1290	1055	1032	1055	968
	R-squared	0.938	0.923	0.935	0.955	0.918
	Joint A + B	0.144	0.209	0.168	0.168	0.206
	p-value	0.00316	7.44e-05	0.000640	0.00144	0.000129
	p-value (Robust)	0.231	0.0174	0.0463	0.0174	0.0214

All regressions control for total province population and year fixed effects. Robust standard errors are presented in italics. All regressions use a sample where Tibet, Hainan and Sichuan are excluded. In Columns (2)-(5), we also exclude all autonomous regions: Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai.

Table A6: Yearly Correlation between Per Capita Production and Mortality Rates

	Dependent Variable: Ln Mortality		
	(1)	(2)	(3)
	Coefficient	SE	Robust SE
Ln Grain Prod x Year = 1949	-0.0396	(0.0515)	(0.0423)
Ln Grain Prod x Year = 1950	-0.0720	(0.0554)	(0.0458)
Ln Grain Prod x Year = 1951	0.0129	(0.0575)	(0.0347)
Ln Grain Prod x Year = 1952	0.0210	(0.0542)	(0.0371)
Ln Grain Prod x Year = 1953	0.0793	(0.0528)	(0.0432)
Ln Grain Prod x Year = 1954	0.0127	(0.0494)	(0.0311)
Ln Grain Prod x Year = 1955	0.0563	(0.0513)	(0.0388)
Ln Grain Prod x Year = 1956	-0.0211	(0.0522)	(0.0504)
Ln Grain Prod x Year = 1957	-0.0032	(0.0546)	(0.0695)
Ln Grain Prod x Year = 1958	0.0537	(0.0497)	(0.0425)
Ln Grain Prod x Year = 1959	***0.2030***	***(0.0502)***	***(0.0900)***
Ln Grain Prod x Year = 1960	-0.0195	(0.0509)	(0.0528)
Ln Grain Prod x Year = 1961	-0.0005	(0.0529)	(0.0335)
Ln Grain Prod x Year = 1962	-0.0125	(0.0518)	(0.0442)
Ln Grain Prod x Year = 1963	0.0112	(0.0556)	(0.0712)
Ln Grain Prod x Year = 1964	0.0255	(0.0550)	(0.0581)
Ln Grain Prod x Year = 1965	-0.0306	(0.0585)	(0.0655)
Ln Grain Prod x Year = 1966	0.0256	(0.0531)	(0.0544)
Ln Grain Prod x Year = 1967	-0.0479	(0.0552)	(0.0516)
Ln Grain Prod x Year = 1968	-0.0977	(0.0564)	(0.0614)
Ln Grain Prod x Year = 1969	-0.0466	(0.0552)	(0.0402)
Ln Grain Prod x Year = 1970	-0.0484	(0.0568)	(0.0451)
Ln Grain Prod x Year = 1971	-0.0800	(0.0555)	(0.0456)
Ln Grain Prod x Year = 1972	-0.0539	(0.0510)	(0.0346)
Ln Grain Prod x Year = 1973	-0.0612	(0.0532)	(0.0348)
Ln Grain Prod x Year = 1974	-0.0931	(0.0566)	(0.0394)
Ln Grain Prod x Year = 1975	-0.1350	(0.0556)	(0.0597)
Ln Grain Prod x Year = 1976	-0.0715	(0.0536)	(0.0259)
Ln Grain Prod x Year = 1977	-0.0932	(0.0518)	(0.0244)
Ln Grain Prod x Year = 1978	-0.0921	(0.0535)	(0.0274)
Ln Grain Prod x Year = 1979	-0.0888	(0.0525)	(0.0273)
Ln Grain Prod x Year = 1980	-0.0907	(0.0526)	(0.0251)
Ln Grain Prod x Year = 1981	-0.0831	(0.0497)	(0.0274)
Ln Grain Prod x Year = 1982	-0.0633	(0.0501)	(0.0374)
Ln Grain Prod x Year = 1983	-0.0627	(0.0478)	(0.0280)
Ln Grain Prod x Year = 1984	-0.0792	(0.0490)	(0.0259)
Ln Grain Prod x Year = 1985	-0.0580	(0.0492)	(0.0239)
Ln Grain Prod x Year = 1986	-0.0780	(0.0493)	(0.0313)
Ln Grain Prod x Year = 1987	-0.0807	(0.0493)	(0.0281)
Ln Grain Prod x Year = 1988	-0.1020	(0.0501)	(0.0267)
Ln Grain Prod x Year = 1989	-0.0684	(0.0504)	(0.0237)
Ln Grain Prod x Year = 1990	-0.0726	(0.0505)	(0.0291)
Ln Grain Prod x Year = 1991	-0.0855	(0.0523)	(0.0272)
Ln Grain Prod x Year = 1992	-0.1050	(0.0514)	(0.0279)
Ln Grain Prod x Year = 1993	-0.0966	(0.0505)	(0.0260)
Ln Grain Prod x Year = 1994	-0.0967	(0.0496)	(0.0279)
Ln Grain Prod x Year = 1995	-0.1050	(0.0492)	(0.0243)
Ln Grain Prod x Year = 1996	-0.0942	(0.0488)	(0.0258)
Ln Grain Prod x Year = 1997	-0.1180	(0.0486)	(0.0237)
Observations		1055	
R-squared		0.935	

All regressions control for year fixed effects. Column (2) presents unadjusted standard errors. Column (3) presents robust standard errors. The sample excludes Tibet, Sichuan and Hainan, and all autonomous regions: Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai.

Table A7: The Yearly Correlations between Suitability for Grain Production and Birth Cohort Size, and Weather and Birth Cohort Size

Dependent Variable: Ln Cohort Size			
	(1) Agric HH Omitting Autonomous Regions		(2) Agric HH Full Weather Subsample
Ln Grain Suitability x Born Year 1949	0.0786		
	(0.0341)		
Ln Grain Suitability x Born Year 1950	0.0420		
	(0.0349)		
Ln Grain Suitability x Born Year 1951	0.208	Ln Mean Spring Precipitation x Born Year 1951	0.00615
	(0.0319)		(0.0817)
Ln Grain Suitability x Born Year 1952	0.188	Ln Mean Spring Precipitation x Born Year 1952	0.0345
	(0.0312)		(0.0426)
Ln Grain Suitability x Born Year 1953	0.181	Ln Mean Spring Precipitation x Born Year 1953	0.0841
	(0.0288)		(0.0344)
Ln Grain Suitability x Born Year 1954	0.144	Ln Mean Spring Precipitation x Born Year 1954	0.0122
	(0.0256)		(0.0218)
Ln Grain Suitability x Born Year 1955	-0.00454	Ln Mean Spring Precipitation x Born Year 1955	0.00749
	(0.0273)		(0.0297)
Ln Grain Suitability x Born Year 1956	0.0971	Ln Mean Spring Precipitation x Born Year 1956	0.0565
	(0.0272)		(0.0152)
Ln Grain Suitability x Born Year 1957	0.0439	Ln Mean Spring Precipitation x Born Year 1957	0.0331
	(0.0251)		(0.0198)
Ln Grain Suitability x Born Year 1958	-0.0934	Ln Mean Spring Precipitation x Born Year 1958	0.00527
	(0.0386)		(0.0314)
Ln Grain Suitability x Born Year 1959	-0.233	Ln Mean Spring Precipitation x Born Year 1959	-0.0472
	(0.0516)		(0.0325)
Ln Grain Suitability x Born Year 1960	-0.292	Ln Mean Spring Precipitation x Born Year 1960	-0.0871
	(0.0651)		(0.0467)
Ln Grain Suitability x Born Year 1961	-0.101	Ln Mean Spring Precipitation x Born Year 1961	-0.0934
	(0.0502)		(0.0349)
Ln Grain Suitability x Born Year 1962	-0.0450	Ln Mean Spring Precipitation x Born Year 1962	-0.0151
	(0.0339)		(0.0163)
Ln Grain Suitability x Born Year 1963	0.0438	Ln Mean Spring Precipitation x Born Year 1963	0.0505
	(0.0278)		(0.0227)
Ln Grain Suitability x Born Year 1964	-0.0736	Ln Mean Spring Precipitation x Born Year 1964	0.0477
	(0.0309)		(0.0345)
Ln Grain Suitability x Born Year 1965	-0.0669	Ln Mean Spring Precipitation x Born Year 1965	-0.000498
	(0.0293)		(0.0240)
Ln Grain Suitability x Born Year 1966	-0.117	Ln Mean Spring Precipitation x Born Year 1966	-0.0177
	(0.0429)		(0.0306)
Observations	17622	Observations	1443
R-squared	0.908	R-squared	0.911
F-Stat Grain Suitability x 1959 and Grain Suitability x 1960	12.84	F-Stat Ln MSPx1959 and Ln MSPx1960	2.292
p-val	0.000	p-val	0.1025

All regressions control for average ln county birth cohort size and birth year fixed effects. Standard errors are clustered at the county level. In columns (1), all autonomous regions (Tibet, Xinjiang, Qinghai, Guanxi, Ningxia and Neimeng) are omitted. Column (2) uses a subsample of all the county-level data for which we have weather data for. Birth cohorts from 1949 and 1950 are omitted from regression three because there are very few weather stations during these years.

Table A8: Historical Production and Procurement from Li and Yang (2005)

	Grain Prod			Retained Grain	Grain Procurement	
	Production (Millions Tons)	Annual Growth Rate	Growth Rate 4MA	(kg/agric laborer)	(Millions Tons)	% of Production
Year	(1)	(2)	(3)	(4)	(5)	(6)
1952	164			260	33	20.12%
1953	167	0.02		242	47	28.14%
1954	170	0.02		228	51	30.00%
1955	184	0.08		256	48	26.09%
1956	193	0.05	0.04	284	40	20.73%
1957	195	0.01	0.04	273	46	23.59%
1958	200	0.03	0.04	268	52	26.00%
1959	**170**	**-0.15**	**-0.02**	**193**	**64**	**37.65%**
1960	143	-0.16	-0.07	182	47	32.87%
1961	148	0.03	-0.06	209	37	25.00%
1962	160	0.08	-0.05	229	32	20.00%
1963	170	0.06	0.00	231	37	21.76%
1964	188	0.11	0.07	256	40	21.28%
1965	195	0.04	0.07	261	39	20.00%
1966	214	0.10	0.08	282	41	19.16%
1967	218	0.02	0.06	281	41	18.81%
1968	209	-0.04	0.03	261	40	19.14%
1969	211	0.01	0.02	259	38	18.01%
1970	240	0.14	0.03	282	46	19.17%
1971	250	0.04	0.04	293	44	17.60%
1972	241	-0.04	0.04	298	39	16.18%
1973	265	0.10	0.06	293	48	18.11%
1974	275	0.04	0.04	303	47	17.09%
1975	285	0.04	0.03	304	53	18.60%
1976	286	0.00	0.04	306	49	17.13%

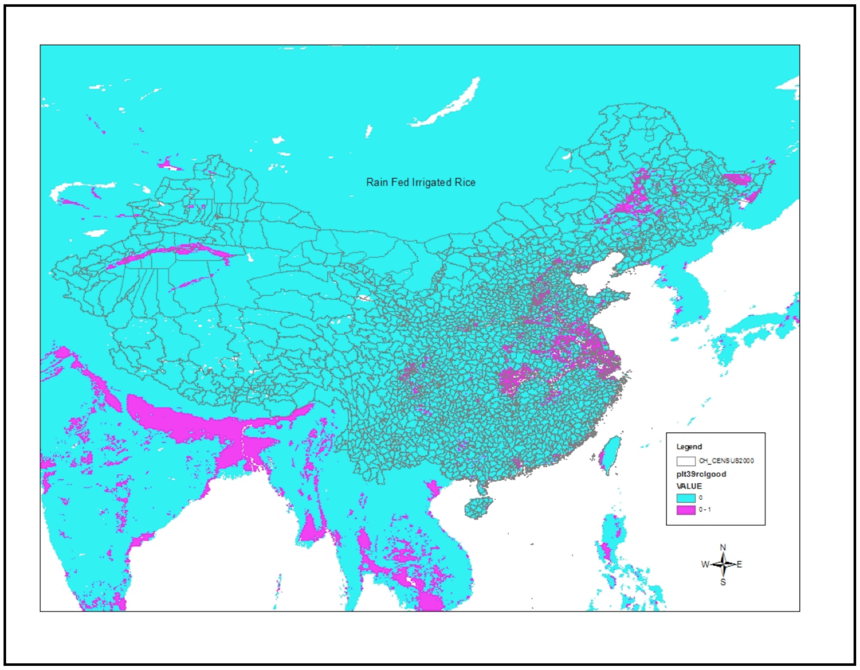
Source: Li and Yang (2005); Original Sources: Ministry of Agriculture (1989). Sample of 24 provinces.

Table A9: The Correlation between Grain Production and Mortality Controlling for Urban Population

		Dependent Variable: Ln Number of Deaths in Year t+1		
		(1) Baseline	(2)	(3)
A	Ln Grain Prod x 1959 Dummy	0.256	0.276	0.321
		(0.0876)	(0.0679)	(0.0594)
B	Ln Grain Prod x 1959 Dummy	-0.0619	-0.0264	-0.0295
		(0.0174)	(0.0157)	(0.0152)
C	Ln Urban Population x 1959 Dummy			-0.327
				(0.254)
D	Ln Urban Population		-0.251	-0.246
			(0.0136)	(0.0129)
	Observations	1055	1043	1043
	R-squared	0.931	0.953	0.953
	Joint A+B	0.194	0.249	0.291
	p-value	0.0271	0.0002	0.0000
	Joint C-D			-0.0813
	p-value			0.750

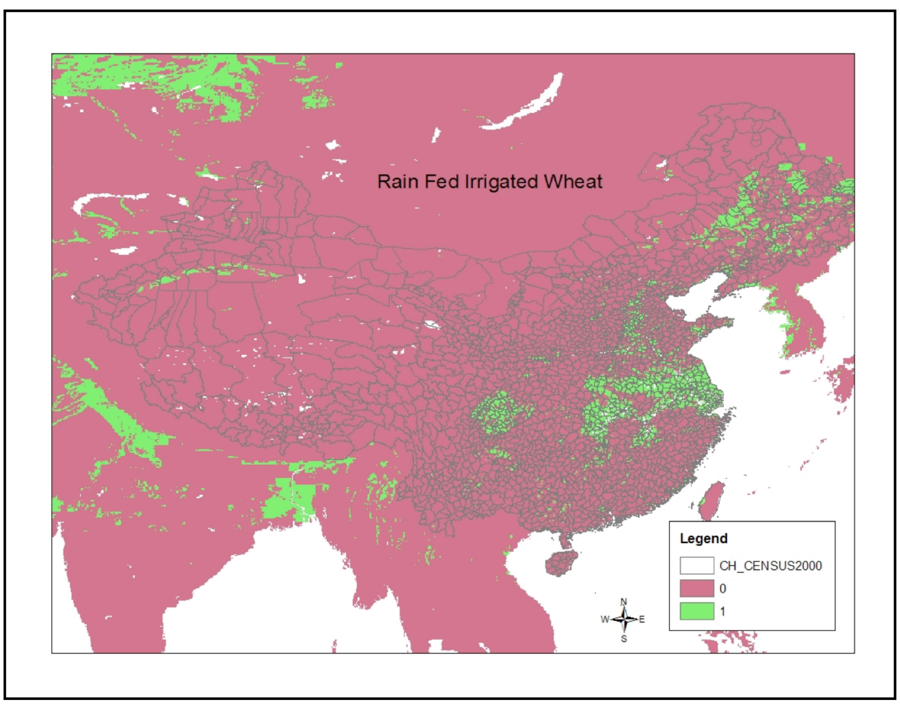
All regressions control for total province population and year fixed effects. Robust standard errors are presented in parentheses. All regressions use a sample where the autonomous regions -- Tibet, Xinjiang, Guangxi, Ningxia, Neimeng and Qinghai -- are excluded.

Appendix Figure A1A: Rice Suitability in China



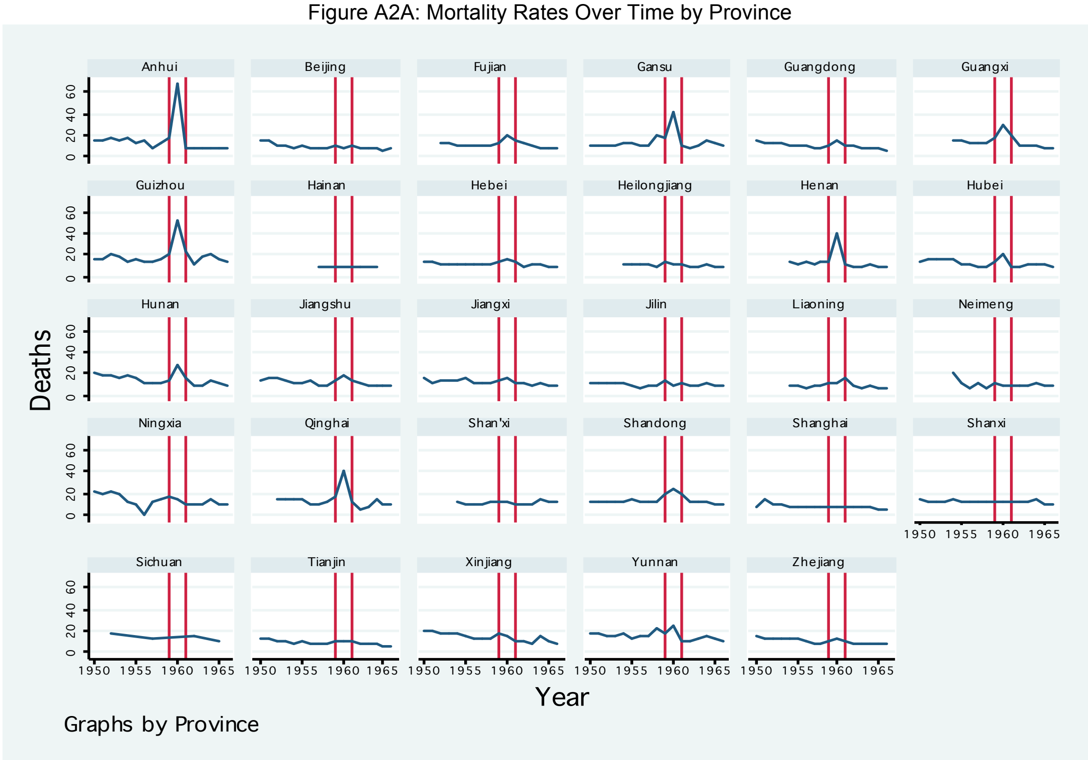
Source: Authors' computation.

Figure A1B: Wheat Suitability in China



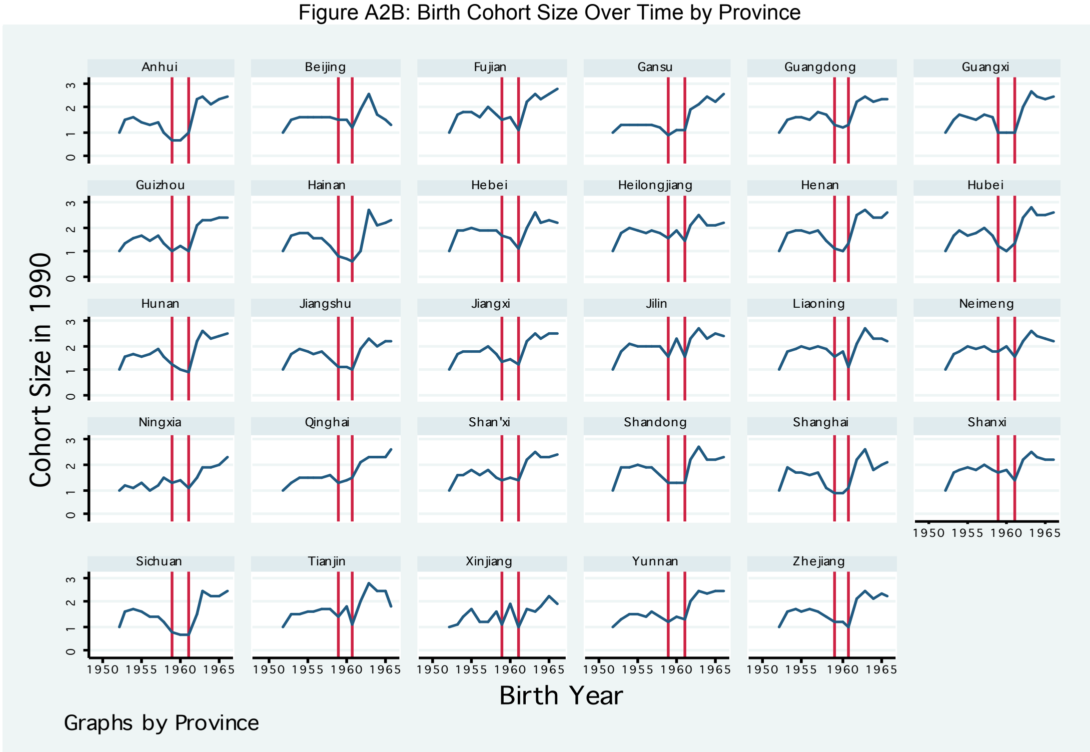
Source: Authors' computation.

Figure A2A: Mortality Rates Over Time by Province



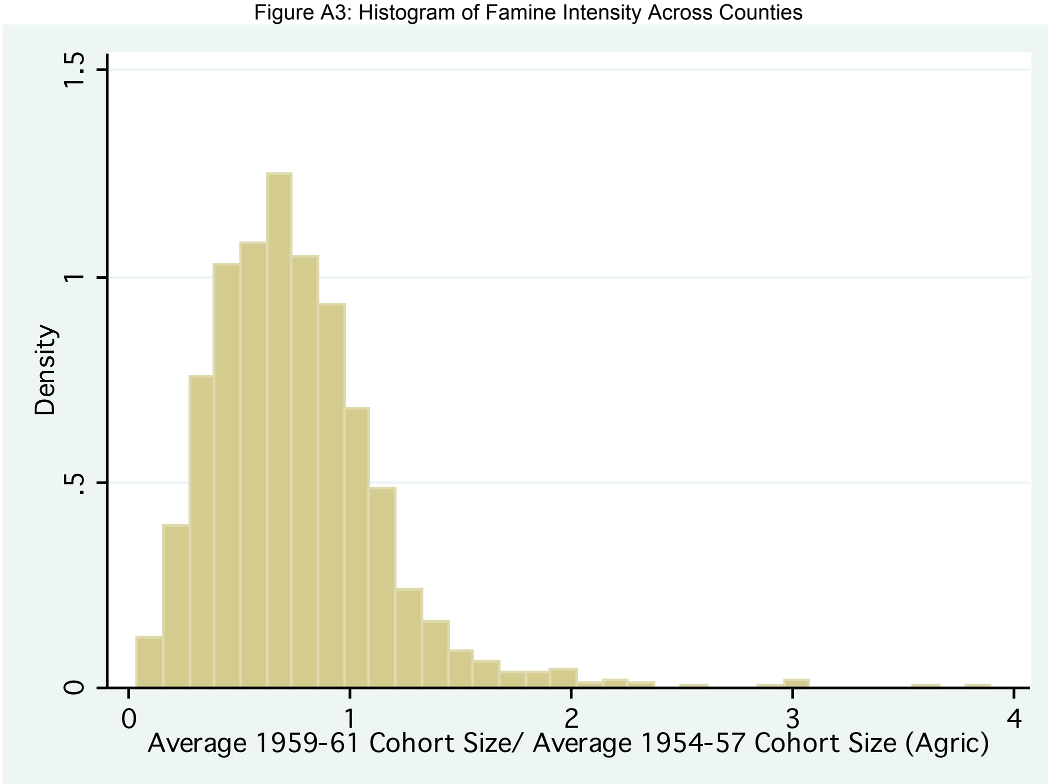
Source: CSDM50

Figure A2B: Birth Cohort Size Over Time by Province



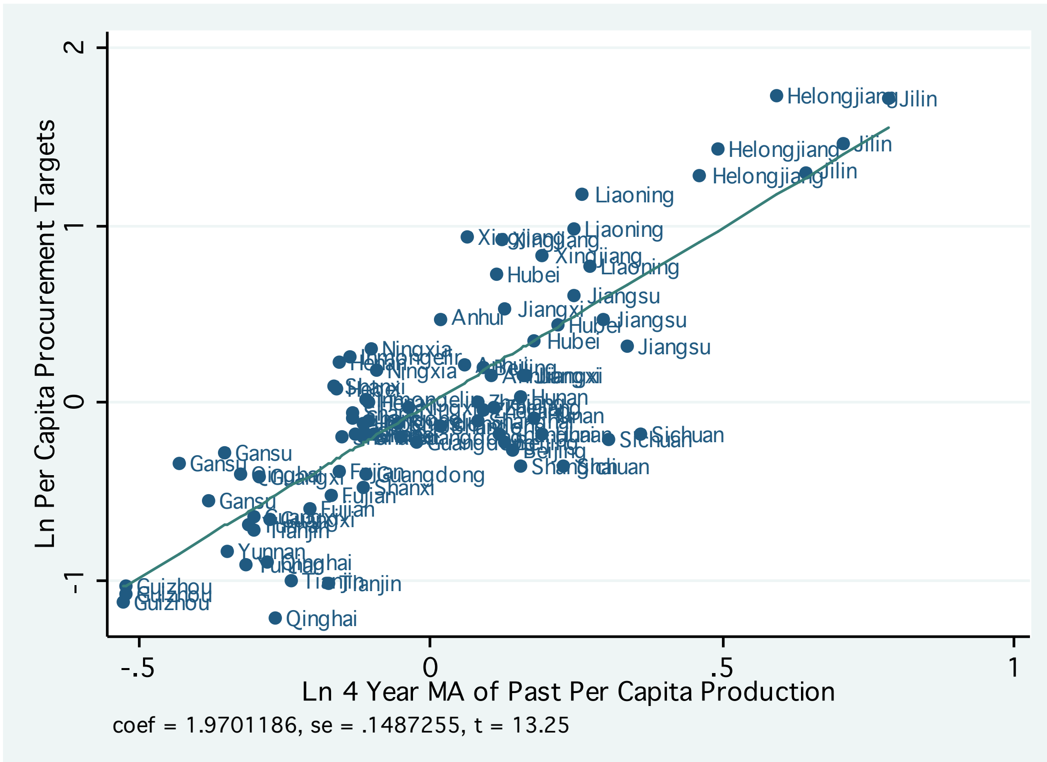
Source: 1990 Population Census

Figure A3: Histogram of Famine Intensity Across Counties



Source: Authors' computation.

Figure A4: The Correlation between Province Per Capita Grain Procurement Targets and a 4 Year Moving Average of Past Per Capita Production 1980-88



Source: Ministry of Agriculture