There Will Be Killing: Collectivization and Death of Draft Animals
(Online Appendix)

By SHUO CHEN and XIAOHUAN LAN

Appendix: Data Sources

Main Data

Our data sources consist of 20 declassified government files, 15 government reports (both internal and public), seven published compilations of statistics, and 1,720 county gazetteers. Appendix Figure A6 shows some pictures of these declassified files. We collect these documents from the National Library of China, university libraries in both mainland China and Hong Kong, the search engine Duxiu with full-text Chinese books for subscribers, and the Kongfuzi website, the largest online market in China for used or antique books and documents.

For 1,323 counties in 20 provinces, we find complete agricultural statistics compiled by the provincial bureaus of statistics or of agriculture. For a centrally planned economy in the 1950s, these statistics were indispensable to the planning committees and governments. There were three waves of provincial compilations of such statistics. The first wave was in 1958, which summarized the first five-year plan (1953-57). The second wave was in 1978-83, which reviewed the first three decades of the People’s Republic of China. The third wave was in 2009, which was compiled to celebrate the 60th anniversary of the republic.

Where provincial compilations are unavailable, we use prefectural compilations of county statistics. A prefecture is an administrative division under a province, and each prefecture consists of several counties. We find 19 prefectural compilations of statistics of 155 counties. When neither provincial nor prefectural compilations are available, we use individual county gazetteers. China has a thousand-year-long tradition of recording local history in gazetteers. The most recent gazetteers were published in the late 1980s and 1990s. Every gazetteer has a section on agriculture that documents relevant policies and some official statistics of agricultural production. Few gazetteers, however, report the annual inventory of draft animals in the 1950s. We find such statistics in 252 county gazetteers. Below are the data sources, all in Chinese, ordered by provincial administrative division code.

11 Beijing Municipality
The statistics are from county gazetteers.

12 Tianjin Municipality
Other statistics of agricultural production and population are from county gazetteers.

13 Hebei Province
Historical Agricultural Statistics in Chengde Prefecture: 1949-1978, the Agricultural Bureau of Hebei Province, the Agricultural Bureau of Chengde Prefecture, 1979
Economic Statistics in Qinhuaungdao Prefecture: 1949-1984, the Statistic Bureau of Qinhuaungdao Prefecture, 1985
The statistics of other counties are from county gazetteers.

14 Shanxi Province
Shanxi’s Economy: Cities and Counties; Shanxi Economy Press, 1992

15 Nei Mongol Autonomous Region

21 Liaoning Province
One Decade of Economic Achievements of Jinzhou Prefecture 1949-1958, the Statistics Bureau of Jinzhou Prefecture, 1959
The statistics of other counties are from county gazetteers.

22 Jilin Province

23 Heilongjiang Province

32 Jiangsu Province
Agricultural Statistics in Jiangsu Province: 1949-1975, Volumes I and II, the Agricultural Bureau of the Revolutionary Committee in Jiangsu Province, 1976

33 Zhejiang Province
Agricultural Statistics in Hangzhou Prefecture (1949-1973), the Agricultural Bureau of Huzhou Prefecture, 1974
Progressive Huzhou Prefecture: Thirty-five Years of Economic Statistics (1949-1984), the Statistic Bureau of Huzhou Prefecture, 1985
Vivid Wenzhou Prefecture: Four Decades of Great Economic and Social Achievements (1949-1988), the Statistic Bureau of Wenzhou Prefecture, 1989
Four Decades of Zhoushan Prefecture (1949-1988), the Statistic Bureau of Zhoushan Prefecture, 1989
The statistics of other counties are from county gazetteers.

34 Anhui Province
The statistics are from county gazetteers.

35 Fujian Province
Fujian Compendium of Statistics (the Agricultural Sector): 1950-1957, the Planning Committed of Fujian Province, the Agricultural Bureau of Fujian Province, and the Statistic Bureau of Fujian Province, 1958

36 Jiangxi Province

37 Shandong Province

41 Henan Province

42 Hubei Province
Agricultural Statistics in Hubei Province: 1949-1975, Volume II, the Agricultural Bureau of the Revolutionary Committee in Hubei Province, 1979
Agricultural Statistics in Hubei: 1949-1978, the Agricultural Bureau of Hubei Province, 1980

43 Hunan Province
Economic Statistics in Hunan Province (Section 2: Agriculture): 1949-1975, Volumes 3-5, the Statistics Bureau of Hunan Province, 1978
44 Guangdong Province
_Agricultural Statistics in Guangdong Province (1949-1981, by Cities and Counties)_ (Statistics Bureau of Guangdong Province, 1982)

45 Guangxi Zhuang Autonomous Region

46 Hainan Province
_Agricultural Statistics in Guangdong Province (1949-1981, by Cities and Counties)_ (Statistics Bureau of Guangdong Province, 1982)

In 1982, Hainan was still a part of Guangdong province.

50/51 Chongqing Municipality and Sichuan Province


The statistics of other counties are from county gazetteers.

52 Guizhou Province


_Fifty Years of Tongren Prefecture: 1949-2009_ (Editorial Committee of the Book, 1999)


The statistics of other counties are from county gazetteers.

53 Yunan Province
_Glorious Sixty Years of Yunan Province, Volumes of Economic Achievements_ (Statistics Bureau of Yunnan Province, Yunnan Press Corporation, Yunnan People Press, 2010)

61 Shaanxi Province

Weather Data

The historical local weather data are compiled by the State Meteorological Society, recorded in 267 weather stations as well as in county gazetteers. We assign these station records to their closest counties, based on the algorithm of Thiessen polygons. This method creates a polygon around each weather station, and these non-overlapping polygons cover all the counties. The counties closest to a station are the counties within the polygon of the station. The weather data use a discrete variable for rainfall: 1 for exceptional floods, 2 for limited floods, 3 for normal weather, 4 for limited droughts, and 5 for exceptional droughts.¹ We define two separate binary indicators for exceptional floods and exceptional droughts.

¹The variable is defined according to the descriptions in local gazetteers or the amount of precipitation, when available. Typical descriptions of events categorized as “exceptional floods” are “countless people and animals drowned in floods,” or “typhoons and heavy rains flood fields and houses,” etc. When the annual amount of precipitation is available, “exceptional floods” are the years in which the amount is higher than a 1.17 standard deviation above the mean.
This figure shows that from 1950 to 1957, among all the sown land, the portion allocated to grain decreased continuously. The cumulative decline, however, was modest - less than 4 percentage points. This decline was neither disrupted nor accelerated by collectivization.

Appendix Figure A2 The Distribution (in percentages) of Lands, Households, and Draft Animals, before and after the Land Reform Movement, by the Classes Assigned

Notes: Panel B shows some slight changes in the ratios before and after land reform. There were typically two or three waves of reform in the same county, with each wave addressing the issues of misclassification in the previous wave. As a result, the ratio of classes was adjusted a bit in each wave. In Panel C, data before the reform are not available.

Appendix Figure A3 The Dynamic Effects of Collectivization on log (draft animals)

This figure shows that collectivization, began at year 0, changes the growth pattern in the animal inventory. After two years since collectivization, the animal inventory drops by 10% from its peak in the year before collectivization (x=-1).

We regress log (draft animals) on the normalized year dummies and plot the coefficients. The dotted lines indicate the 95 percent confidence intervals. The reported coefficients reflect the changes in log (draft animals) relative to its level in the base year, the third to last year prior to collectivization. All regressions include log(rural population), log(arable lands), flood, drought, calendar year dummies, county-specific trends, and a constant.
GLF stands for the Great Leap Forward movement in 1958. “Adjustment” is the three-year “adjustment period” from 1962 to 1964.

*Source:* Ministry of Agriculture (1990)
The collectivization movement in the Soviet Union was between 1929 and 1933. Draft animals include horses and cows, excluding milk cows.  

Source: Li (1981)
The picture on the left shows the three volumes of the *Economic Statistics in Hunan Province (Section 2: Agriculture): 1949-1975*. The picture on the right highlights the classification level of the file, *Top Secret*, on the front cover.
## Appendix Table A1 Placebo Test: The Effects of Collectivization on Rural Population and Arable Land

<table>
<thead>
<tr>
<th></th>
<th>log (rural population)</th>
<th></th>
<th>log (arable land)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>collectivization</td>
<td>0.001</td>
<td>-0.000</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>flood</td>
<td>-0.008***</td>
<td>-0.002</td>
<td>-0.021***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>drought</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>constant</td>
<td>12.105***</td>
<td>12.082***</td>
<td>3.712***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>county FE*trend</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>N</td>
<td>8,826</td>
<td>8,826</td>
<td>8,623</td>
</tr>
</tbody>
</table>

Standard errors are clustered at the county level

***p<0.01

This table reports the results of a placebo test. It shows the difference-in-differences estimates of the effects of collectivization on log (rural population) and log (arable land). The collectivization dummy equals 1 for the year that a county began to establish advanced cooperatives and for the years after. All regressions include a set of year dummies and county dummies.
Appendix Table A2 The Effects of Collectivization on Log (draft animals): by Two Groups of Provinces

<table>
<thead>
<tr>
<th></th>
<th>Agricultural Machinery in 1958</th>
<th>Change of Crop Mix after Collectivization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) More</td>
<td>(2) Less</td>
</tr>
<tr>
<td>collectivization</td>
<td>-0.042***</td>
<td>-0.048***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>log(rural population)</td>
<td>-0.067</td>
<td>0.120*</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.069)</td>
</tr>
<tr>
<td>log(arable lands)</td>
<td>0.550***</td>
<td>0.270***</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>flood</td>
<td>-0.002</td>
<td>-0.012***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>drought</td>
<td>-0.000</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>N</td>
<td>1,002</td>
<td>6,995</td>
</tr>
</tbody>
</table>

Standard errors are clustered at the county level

* p<0.1, ** p<0.05, *** p<0.01

This table shows that the effects of collectivization on the animal inventory do not vary across two groups of provinces. Columns (1) and (2) divide the provinces according to the extent of their mechanization. According to the National Bureau of Statistics (1980), 21 percent of tilled areas in Beijing were tilled by machines in 1958, 17.4 percent in Heilongjiang, 15.4 percent in Xinjiang, 13.1 percent in Hebei, and 10.4 percent in Qinghai. In all other provinces in column (2), the percentages were lower than 4.2. Columns (3) and (4) divide the provinces according to the changes in the ratio of land sown with grain among all the sown land after collectivization. The changes are modest in all provinces, ranging from -0.04 to 0.01 (National Bureau of Statistics, 2010). The mean change is -0.01. The total number of observations in columns (3) and (4) is smaller than the total number of observations in columns (1) and (2), because data of crop mix are not available in Hainan province and Tianjin municipality. All regressions include year dummies, county dummies, and a constant.
Appendix Table A3: The Effects of Collectivization on Log (draft animals), in a Truncated Sample

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collectivization</td>
<td>-0.051***</td>
<td>-0.037***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Collectivization*ratio of middle peasant households†</td>
<td>-0.092*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td>Collectivization*log(number of people in a co-op)†</td>
<td>-0.028***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>Collectivization*log(distance from the capital city)†</td>
<td>-0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Collectivization*dummy for a revolutionary base</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>Collectivization*ratio of ethnic minorities</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td></td>
</tr>
</tbody>
</table>

| N                        | 6,144         | 3,134         |

Standard errors are clustered at the county level

*p<0.05, ***p<0.01

These regressions estimate the immediate effect of collectivization. We drop the observations in 1956 and 1957 for counties that started collectivizing in 1955, and the observations in 1957 for counties that started in 1956. All regressions include log(rural population), log(arable land), flood, drought, year dummies, county dummies, county-specific trends, and a constant.

† The variables are the deviation from their mean. This transformation does not affect the coefficients of the interaction terms, and the coefficients of the collectivization indicator are the effect at the mean level of these variables.
This table shows how collectivization affects the growth rate of the animal inventory, year on year. We regress $\Delta \log$ (draft animals) on a set of normalized year dummies. The reported coefficients reflect the changes in $\Delta \log$ (draft animals) relative to the base year. In the first column, the base year is the second to last year prior to collectivization. In the second column, we use an unbalanced panel data set in which some counties have data extending back to 1949. The base year is the fifth to last year prior to collectivization. All regressions include $\Delta \log$(rural population), $\Delta \log$(arable land), flood, drought, calendar year dummies, and a constant.