Title: Further exploration of the skewed sex ratio in China
Name: Sixue Liu
Date: May 2019

ABSTRACT

In this paper, I select the data of sex ratio at birth and the time of collectivization in different stages from 1949 to 1965 in China, including 5 provinces (46 counties) in total. To answer my research question: Did industrialization and agricultural collectivization normalize the skewed sex ratio in China? In order to answer this question: I focus on three main stages of the collectivization: mutual aid group of agricultural production; the primary agricultural production cooperative and the advanced agricultural production cooperative, and also their influence on the sex ratio at birth at that time. Based on the skewed sex ratio at birth in China now, this paper combines the awakening of feminism during the collectivization time, the basic social structure of transition from private ownership to collective ownership, and also the related variables in the historical development of agriculture and education, then using the sample OLS and fixed effect model to regress and analyze the multivariate variables. As a result, I find that when collectivization was carried out in the 50s of last century, especially before and after the establishment of the advanced agricultural production cooperative, the trend of the skewed sex ratio was relieved, and even a period of normal sex ratio at birth appeared in the biological sense. Therefore, I also put forward a new factor that could have infulence the gender selection: the movement of collectivization in the whole country in 1950s. In this paper, through my study of 46 counties in 5 privinces in mainland China, I found that the establishment of the advanced agricultural production cooperative could reduce the sex ratio at birth at around 3 percentage points.

KEY WORDS: Sex ratio; Collectivization; The advanced agricultural production cooperative; Fixed effect model

1. Initial Methods

Among the 46 counties in the five provinces I studied, each county started to implement the collectivization in a different period of time. Therefore, I use a fixed effect model to explore the effects of different stages in collectivation on the sex ratio at birth after capturing the characteristics of different county and year changes.

$$SRB_{it} = \beta \cdot collect_{it} + year_t + d_i + X_{it} + \varepsilon_{it}$$

For the county i at the year t, $collect_{it}$ is a dummy variable used to measure whether the various stages of the collectiviation are carried out. This means that the county developed the mutual aid group (the primary agricultural production cooperative /the advanced agricultural production cooperative) at the year t. For most of the counties I studied, the three main stages of the collectivization movement took place between 1951 and 1956. β is the coefficient I care most about, and it captures the changes in the sex ratio at birth in the collectivization period. Secondly, we use the $year_t$ and the county administrative division code (1986 Standard Edition) d_i to capture fixed features that didn't change with time and different counties. Finally, I will add some control variables X_{it} that changed with time and county level, including agricultural variables such as per capita arable land area, per capita agricultural output value and per capita food crop yield, and the number of primary school students per 1000, the number of junior high school students and the number of secondary schools. Those control variables are selected by feature contribution selection. Such educational variables are variables that may have an impact on the sex ratio at birth in that year. I use the robust standard errors clustered at the county level in the regression at the county level.

2. Data

2.1 Data source

The main data source is from China Gazetteer Project which is affilated to Harvard University's Department of Economics' China Economic History Big Data Research Project. This database collects the China's development in population, education, and agriculture of nearly 2000 counties from 1949 to 1990. The project aims to assess the role and impact of various government policies by digitizing China's county-level and city-level data and establishing a relatively complete Chinese socio-economic database, as well as studying economic and social changes in different parts of China using quantitative analysis. At present, the project has initially established a database covering 2,000 counties and cities for 65 years and a total of more than 120 variables. The data source of this project is the original county records of Harvard Yanching Library. There are two kinds of data sources for the county records: one is the data collected from the county and village data recorders (bottom to up), and the other is from the National Bureau of Statistics of China (up to bottom). In addition, in order to ensure the quality of the project data, we also arranged a lot remote county interviews to specifically interview the county editors, and raised our questions concerning about the data quality and also other problems. In summary, our data quality is relatively perfect and guaranteed. The main variables used in this paper are from the "population" part of the database, and some representative variables are selected from the other two parts "agriculture" and "education" as control variables.

2.2 Desciptive statistics

Our main dependent variable is the sex ratio of the county-level birth population from 1949 to 1965, based on the data published by the China Gazetteer Project and the United Nations (The sex ratios greater than 5 and less than 0.1 are removed. Extreme values may be due to manual input errors).

In the county-level data, the sex ratio at birth in China during the collectivization period is close to the normalization of biological significance. In 2005, the overall male-female ratio in China was 118.88:100. The sex ratio at birth in Jiangxi, Guangdong, Hainan, Anhui, and Henan provinces even exceeded 130:100. In most countries, the male-female ratio is between 97.0:100 and 100.3:100, while the world average is between 105-107:100. It's not hard to see, the male-female ratio in China has been highly distorted since the 1980s. This situation has continued to today (Loh and Remick, 2015). From this, we can see that it was indeed before the 1980s, that is to say, the sex ratio in the collectivization period was close to the normal range in the biological sense.

Table 2-1 shows the time at which each county was undergoing different stages of collectivization. At the same time, we can also see that the average number of population in each county is 269,700, which is almost the same as 255,100 people per county calculated in Chen and Lan (2017). And our data calculated the per capita arable land area is 0.189 hectares, the per capita agricultural output value is 0.048 million yuan and the total output is 0.456 tons per capita grain crop and the numbers in the paper by Li and Yang (2015) are 0.159, 0.023 and 0.347.

But the numbers I got are slightly larger than theirs. This may be because I have collected more data until 1965, so the total agricultural output, food crop production and total population had increased. Therefore, it can also be explained that the data I selected is more representative.

Table 2-1　Descriptive statistics of raw data

	Obersvations	Mean	S.d.	Min	Max
Mutal aid group	540	0.85	0.36	0	1
The primary cooperative	540	0.73	0.44	0	1
The advanced cooperative	540	0.63	0.48	0	1
Population (Thousand people)	878	269.7	204.7	7.75	998.6
Total cultivated area (Hectares)	527	50.98	50.26	7.45	645.32
Gross agricultural output value (Ten thousand yuan)	539	13000	4097.86	310	298,555
Grain crop yield (Ton)	527	123.08	180.23	8.57	195.01
Primary school students (Thousand people)	1006	36.86	39.14	0	400.83
High school student (Thousand people)	1023	3.30	5.46	0	52.46
Nuber of high schools （Number）	1037	16.54	43.61	0	364

3. Initial results

3.1 The influence of collectivation on gender ratio

Table 3-1 describes the statistical regression results. From columns (1) and (2), we can see that the presence of advanced agricultural production cooperative has a significant negative impact on gender ratio, regardless of whether or not the control variables are added. The primary agricultural production cooperative also has some influence, and the mutual aid group nearly has no impact. That is to say, at a significant level of 95%, the establishment of an advanced agricultural production cooperative can reduce the sex ratio at birth by about three percentage points. In column (3) of the table, I control the fixed effect at the county level. It can be clearly seen that the influence of advanced agricultural production cooperative on the sex ratio is still stable, and the influence of the primary agricultural production cooperative is also significant, but the significance level is lower. But it also shows that our results are not affected by a particular county. Similarly, I add the fixed effect of county level and year to the regression (i.e. column (5)). I found that the influence of the establishment of the advanced agricultural production cooperative on the sex ratio is still very significant, and there is almost no statistically significant difference between the previous results.

Table 3-1　Statistical results of the collectivation's influence on gender ratio

	（1）	（2）	（3）	（4）
Mutual aid group	0.051	0.035	-0.094	-0.072
Primary cooperative	-0.018	-0.021*	-0.039	-0.029
Advanced cooperative	-0.033*	-0.026**	-0.034**	-0.033**
Cultivated area per capita		0.001	0.062	0.052*
Agricultural output value per capita		0.069*	0.085	0.024*
Annual crop yield per capita		-0.001	-0.002	-0.001
Number of primary schoold student per thousand people		0.003*	0.002*	0.003*
Number of high school students per thousand people		0.129	0.007	0.006
Number of high		0.018	0.033*	0.036

school per thousand people				
Constant	9.923***	8.234	7.525	10.232
County-level fix effects	N	N	Y	Y
Year fix effects	N	N	N	Y
Observations	540	515	501	501

Note: * represents significant level at 10%; ** represents significant level at 5%; *** represents significant level at 1%.

3.2 Results from Shaanxi province

Table 3-2 describes the statistical regression results of Shaanxi Province. From the columns (1) and (2), we can see that whether or not the control variables like agriculture and education are added, the primary and the advanced agricultural production cooperatives both have significant negative impact on the sex ratio, which is consistent with the results of other provinces I have done. That is to say, the establishment of the advanced agricultural production cooperative in different counties will lower the male-female ratio at birth by about 2.8%, which is more normal in the biological way. In column (3) of the table, I still control the county-level fixed effect on the data of Shaanxi Province. It can be clearly seen that the influence of advanced agricultural production cooperative on the gender ratio is still stable, and the influence of the mutual aid group and the primary agricultural production cooperative are not. Obviously, that is to say, for Shaanxi Province, the time established by the advanced agricultural production cooperative also has a significant impact on the gender ratio, and at the same time it can also show that my results are not affected by a particular county. Similarly, I add the fixed effect of the year in the regression (i.e. column (4)) and the fixed effect of the county and the year at the same time (i.e. column (5)), I find that the establishment of the advanced agricultural production cooperative on gender ratio is still very significant, which is almost indistinguishable from the results of other provinces.

Table 3-2　Statistical results of the collectivation's influence on gender ratio in Shaaxi

	（1）	（2）	（3）	（4）
Mutual aid group	0.069	0.025	0.021	0.038
Primary cooperative	-0.046	-0.049*	-0.049	-0.039
Advanced	-0.038**	-0.037**	-0.039**	-0.038**

cooperative				
Cultivated area per capita		0.024*	0.028*	0.018
Agricultural output value per capita		0.163	0.054	0.127
Annual crop yield per capita		0.003	0.004*	0.002*
Number of primary schoold student per thousand people		0.001*	0.001	0.003
Number of high school students per thousand people		-0.004	-0.004	-0.003
Number of high school per thousand people		-0.092	-0.081*	-0.064*
Constant	5.812***	4.988	6.283	5.398
County-level fix effects	N	N	Y	Y
Year fix effects	N	N	N	Y
Observations	155	155	155	155

Note: * represents significant level at 10%; ** represents significant level at 5%; *** represents significant level at 1%.

In addition, we can visually see the effect of collectivization on gender ratio over time by the mean shift of β obtained by regression equation. If there is no influence of the collectivization, then the gender ratio at births in the 1950s should fall within a relatively stable range. That is to say, there is no major change before. Therefore, we now use the year in which the collectivization is carried out as a benchmark to observe whether the gender ratio at birth had changed.

I use ΔSRB_{it}, which is to measure the change in male-female ratio at birth by using the difference between this year and the previous year. Because the establishment of high-level agricultural production cooperatives has a greater impact on the birth ratio in Shaanxi Province, I only measure the changes before and after the establishment of the advanced agricultural production cooperatives:

$$\Delta SRB_{it} = \sum_{k} \beta_k \times I(CY_{it} = k) + year_t + X_{it} + e_{it}$$

I use ΔSRB_{it} to filter out the county-level fixed effect d_i in orginal regression equation. And I use a series of dummy variables $I(CY_{it} = k)$ which can be used to indicate the k_{th} year in the advanced agricultural production cooperatives; $CY_{it} = 0$ means that the advanced agricultural production cooperatives was established; and $CY_{it} = 1$ means the second year after the establishment of the senior society; $CY_{it} = -1$ means the year before the establishment of the advanced agricultural production cooperatives.

I now use the earliest year as the base year. If the male-female ratio at birth is not disturbed by any policy or social environment, then for $k < 0$, $\widehat{\beta_k}$ should be statistically no different from zero. However, due to the influence of the establishment of the advanced group on the sex ratio, $\widehat{\beta_0}$ and $\widehat{\beta_1}$ should be negative. X_{it} contains the difference between other control variables as a function of year. Figure 3-3 shows the estimates of $\widehat{\beta_{-1}}$, $\widehat{\beta_0}$ and $\widehat{\beta_1}$ in the regression and the 95% confidence level interval.
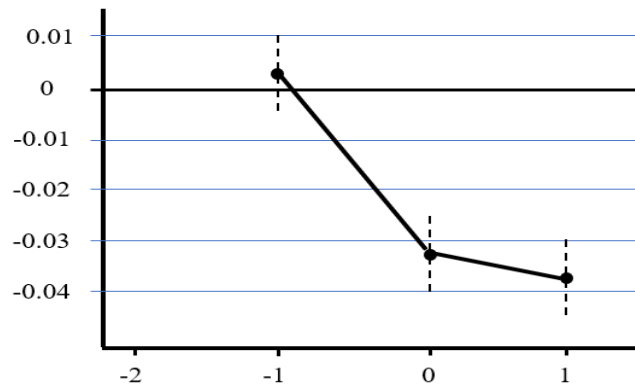

Figure 3-3   Dynamic change of the influence of collectivization on gender ratio

We can clearly see from the figure that the gender ratio at birth is maintained at an almost constant level without the influence of the establishment of the advanced group. In the years when the advanced agricultural production cooperatives was established and the years after it, there was a significant downward trend in the sex ratio. This downward trend is also consistent with the results in other provinces I have done before. In other words, without the influence of the establishement of advanced agricultural production cooperatives, the birth ratio will remain at this higher level.

Reference:

Charis, Loh, and E. J. Remick. "China's Skewed Sex Ratio and the One-Child Policy."
China Quarterly 222(2015):295-319.

Chen S, Lan X. There Will Be Killing: Collectivization and Death of Draft Animals[J].
American Economic Journal Applied Economics, 2017, 9(4):58-77.