

**China's Rural and Urban Household Survey Data:
Collection, Availability, and Problems**

Cheng Fang¹, Eric Wailes and Gail Cramer

CARD Working Paper 98-WP 202
Center for Agricultural and Rural Development
Iowa State University

**Presented at the Free Section entitled "Availability and Discrepancy of China's
Agricultural Statistics, 1998 AAEA Annual Meeting, Salt Lake City, August 5, 1998**

¹ Cheng Fang is an assistant scientist in the Center for Agricultural and Rural Development at Iowa State University, Eric. J. Wailes and Gail L. Cramer are professors in the Department of Agricultural Economics at University of Arkansas.

CHINA'S RURAL AND URBAN HOUSEHOLD SURVEY DATA: COLLECTION, AVAILABILITY, AND PROBLEMS

Since reform and open-door policies were implemented in the late 1970s, China has become a more important player in international markets. Increasing numbers of researchers are exploring China's agriculture. The data from China's rural and urban household survey are some of the most important data for researchers both inside and outside China. These data sets have been widely used to analyze rural and urban consumption behavior, savings and investment behaviors, and China's food projections. Researchers, however, still encounter a number of difficulties in using the household survey data. Problems of accuracy and availability have seriously limited its use, made the analysis inexact and the research results difficult to interpret. As a result, consumption parameters estimated in various studies widely different. Incomplete reported meat consumption data has caused difficulty in estimating China's meat output and feed grain demand.

To help explain China's household data, we discuss some issues about rural and urban household surveys. We first provide an overview of the history of household surveys in China and the sampling methods. Then we describe the content of the household survey, the main changes in recent years, the issue of data availability, and identify problems in data collection and use.

An Overview of the Urban and Rural Household Survey in China

History of the Household Survey in China

China began the survey of rural households in 1955 in order to provide information for the five-year plans for the Chinese economy and to check final results of these plans. To support the survey, rural statistical institutions were established within the State Statistical Bureau (SSB) and various local statistical bureaus within the county levels. Rural statistical divisions were also

set

up in related government agencies at various levels such as the Ministries of Agriculture and Forestry. The main statistics units are the production team, village, and community (township). The statistical system works by reporting through levels for the statistical bureau and statistical departments in other sectors of the economy.

During the 10-year Cultural Revolution, which started in 1966, the SSB was closed and the survey of rural households was suspended. In 1978, when rural economic reforms started, the SSB was reinstated, and the rural household survey resumed.

With rural economic reforms, families became the basic production units with implementation of the household responsibility system. The original information system could not efficiently collect sufficient data to meet the needs for the economic and policy analysis. In 1983, the General Organization for the Rural Socioeconomic Survey under the SSB was established to conduct rural surveys and collect agricultural and related statistics. Corresponding institutions and rural social and economic survey teams were also established in 30 provinces and 857 sample counties. The Rural Social and Economic Survey Organization in SSB is responsible for the design and implementation of plans for rural economic and social surveys, while the organizations at the provincial level conduct these surveys.

As in the case of the rural household survey, the surveys of urban households that started in 1956 were suspended from 1966 to 1979 and resumed in 1980 (He Jhuang 1985). The survey was conducted by the Price Statistical Division and the Division for the Survey of Workers' Family Life in the Department of Trade and Price Statistics under the SSB. In 1984, the Urban Social and Economic Survey Organization was set up. The corresponding survey teams for urban surveys were established in 30 provinces, 146 sample cities, and more than 80 counties that were selected for the social and economic surveys (Han, Wailes, and Cramer 1995).

The sample size for the rural household survey was increased each year from 1978 to

1985. The number of households surveyed in rural areas was 6,095 in 1978, and increased to 66,642 in 1985, and has been maintained at this level ever since. The number of urban households surveyed increased from 8,715 in 1981 to 32,855 in 1987 and has remained about the same over the last 10 years (Table 1).

The Sampling Method Used in the Household Survey

Households surveyed in the SSB rural survey are selected by using a three-stage stratified and systematic sampling method. Counties are chosen from each province; villages are drawn from counties; and households are selected from the villages. Using cumulative figures, such as population and desired sample size, the sample interval can be determined and units falling in the midpoint of each interval are included in the survey (Tuan and Crook 1984).

Two methods are used to select counties from provinces. One is the cumulative average of grain cropping area, calculated from the average grain yield per mu for the latest three years, assembled in ascending order, and using grain cropping area as cumulative figures of unit field. The other is the average net income per capita for the latest three years in ascending order, and using rural population as unit field. The indicator with the greater variation is selected as the key indicator for compiling the sample. The counties included in the sample are about 35 percent (857) of the total number of counties (SSB 1992).

Table 1. Number of households in survey

Year	Households surveyed	
	Urban	Rural
1957	5,350	17,378
1964	3,537	n.a.
1978		6,095
1979		10,282
1980		15,914
1981	8,715	18,529
1982	9,020	22,775
1983	9,060	30,427
1984	12,500	31,375
1985	24,338	66,642
1986	n.a.	66,836
1987	32,855	66,912
1988	34,945	67,186
1989	35,235	66,906
1990	35,660	66,960
1991	36,730	67,410
1992	36,290	67,490
1993	35,390	67,570
1994	34,940	67,420
1995	35,520	67,340
1996	36,370	67,610

Source: SSB various issues.

The samples selected for the rural household survey are the same set as those used for the agricultural production survey. There are three ways of sampling (Han 1996). The first method is based on the average yield per mu as the key indicator, and uses the accumulated average of grain cropping areas as the unit to determine the size. The second method is based on the average income per capita as the key indicator, and uses the accumulated average of grain cropping areas as the unit to determine the size. The third one is based on the average per capita income as the key indicator, and uses the accumulated average rural population as the unit to determine the size. The survey villages are fixed by the symmetrical equidistant sampling method up to the predefined number of samples. The sample includes 18 to 30 villages from each county, with the total including more than 20,000 villages in the national sample.

The rural households are selected from the village according to per capita net income as the key indicator, and accumulated number of residents in the village as the unit of size using a random starting point and symmetrical, equidistant sampling. Sample households are fixed to

the predefined number of samples.

The Content of the Rural Household Survey and Change

Including rural households in the survey provides feedback for government planning. The main contents of early rural household surveys were farmers' income and expenses. In 1984, some indicators that reflected the social and economic activities of rural households such as production variables, quantity of consumption goods, and general economic conditions of households, were added to the survey. In 1993, more rural marketing variables were added to the rural survey.

Rural household surveys before 1993 consisted of basic information on rural households, agricultural production and agricultural product sales, income and expenditures of households, cash balances, grain balances, purchases of commodities, consumption of in-kind commodities, and possession of durable goods. Table 2 summarizes the various major indicators in groups.

To reflect the information demanded by the Chinese government about macroscopic adjustment and control and microscopic activities under the economic reform and market economy system, the rural household survey was substantially amended in 1993. About 400 more variables were added to the survey questionnaire. The main additions include the prices of outputs and inputs in the rural household and more specific information about food consumption and nutrition.

The price section is completely new in the survey. The price information is grouped into two tables: output sales prices and input purchase prices. Both the quantity and value for sales are recorded for all commodities sold. The outputs sold include the main crops, livestock, and aquatic products. The main crops are wheat, rice, corn, sorghum, barley, soybeans, potatoes, millet, cotton, peanuts, sesame, rapeseed, cottonseed, tea, and sunflower. Three types of rice included in the survey are indica, japonica, and glutinous. Input purchases are itemized in some detail; for example, there are up to eight items under chemical fertilizers and seven items under pesticides. The data about fuel, machinery, and other farm inputs are also detailed.

Table 3 summarizes the changes in food consumption and nutrition before and after 1993.

Grain, beans, and bean products were not added to the survey until 1993. Livestock products, vegetables, and all other foods were described in greater detail after 1993.

Many new items were added to the consumption expenditures along with more detailed information for groups. For example, staple and nonstaple foods before 1993 are itemized into grain, beans, grain products, vegetables, bean products, fat and oil, sugar, meat and products, eggs, and aquatic products.

Table 2. The Content of the Rural Household Survey

Comprehensive Indicators			
geography		remote location	
geographical position		location of household	
household size		the highest cultural status of labor force	
type of household		household with a new house	
early liberated area		household using electric power	
Population and Labor			
number of permanent residents		illiterates and semi-illiterates	
number of staff and works in permanent residences			
Land, Major Fixed Assets for Production, House			
area of cultivated land		motor vehicles at year-end	
hilly fields		number of newly built houses within the year	
water areas		number of houses built at year-end	
original value of fixed assets			
Production and Sales for Plant and Forestry, Animal Husbandry			
sown area and yield by crop		output of livestock by species	
output of fruit by crop		quantity and value of sales for major commodities	
Grain Balances of Rural Households			
beginning inventory		purchases	
production		sales and inventory	
Total Income, Net Income, Total Expenditures			
net income		income by source	
expenditure for household production		collective production	
expenditure for production fixed assets		economic union	
taxes paid		household production	
expenditures for consumer goods		other nonproduction income	
other nonproduction expenditures			
Cash Balances of Rural Households			
beginning deposits		annual cash expenditure	
beginning cash on hand		ending cash on hand	
annual cash income		ending deposits	
Purchases of Commodities by the Rural Household			
food		goods for daily use	
cloth		cultural goods	
Per Capita Food Consumption			
vegetables	milk	sugar	candy
vegetable oil	poultry meat	liquor	fruit
animal fat	eggs	tea	cake
pork, beef, and mutton	fish and shrimp	cigarettes	
Number of Durable Consumer Goods Owned			
bicycles	sew machines	clocks	wrist watches
electric fans	washing machine	refrigerators	motorcycles
sofas	wardrobes	desks	radios
TVs	radio cassette players	cameras	

Table 3. Comparison of Indicators of Food Consumption and Nutrition Before and After

1993 Before 1993	1993 and After
No Grain	Grain wheat, rice, corn, sorghum, tubers
No Beans	Beans soybeans, others
No Bean Products	Bean Products
Vegetables	Vegetables Fresh Vegetables, 1-25 items Dried Vegetables, 1-5 items
Edible Vegetable Oil Edible Animal Oil	Edible Vegetable Oil Peanut oil, soybean oil, rapeseed oil, sesame oil Edible Animal Oil
Pork Beef and Mutton Poultry	Pork Beef and Mutton Poultry Meat related products
Eggs and Products	Eggs and Products
Liquor	Liquor and Beverages white liquor, beer, fruit wine, others, tea, other beverages
Milk and Products	Milk and Products
Fish and Shrimp	Fish, 1-5 kinds of fish Shrimp Other sea products
Sugar Candy	Sugar Candy Others
Fruit	Dried, Fresh Melons, and Fruits, 1-10 types Nuts and Grains

The Categories within the Urban Household Survey

The main content of the urban household survey includes the basic conditions of urban households, such as living expenditures for consumption, purchase of major commodities, durable consumer goods owned at year-end, durable consumer goods purchased, housing condition at year-end, and cash income and expenditures. The major indicators in each group are listed here.

Table 4. The Categories within the Urban Household Survey

Basic Conditions of Urban Households			
number of households surveyed	household size	number of employees per household	per capita annual income, etc.
Living Expenditures for Consumption			
food	daily articles	books and magazines	medicine
clothing	cultural and recreational articles	construction materials	fuel
other goods	noncommercial goods		
Purchases of Major Commodities			
grain, wheat, and rice	coarse grains	dried vegetables	pork, beef, and mutton
edible vegetable oil	fresh vegetables	dried vegetables	poultry meat
eggs	fish	sugar	cigarettes
liquor	fruit wine	beer	fresh melons and fruits
cake	milk	clothing	
Durable Consumer Goods Owned at Year-End			
clothing	furniture	bicycles	sewing machines
washing machines	refrigerators	TV sets	air conditioners
showers	dusters		
Durable Consumer Goods Purchased			
clothing	furniture	bicycles	sewing machines
washing machines	refrigerators	TV sets	air conditioners
showers	dust catchers		
Housing Conditions at Year-End			
area	water supply	washroom	
Cash Income and Expenditures			
disposable	wages of staff and workers in state-owned units	wages of staff and workers in collectively-owned units	wages of staff and workers in other ownership
other income of staff and workers from working units	transfer income	cash expenditures	

Availability of Household Survey Data

The SSB began publishing information from household surveys in 1980 (He Juhuang 1985). All published data are aggregated at both the national and regional levels. The urban household data are published in the SSB's *Chinese Urban Household Budget Surveys Yearbook*. The best source of detailed data is from this yearbook. Unfortunately, the SSB did not publish 1992 and 1994 yearbooks. Similar data for 1995 and 1996 were published in the *Chinese Price and Urban Household Budget Survey Yearbook*.

The *Chinese Urban Household Budget Surveys Yearbook* includes comprehensive data, data by region, and data by city. The national comprehensive data are reported for seven

different income groups. Each income group contains an equal percentage of the total households recorded. Comprehensive data are also reported for all cities, capital cities of each province, and capital city (township) of each county.

The more simple or grouped urban household data are published in the SSB's *Chinese Statistics Yearbook* and appear in Table 5. These data include basic indicators of the urban household, urban household annual living expenditures per capita, urban household annual per capita purchases of major commodities by level of income, urban household year-end possession of major durable consumer goods per 100 households by level of income, urban household year-end possession of major durable consumer goods per 100 households by region, urban household annual per capita sources of income by region, and urban household annual per capita living expenditures by region.

The detailed aggregated rural household survey data were published for only 1991 data, in the 1992 *Chinese Rural Household Budget Survey's Yearbook*. The same grouped data are published in the SSB's *Chinese Statistics Yearbook*, the SSB's *Chinese Rural Statistics Yearbook*, and the Ministry of Agriculture's *Chinese Statistics Yearbook*. The contents of SSB's *Chinese Statistics Yearbook* include basic indicators of the rural household, total and net income of the rural household, rural household net income per capita by source and region, rural household per capita living expenditures by region, rural household per capita living cash expenditures by region, rural household per capita consumption of major food by region, rural household per capita consumption of clothing by region, and rural household year-end possession of major durable consumer goods per 100 households by region. The rural household survey data published in *China's Agricultural Yearbook* and *Rural Statistics Yearbook* are basic indicators of rural households, per capita net income of rural households by region, per capita living expenditures of rural households by region, per capita living cash expenditures of rural households by region, per capita consumption of major food of rural household by region, and possession of main durable consumer goods per 100 farmers by region.

Published aggregated data are not adequate for many types of research projects. For example, aggregate data are not suitable for estimating certain types of equations (Rozelle 1995). In recent years, the SSB has adopted a more open policy to promote use of statistical data both

inside and outside China. However, these data are not provided free of charge (Chern 1994). The transaction costs associated with accessing the rural household data have been high (Wailes 1994).

Problems in Using Household Survey Data

In addition to the difficulty in accessing the data, there are other problems related to using the household survey data.

Inconsistent Data in Group Indicators

The reported data for some group indicators are inconsistent (Table 5). Before 1993, total expenditure was separated by commodity and noncommodity expenditures. However, these expenditures were combined. Most group indicators were regrouped and definitions were changed.

Incomplete Time Series and Incomplete Information for Consumption Analysis

More detailed aggregate data in some years were not published, which together with the definitional changes make it extremely difficult to find a long enough time period for analysis.

Regional consumption data for many food groups in the rural household survey were reported, but the corresponding expenditure data were not. On the contrary, the regional expenditure data for the main food group in the urban household survey were reported, but not for consumption data. For food demand analysis, the price indexes for food groups are generally used as a substitute. But detailed prices are not available, either.

Incomplete Survey of Food Consumption

There is a large discrepancy between meat production and consumption data in China because of incomplete food consumption data. Per capita food consumption tabulated from the survey data does not adequately account for away-from-home consumption. Away-from-home consumption includes eating in restaurants, food shops, in meetings, factory, school, and institutional cafeterias, and food for guests. Another missing category is for the migrant workers

who work far from home. These people are classified as rural residents, but live in urban areas. They have much higher incomes on average and also have much higher nonstaple food consumption. As the market economy develops, this part of consumption will become more important in total consumption. In the urban survey, the indicators only include food purchased, and all away-from-home consumption is excluded.

Table 5. Change in Category Definition of Group Indicator in Broad Group of Urban China

Before 1993		1993 and Later
Urban Household	Rural Household	Both Rural and Urban Household
Food	Food	Food
Clothing	Clothing	Clothing
Articles for daily use	Articles for daily use	Household facilities, articles and services
Housing and construction	Housing	Residence
Materials		
Fuel	Fuel	
Medicine and medical articles		Medicine and medical services
Books and magazines		
		Recreation and communications
Noncommodity expenditure	Noncommodities	
Rents, water, electricity, gas		
Transportation fee with city	Expenditure	
Other transportation fee		
Postage		Other commodities and services
Medical service fee		
Tuition and nursery		
Recreation fee		
Fix and maintain fee		
Other noncommodity fees		

Some Important Indicators Not Included in the Household Survey

Although reforms developed an indicator system reflecting social and economic changes since 1984 (especially in 1993), there is still a big difference between the survey indicator system and the need for analysis of the socioeconomic situation. Many important variables, such as those for feed consumption, are not included in the current survey.

Factors such as the feed conversion ratio and feed composition are a puzzle in China's livestock production analysis. Although some surveys were conducted by some institutions, such as the University of Arkansas and RCRE with support from CARD and USDA (Wailes et al.

1998), the sample is relatively small and also lacks adequate time series data for dynamic analysis. Only one indicator, total feed-grain consumption, can be found in existing questionnaires of the rural household survey.

Other important issues, which are still incomplete or are wholly missing, include agricultural labor force immigration, agricultural industry, rural transportation, and rural services.

Accuracy and Quality of the Household Survey Data

Administrative influence on the complete reporting system is likely to cause some bias in the collected data. Another problem is the lack of professional workers to conduct the survey. There is no statistical oversight entity, but rather part-time personnel assigned for statistics duties in basic rural units (townships or villages). There are still many things to do on sample frame selection, estimation, evaluation, and quality control. The sample methodology is being improved not only theoretically, but also practically.

Every family selected for a household survey has to sign an agreement with the government to guarantee their accuracy in recording their daily income and expenditures, and in auditing all monthly and quarterly records. However, since the income comes from many sources, some incomes are easily hidden. Thus, this coercion by government may cause the data to have errors (SSB 1991).

Concluding Remarks

There are more researchers studying China's food and agriculture economy. The SSB household survey data sets have become important tools for this research.

Both the rural and urban household surveys have been conducted for many years, except for the 10 years of the "Cultural Revolution." The survey samples are large, with more than 60,000 households in the rural survey, and more than 30,000 households in the urban survey. The sample covers about 3.5 of each 10,000 families. Households are selected using a multistage, stratified, and systematic sampling method. Chinese officials recognize that the procedure in selecting samples is not strictly in accordance with the principles of random sampling but claim that the impact is not significant (Tuan and Crook 1984).

The survey is designed to help serve government planning interests. Key contents in the initial rural household surveys were farmers' income and expenses. In 1984, some indicators that reflect the social and economic activities of the rural household such as production variables, quantity of consumption goods, and general economic conditions of households, were added to the survey. More variables reflecting rural marketing variables were also added, including data on prices in the rural household, more itemized data on food consumption and nutrition, and more information on rural household expenditures.

Access to data has continued to be a problem for researchers, especially for those outside China. Detailed household-level data are not easy to access. Aggregated data are also not published for all years and do not contain the same items across years. The problem is more serious for rural household survey data. Fortunately, access to data has greatly improved for most interested users. The cost, however, is generally high.

One solution to data access is to reinforce the cooperation between institutions in the United States and the SSB, and among institutions in the United States. China has adopted a more flexible policy to release the data to institutions outside of China by cooperative research and training agreements.

Another problem in using household survey data is the inconsistent group indicator having an incomplete time series. Researchers can use short-term data instead of whole period data or try to get some years of unpublished data to make the data consistent. To conduct

consumption analysis, the published price indexes can be used to construct the unpublished consumption data for urban areas or expenditure data for the rural survey at the provincial level.

Although the SSB has extended and added many new indicators to the household survey, many important data are still not included. For example, food consumption data are incomplete and feed-grain consumption data are completely absent. A long-term solution is to provide support to the SSB to add these items to the survey. A short-term solution is to collect the data by special surveys. The University of Arkansas has collaborated with the Research Centre for Rural Economy to conduct a livestock production survey. The data are available now. CARD is currently conducting a meat consumption study with the Chinese Academy of Agricultural Science. The data will be available soon. Other U.S. institutions are also collecting their own data from China now.

The quality of household survey data is in general is good. However, due to some statistical problems, administrative interference, and some political and social factors, errors in the survey are inevitable.

REFERENCES

- Chern, Wen S. 1994. "Consumption and Demand Data in the People's Republic of China." *China as a Market and Competitor* Proceedings of WRCC-101 Symposium #1, Reno, Nevada, November 21.
- Han, Tong. 1996. "Rural Household Food Consumption In China: Evidence from the Rural Household Survey." Unpublished M.S. thesis, University of Arkansas, Fayetteville.
- Han, Tong, Eric J. Wailes, and Gail L. Cramer. 1995. "Rural and Urban Data Collection in the People's Republic of China." In *The China Market Data and Information Systems*. Proceedings of WRCC-101 Symposium, Washington, D.C.
- He, Juhuang. 1985. "An Analysis of Chinese Resident's Demand." Paper presented to the Bilateral Australia-China workshop in the Role of Economic Modeling in National Economic Management, University of Melbourne, November.
- Rozelle, Scott. 1995. "Rural Household Data Collection in Reform China." *The China Market Data and Information Systems*, Proceedings of WRCC-101 Symposium.
- State Statistical Bureau (SSB) various issues. *Chinese Statistic Yearbook* Beijing: SSB.
- _____. various issues. "Questionnaire of Rural Household Survey." Beijing: SSB.
- _____. 1992. *1992 Chinese Rural Household Survey Yearbook*. Beijing: China Statistics Publishing House
- _____. various issues. *Chinese Urban Household Survey Yearbook*. Beijing: China Statistics Publishing House.
- Tuan, F.C. and F.W. Crook. 1984. "Planning and Statistical Systems in China's Agriculture" Foreign Agricultural Economic Report No. 181. Washington, DC: Economic Research Service, U.S. Department of Agriculture,.
- Wailes, Eric J., 1994. "The State Statistical Bureau and Other Data Collection Institutions." In *China as a Market and Competitor*. Proceedings of WRCC-101 Symposium #1, Reno, Nevada, November 21.

Wailes, Eric J., Cheng Fang, Zhang Xiaohui, Cao Liqun, Chen Hansheng, Wu Zhigang, and Guo Jianjun. 1998. AChina=s Livestock Feed Use Relationships: Preliminary Results from a Survey in Seven Provinces.@ Paper presented at the WRCC-101 Annual Meeting, Honolulu, Hawaii, January 12-13.