

Björn Gustafsson
Department of Social Work
University of Gothenburg
P.O. Box 720
SE 405 30 Göteborg
Sweden
And Institute for the Study of Labor (IZA)
Bonn, Germany
E-mail: Bjorn.Gustafsson@socwork.gu.se
Fax (+46) 31 786 1888
Phone (+46) 31 786 1890

Li Shi
School of Economics and Business Administration
Beijing Normal University
19 XinJie Kou WaiDajie
Beijing 100875
China
And Institute for the Study of Labor (IZA)
Bonn, Germany
E-mail: lishi@bnu.edu.cn

Hiroshi Sato
Graduate School of Economics
Hitotsubashi University
Tokyo, Japan
E-mail: sato.zuoteng@r.hit-u.ac.jp

December 2013

Data for studying earnings, the distribution of household income and poverty in China

Abstract

This paper discusses data used in publishing statistics on earnings, the distribution of household income and poverty in China by the National Bureau of Statistics (NBS) which is widely used by policy makers, international agencies and researchers. Unlike many other countries, China has had a dual system of household surveys with one rural system and one urban system up to now. This has consequences for providing official data on wages, income and poverty which we discuss along with other challenges. Since the end of the 1980s, researchers have been active in the construction of some larger databases aimed at mapping earnings, household income and poverty, which is also presented in the paper.

JEL Classification: C83 D31 I32 P36

Keywords: Earnings; Income; Inequality; Poverty; Data; China, National Bureau of Statistics

1. Introduction

How does China's rapid economic growth affect individual and household well-being? Looking at the big picture, there is consensus that after economic reform was initiated three decades ago, China's rapid economic growth has taken a very large number of households and persons out of poverty. There is also consensus that during the same period China has become a more unequal society. However, answers to more detailed questions are open for discussion and might depend on aspect, period studied and in some cases, data used. Examples of such questions are: How have earnings and earnings inequality among the urban population developed? How large is the gender wage gap and how has it developed? How large is the income gap between urban and rural areas, and between different regions of China? Which groups in the population are more likely to be poor and what kinds of changes in poverty over time can be observed? Researchers and other observers can offer some answers to these questions by consulting statistical information available in official publications and web pages, but the answers would be more specific and detailed if microdata from household surveys was made available. One task of this paper is to provide a brief introduction to the statistical system for monitoring earnings, income and poverty in China as well as to discuss some of its limitations and challenges.

Before reform began at the end of 1970s, researchers interesting in making empirical studies on earnings, the distribution of household income and poverty in China did not have access to microdata. Now, however, the situation has changed radically and a large number of studies based on microdata can be found in the literature. Furthermore, certain researchers have been given the opportunity to work with the microdata NBS has collected. Still another development has been that researchers have taken the initiative to collect surveys covering large parts of China. By putting these in the public domain they have become available for public use. The first academic studies on the topics we cover here were based on single cross-sections of data or repeated cross-sections of data. The latter alternative made it possible to address questions on changes over time. One highly interesting development concerning survey design is that of collecting panel data for households and individuals. This allows analysts to ask new questions and make new types of analyses such as dynamic studies of poverty aimed at understanding the duration of poverty and reasons for moving into and out of poverty. Panel data also puts the analyst in a better position for trying to find causes and effects of differences in earnings and income, and poverty. Another task of this paper is to introduce databases with microdata which researchers can access relatively easily for conducting empirical studies.

The rest of the paper is laid out as follows. The next section presents how the system of official statistics in China collects information on earnings and household income. Section 3 discusses some limitations and challenges of the system. The topic for Section 4 is microdata that is the outcome of researchers' activities. The paper ends with some concluding comments.

2. Data collected by NBS¹

Most countries in the world have official statistical systems which produce and publish information on the populations in which they operate. In the People's Republic of China the

¹ This and the next section build partly on Gustafsson and Li (2006). Cook and Keeley (2007) provide information on additional number of data sources which are not covered here.

NBS (National Bureau of Statistics, earlier known as the State Statistical Bureau, SSB) is responsible for country-wide reporting and has counterparts in each of the 23 provinces, 4 municipalities and 5 autonomous regions. There are also statistical authorities at the sub-province level. The province level statistical bureau is not only responsible for collecting data, but also for publishing information for the jurisdiction in which it is active. A researcher interested in the economic and social situation of one single province is therefore advised to consult the publications from the particular province level unit.

China differs from most other countries, however, in that institutional arrangements for rural and urban areas are rather different; the *Hukou* system has kept the two populations separated. People living in the cities live a different life than their peers in the rural areas. This is the background for why China for many years has had one Rural Household Survey covering rural areas and one Urban Household Survey for the urban areas - not a unified system. For these surveys the statistical bureaus in each province level unit draw one sample for rural households and another for urban households to produce the national statistics. Some Chinese provinces have a population the size of a country, so statistical bureaus in such provinces draw additional samples for producing more detailed information which is published in the statistical yearbook of the province.

Generally, the strategy of sampling households and individuals is to apply a multi-stage approach striving for the information to be representative at the province level. In rural areas counties are chosen after categorizing all counties in a province by average income. Choosing villages in each county follows the same method. In each village the target is to survey 10 households. Since 2010 not fewer than 74 000 households from 869 counties have been chosen (NBS, 2011). A household chosen for the rural survey is included for a period of five years typically after which it is rotated out. The rural households are not only asked to provide detailed income information, but also to keep detailed records of their expenditures as well. While the rural survey assembles information on the demographic characteristics and the education of each household member, it asks each household member questions on earnings which can only be attributed to him/her. Many thousands of assistant enumerators are involved in helping the households keep good records.

Household samples for the Urban Household Survey are drawn from a large sampling frame of households having an urban *Hukou*. A two-stage stratified systematic random sampling scheme is applied. In the first stage, cities and counties are classified into five categories by population size. They are then grouped into six administrative regions. In each administrative region the cities and counties are arranged by average wages of staff and workers in the urban area. Based on this classification, cities and counties are selected by a systematic sampling scheme. At the second stage, sample households are selected. This results in 66 000 households in 474 cities being selected by NBS (NBS, 2011).

Households selected for the Urban Household Survey typically remain for a period of three years after which they are rotated out. As in rural areas the household is regularly visited by an enumerator who asks questions and assists in bookkeeping. As in urban areas, the respondent and his or her household has to record all income and expenditures in a very detailed way. For example questions on earnings are put to each person in the household and the answers constitute the basis for tables published by NBS. However, as is true for the rural survey, the urban survey does not include questions on household wealth.

The system of separate rural and urban surveys was used in China until the end of 2012 when a new system was initiated. Our understanding is that there have been changes in the sampling procedure so the description above refers to the most recent years prior to 2013.² In addition, the income concept adopted by NBS when preparing the published tables has been broadened over time - a fact that should be regarded as progress. However, as a side effect, comparisons over time might have been affected. Thus one (unknown) part of the recorded increase in average income is not real, but due to changes to a more improved definition of disposable income.

The major challenge to the separate rural and urban survey systems is rapid development of rural-urban migration. In order to deal with increasing mobility of the population, NBS launched the National Rural-Urban Migrant Monitoring Survey (*quanguo nongmingong jiance diaocha*) at the end of 2008. This survey is a very , quarterly-based sampling survey of the rural labour force. The survey's purpose is to capture out-migration flow at the place of origin. Its sampling frame includes more than 7 500 villages in 899 counties in 31 province level administrative units and covers approximately 200 000 rural labourers. Aggregated data containing volume of out-migrants, distributions of places of origin and destinations, and individual characteristics of out-migrants such as gender, age, and education are published annually (NBS 2012).

It is this kind of information collected in the statistical system that is used for tables in the statistical yearbooks of China. For example, the widely cited rates of poverty in rural China are obtained by comparing the household's income with a cutoff. The poverty line was long set at a very low level, for example at 627 Yuan per person per year in 2002. Starting in 2000 NBS also applied a "low income line" which for 2002 was set at 869 Yuan per person per year. A new poverty line set at 2 300 Yuan per person per year was introduced in 2011. This poverty line is close to the 1.75 dollar per day poverty line in ppp terms.³ In contrast, in its monitoring of urban poverty, NBS does not regularly publish poverty rates for urban China and there are no official statistics on the extent of poverty in China as a whole.

Monitoring of rural poverty by NBS has also been conducted as the Rural Poverty Monitoring Survey (*nongcun pinkun jiance diaocha*) started in 1997. This survey focuses on all nationally designated poor counties (592 counties) and covers 54 000 sample households distributed throughout 5 400 villages. Aggregated data on household income and expenditures and poverty-related indicators such as poverty head count ratio have been published annually in the Rural Poverty Monitoring Report (Rural Survey Department of NBS 2011). Although sample households of this survey are not representative of the entire rural poor population, the survey is useful for tracing changes in economic conditions of households living in counties with special policy treatment.⁴

The National Bureau of Statistics also produces other data and publishes statistics based on it that are of relevance for researchers interested in people's well-being. Since the mid-1980s NBS has conducted large-scale population surveys during inter-census years, typically in the 5th year after a population census. The 2005 One-Percent Population Survey is the latest such

² For information on earlier years see Chen and Ravallion (1996:26-31) and Bramall (2001) who assess the quality of China's household surveys.

³ However, note that while NBS bases its estimate on rural poverty on household income, when assessing poverty in China the World Bank typically makes its assessments based on household consumption.

⁴ For a detailed introduction of the Rural Poverty Monitoring Survey and empirical analyses of rural poverty using the survey, see Yue et al. (2007).

survey. It includes more detailed information related to demographic, geographic, economic and housing information than the censuses provide. This survey is well developed when it comes to migrations flows, but the income information is limited.

Some countries have conducted regular labour force surveys for many years, from which statistics on employment rates and unemployment rates are regularly derived. China does not belong to such a category and no long time series on the unemployment rate based on household surveys in urban areas has been published. However, international scholars in cooperation with the Institute of Population and Labour Economics at the Chinese Academy of Social Sciences have in cooperation with local NBS carried out the China Urban Labour Survey (CULS) for the years 2001, 2005 and 2010. These surveys cover each of the years' five provincial capital cities (Shanghai, Wuhan, Shenyang, Fuzhou and Xi'an). In 2005 seven municipal cities were added (Wuxi, Yichuang, Benxi, Zhuhai, Shenzhen, Baoji and Daqing), and Guangzhou was added to the 2010 survey. Separate samples of persons with urban residence status and with migrants were derived using a two-stage procedure by area, then households were selected and interviewed.⁵

For researchers the strategy of working with data already collected by NBS can be advantageous in many respects. Statistical bureaus typically have a long-term commitment to their work leading them to be experts at all stages of the data gathering process and thus have an advantage over individual researchers or research groups. Added to this is the fact that the NBS statistical system covers the entire country. In many countries, a division of labour exists between the statistical agency collecting data and publishing descriptive reports and persons in academia who work with deeper analyses. Where this occurs, routines are in place for giving researchers access to the data collected by the statistical agency. However, such a description does not fit China where requests for access to microdata appears to be handled on a case by case basis. Initially, access was typically given to only a small number of regions, making it difficult to judge to what extent results could be generalised to large parts of China.⁶ However, more recently there are examples of research groups that have been successful in gaining access to microdata from NBS covering large parts of China for a substantial number of years.⁷

Other public agencies in China also produce statistics of relevance for judging the well-being of Chinese households and their members. Among such data most frequently referred to by researchers is the Rural Fixed Observation Points Survey (*nongcun guding guanchadian*) conducted by the Ministry of Agriculture since 1986.⁸ This survey is a longitudinal survey of villages as well as rural households and currently covers 23 000 households living in 360 villages distributed in 30 province level administrative units excluding Tibet. **The sampling frame of the survey is as follows: firstly, counties belonging to high-, middle-, and lower-income categories are selected and, secondly, "representative" villages are selected in each sample county. Then 40-120 households are selected in each village depending on village size. In addition to regular annual surveys on households and villages collecting basic**

⁵For further information see Giles et al (2005); see also Liu (2012) who uses CHIP-data presented in the next section to map labour force participation and unemployment in urban China.

⁶ See for example Tsui (1998), Yang (1999), McCulloch and Calandrino (2003), Kung and Lee (2001), Aaberge and Li (1997), Aaberge and Zhu (2001), Coady and Wang (2000) and Fang et al (2002).

⁷ See for example Zhang et al (2005), Li et al (2006), Chi and Li (2008) Meng (2012) and Meng et al (2013).

⁸ From the early 1990s, the survey is jointly managed by the Ministry of Agriculture and the Policy Research Office of the Central Committee of the Communist Party of China. The survey is managed by the Research Center for Rural Economy (RCRE) at the Ministry of Agriculture.

information such as land utilization, labour inputs, agricultural production, income, and expenditures, other surveys focusing on specific issues are carried out intermittently. Aggregated data of regular annual surveys are published (see for example the Policy Research Office of the Central Committee of the Communist Party of China and the Ministry of Agriculture 2001). Microdata of the surveys is also utilized in previous literature on rural household income, poverty, and other various issues on rural economy.⁹ Although availability of data to researchers is very limited, other public agencies also conduct similar fixed observation point-based surveys. For example, the National Health and Family Planning Commission has recently started a nationwide monitoring survey on migrant households focusing on reproductive health and economic well-being.¹⁰

3. Challenges for official data on wages, income and poverty

How well do NBS and its regional counterparts succeed in producing impartial, high quality and easily assessed information? These are ideals for statistical bureaus in any country, and on these broad questions we here only offer a few comments and note that a primary motivation for creating a statistical bureau is that the government should not be able to manipulate statistical information. However, the separation from the political leaders and their staffs can never be complete as it is the government which defines the task for the statistical bureau and also provides its resources. This is true for statistical bureaus around the world whose staffs know that some results are more welcome than others which can lead to self-censorship.

The degree of data quality is the outcome of all steps in the data production process: sampling, construction of instruments, fieldwork, coding and editing the data. Our subjective impression from having contact with NBS staff is that they have high ambitions at every step of the process. However, NBS has unfortunately not produced many documents describing the various steps in detail and it is therefore difficult to make a detailed and qualified evaluation of the data-generating process.

International standards for compiling macroeconomic statistics used to produce information on GDP, for example, have existed for many years. China has now adopted the System of National Accounts (SNA) meaning that such information is internationally comparable to a large degree. In contrast, comparability across countries on earnings and income data collected at the household level is lower than in national income data, although some international recommendations exist.¹¹

The largest problem with the Chinese system of collecting and publishing household statistics is that it is a dual system. China is treated as two different countries when it comes to household statistics used in policy making. Because of differences in data collection it is difficult to find statistical descriptions that refer to the entire population. While NBS publishes results on income inequality in rural areas and urban areas separately, the agency does not regularly publish information on the extent of income inequality in China as a

⁹ As for recent literature utilizing the merit of panel data characteristics, see for example, Tao et al. (2011). See Bramall (2001) for comparisons of income data of the NBS rural household survey and the fixed observation point survey.

¹⁰ From 2010 on, 106 city-based fixed survey points were built (the total number of the sample is approximately 100 000). See the National Health and Family Planning Commission (2010) .

¹¹See Canberra Group (2011).

whole. This is highly problematic, as all evidence indicates that much of the inequality in China as a whole is due to the large income gap between urban and rural areas.¹²

The lack of a survey system integrated for the entire population has meant that the urban system has not covered people with a rural *Hukou* permanently living in urban areas; the long-term rural-urban migrants. As during the planning époque restrictions for urban people to live in urban areas were strong, this might have been a relatively minor problem at the time. However, since China started to move towards a market economy ever larger numbers of rural residents have moved into the cities where some intend to stay temporarily, but many remain for an extended period (see for example Chan, 2012). Typically such persons find jobs, but they are paid less than urban residents. From this follows that all official statistics on income overestimate average household income among persons actually living in urban areas. Most likely official statistics also underestimate inequality in income among persons who actually live in the cities. It is a great challenge for the statistical agencies to try to include migrants in their household surveys as by definition they cannot be identified in official registers. However, as we will discuss later in this section, staff at NBS are working on how to develop the data collections.

It is an intriguing task to collect statistical information on migrants in China for several reasons. There are issues of definition. For example: What distinguishes migration from commuting? Migration can be characterised by origin and destination as well as by *hukou* status. There is the issue of how to sample migrants at the destination where some have accommodation similar to urban residents while others live, for example, on construction sites. As migration typically is temporary, by and large it is difficult to collect panel data on migrants. However, such problems have not hindered researchers in collecting data on earnings and income among migrants in China.¹³

Having two systems for collecting income information, one for rural households and another for urban households, gives rise to another problem. While the designations of rural and urban locations in China are fixed at one point in time, this can change. As people move into the cities and buildings are constructed on what was previously farmland, rural locations can receive urban status. As such changes are not frequent during a one-year period, they should have little effect on how to interpret year to year changes in household income. However, this may not be the case regarding long-run changes.

During the planning époque, almost all workers in urban China were employed by State Owned Enterprises or by Collective Enterprises. A substantial part of workers' compensation was given in kind: goods, access to very low-price housing, etc. Such compensation was typically not included in the definition of household income applied by NBS which can be seen as a substantial weakness. However, as many of these benefits have been phased out, the problem of under-recording has become much smaller over time. In contrast, China's transformation has led to a rapid increase in home ownership in urban areas. It is true to say that among economists there is consensus that imputed rents from homeownership should constitute a component of household income (although methods for its implementation differ). However, such a strategy is not taken by NBS.

¹²See for example Sicular et al (2007) and Li and Luo (2010).

¹³Gustafsson and Li (2006) describes six such surveys.

To repeat - in defining household income NBS does not include the value of benefits in kind received from employers nor imputed rent from owner-occupied housing for the information the agency publishes. However, researchers at times try to include such income components to come closer to the real situation. As a consequence one cannot expect to find full agreement between what such sources report and what NBS publishes. This is the case not only regarding income and income inequality per se, but also on their changes over time.

The coexistence of different household survey schemes between urban and rural populations has increasingly been inappropriate for the actual socioeconomic conditions in China. In addition to the vast flow of rural-urban migrants, many regions have carried out *Hukou* system reform to unify urban and rural *hukou* registration. In such regions, all household members are newly registered into the resident *Hukou* (*jumin hukou*). Against this background, NBS has recently started a significant alteration in the structure of household surveys, which is called reform of household survey for “urban-rural integration” (*chengxiang zhuhu diaocha yitihua gaige*). The new household survey scheme was announced in 2012 and is in effect from 2013 on. Ma Jiantang, the director of NBS, states that the aim of the reform is to build a unified, nationwide household survey system that can represent changes in urban-rural, interregional, and intersectoral income disparities and inequalities between different social groups (Ma 2012).

According to Ma (2012) and other related materials, the major points of this reform can be summarized as follows: First, the separate urban and rural sampling frames will be unified into one national frame by building base sampling units based on the 2010 Population Census. This means rural-urban migrant households are to be incorporated into the regular sampling frame for the first time. In addition, distribution of cities/counties surveys is to be adjusted to reflect recent changes in geographical structure of population caused by vast labour migration. Second, statistical indicators, which have not been comparable between urban and rural surveys, are to be unified. For example, “disposable income” (*kezhipeishouru*)” instead of “net income (*chunshouru*)” is to be used as the basic indicator for rural household income.¹⁴ Indicators of transfer income/expenditure are to be re-examined to more accurately reflect the impact of social security and other public policies. Incorporation of imputed rental income of owner-occupied housing is treated as another important issue. In addition, unification of rural and urban income indicators naturally requires reconsideration of in-kind income and expenditures (especially services) of urban households.

This reform is certainly an important milestone in the NBS household survey system. However, since relevant data collected through the “urban-rural integrated” scheme has not yet been published, it is unclear to what degree the new scheme will change the overall picture of income distribution. The perhaps most important issue will be the coverage of migrants. Considering the high mobility of migrants and difficulty in incorporating single migrating households into a household survey system, a complementary survey scheme for migrants (or for single households generally) might be necessary. Another issue is to what degree longitudinal comparability of published data is secured after 2013.

¹⁴Specifically, public transfer expenditures and interest payments of rural households which are not deducted in “net income”, should be deducted in disposable income.

Common to statistical authorities in other countries, NBS faces the problem of some households not being willing to participate in their surveys, see for example Gibson et al (2003). A related problem is that not all incomes might be recorded in the surveys. Most likely, non-response and underreporting lead to inequality in income being underestimated.¹⁵ Another not trivial issue of relevance to discuss is the fact that China is a very large country and prices differ across its territory. One can therefore argue that it is meaningful to correct statistical information on earnings and income by a spatial price index. While this is not done in the information published by NBS, some researchers are doing this using a spatial index developed by Brandt and Holz (2006).

The present international discussion on how statistical agencies best can measure economic performance and social progress is greatly inspired by the report of Joseph Stiglitz, Amartya Sen and Jean-Paul Fitoussi commissioned by French president Nicholas Sarkozy published in 2009 (Stiglitz et al, 2009). We cannot judge whether this report has had or will have implications on the statistics NBS collects and publishes. However, related to the issue of providing statistics that measure the household's well-being and gender differences, one can for example note that NBS conducted a time-use study for 2008 in a manner allowing for international comparisons.¹⁶

4. Researcher initialised databases.

While data collected by NBS is widely used by policy makers, international agencies and many researchers it also has limitations. Some of these limitations were discussed above and we can stress that individual researchers and research groups cannot count on having access to microdata from NBS. To the limitations already discussed one should add that in many cases NBS does not collect information matching concepts and circumstances researchers are interested in. This provides the background for why many researchers have involved themselves in the data-collecting process.¹⁷ Such data typically refers to one geographical area and one point in time. Limiting the fieldwork to one location and one point in time makes the sampling strategy and data collection manageable. One limitation with such a strategy is difficulties with generalising the results to a wider territory, another is that such data typically does not make it possible to follow and analyse changes over time.

However, there are now several examples of larger research groups being involved in collecting household data covering large parts of China. Table 1 describes five such projects by design, reference year, sampling and sample size as well as types of questionnaires used. The table also indicates the availability of surveys on characteristics of the community in which the respondents live. Much of the data catches are cross sectional. However, in some cases retrospective questions are also posed, and there are examples of a prospective panel design.

¹⁵ There has been considerable discussion in recent years about the underreporting of income of urban high-income groups, sometimes referred to as "gray income", see Luo et al. (2011) and Wang (2010).

¹⁶see <http://australian-time-users-group.org/assets/docs/chinese-tus08.pdf>

¹⁷For example Gustafsson and Li (2006) survey not less than 25 articles published in academic journals from 1997 to 2003 which have used such data, and subsequently many articles using this approach to data-access have been published

/Table 1 about here/

Among the five surveys, the first were the China Household Income Project (CHIP) and China Health and Nutrition Survey (CHNS), both started at the end of the 1980s. CHIP was initiated by a group of international scholars and scholars who at the time were based at the Institute of Economics, Chinese Academy of Social Science, Beijing. Now it is housed at Beijing Normal University, Beijing. CHNS is a project conducted by the University of North Carolina at Chapel Hill, USA and the National Institute of Nutrition and Food Safety at the Chinese Centre for Disease Control and Prevention, China.¹⁸ Microdata from both is available to researchers and a large number of studies using one or the other of the two databases have been conducted and published.

CHIP and CHNS differ in several respects. As indicated by their names, CHIP is focused on household income and CHNS on health and nutrition, with less detailed income information. CHIP has taken advantage of working with NBS in many stages of the data generating process, while this is not the case for CHNS. Households selected for the rural and urban surveys of CHIP are subsamples from NBS's larger surveys and cover many, but not all, province level units.¹⁹ The coverage of province level units is smaller in CHNS and does not cover for example any of the four municipalities. CHNS has panel data characteristics while CHIP is a repeated cross-section survey, although some retrospective data on household income is collected in each wave.

Being a multi-purpose survey containing rich information, CHNS has had large use outside a narrow circle of economists. Many users of CHIP have worked with concepts of income that are broader than that adopted by NBS in line with what is used in international studies. This has been done by including the value of benefits in kind received by urban households and including the imputed rent of owner-occupied housing. In the waves now available for researchers, income information refers to the years 1988, 1995, 2002 and 2007. In the first three waves there is also information on household wealth. As shown in Table 2, all waves of CHIP consist of separate surveys on rural and urban households. The 2002 wave and the 2007 wave also include a sample of migrant households, which are not the subsample of the NBS survey. For further details see Eichen and Zhang (1993), Li et al (2008) and Luo et al (2013). Usage of microdata from the CHIP surveys can be applied for from ICPSR and the China Institute of Income Distribution at BNU²⁰; microdata from CHIP 2002 has recently been made available from LIS Cross National Data Centre in Luxembourg where variable definitions as far as possible have been harmonised with similar surveys made for other countries in order to facilitate cross-country comparisons of earnings, income and poverty.²¹

/Table 2 about here/

Rural-Urban Migration in China and Indonesia (RUMiCI) was constructed to provide answers to questions on the impact of internal migration on income mobility, poverty alleviation, education, health and nutrition of children, and the assimilation of migrant

¹⁸See <http://www.cpc.unc.edu/projects/china>

¹⁹See Li et al. (2013) as the latest collections of papers based on CHIP.

²⁰See for example <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/21741>. the China Institute of Income Distribution can be accessed at <http://www.ciidbnu.org/>

²¹ See <http://www.lisdatacenter.org/news-and-events/first-chinese-dataset-in-lis/>

workers in the city.²²This panel survey housed at the Australian National University has the first survey for 2008. The sample for China covers the nine largest provinces sending and receiving migrants: Shanghai, Jiangsu, Zhejiang, Hubei, Sichuan, Guangdong, Henan, Anhui and Sichuan. The design includes one urban survey, one rural survey (both conducted by NBS) and one migrant survey (conducted by a professional survey organisation).²³ Along with standard questions on household characteristics, special questions on education, income and migration history are included in the migrant survey. In addition, the first wave of the Chinese General Social Survey refers to 2003. The survey is a collaboration between Renmin University of China and Hong Kong Sciences and Technology University, along with seven other Chinese universities. The basic structure is repeated cross sections. It contains core modules in each survey including questions on labour market and social security as well as topic modules that vary across surveys. Bian and Li (2012) provide further details on background and sample design. Since it is basically a sociological survey, more attention is devoted to subjective aspects of well-being including questions on perception of income inequality.

The China Family Panel Study (CFPS) launched in 2010 after preliminary studies conducted in 2008-2009, is conducted by the Institute of Social Science Survey at Peking University (ISSS). The study covers approximately 16 000 households distributed in 25 province level administrative units. CFPS has taken advantage of its panel data characteristics and unique questionnaire design that combines main questionnaires (separate adults and children questionnaires) and additional questionnaires focusing on specific population groups.²⁴ The second wave of the survey was conducted in 2012 and the data from the first two surveys are available for public use. The results and descriptive statistics from the surveys are presented in a new publication (Xie Yu et al, 2013).

Some researcher-conducted surveys collected information on community level (village, urban neighborhood) simultaneously with household data. Since household economic behavior and its outcomes are affected not only by macro and regional economic conditions, but also by economic and sociopolitical contexts at the community or small-region level (see, for example, Treiman 2009), it will be useful to employ a multilevel research framework by combining microdata and community-level/small-region data. The CFPS employs multilevel survey design and includes both urban and rural community surveys (residence committee and administrative village). CHNS has included community questionnaires from the start of the project. CHIP began to include a rural community (administrative village) questionnaire from the 2002 wave on. The 2007 wave of CHIP also includes an urban community (residence committee) survey to best capture the relationship between migrants and local urban residents. As for data provided by an official agency, the rural fixed observation-point survey applies a multilevel framework.

²² See <http://rse.anu.edu.au/rumici/>

²³ For a description of the sampling procedure see http://rse.anu.edu.au/rumici/pdf/Census%20manual_China_English08.pdf

²⁴ See official web site of ISSS. <http://www.iss.edu.cn/index.php>

5. Conclusions

To a large extent the research initialised databases focus on different aspects and are thus complementary to each other. They also complement the data that NBS collects. We conclude this paper by summarizing several basic issues that both [official surveys](#) and [researcher initiated surveys](#) need to address.

Firstly, the major challenge both for official surveys and researcher initiated surveys is how to incorporate rural-urban migrants. As discussed above, recent reform for “urban-rural integrated” household surveys initiated by NBS is certainly an important milestone in Chinese statistical systems. Researcher initialised surveys such as recent waves of CHIP and RUMiCI have also made progress, although the transient, informal and diverse situations of migrants make it difficult to capture this population group well.

Secondly, along with increases in income level and social mobility, probable biases due to rejection of survey, non-response, and underreporting of income (especially by the high-income group), which are universal issues in household surveys in other countries, will become more and more important for Chinese household data. Improvements in operation process as well as sampling process are needed for NBS household survey. NBS has already tried to improve operation process, for example, by introducing web-based bookkeeping. Regarding possible downward bias caused by the underreporting of income of the high-income group, some studies have tried to combine household survey data with other complementary data focusing on the high-income group (Li and Luo 2011).

Thirdly, data on wealth is relatively weak in both the NBS household survey and existing researcher initialised surveys. Wealth inequality has been increasingly important for the analyses of household well-being. For example, recent waves of CHIP show that imputed rent of owner-occupied housing has a strong impact on intra-urban income inequality and rural-urban income disparity (Sato et al. 2013). In addition, under the current institutional setting, inequality in wealth has most likely very strong implications for intergenerational transmission of inequality.

References

- Aaberge, R. and Li, S. (1997) "The Trend in Urban Income Inequality in Two Chinese provinces, 1986-90", Review of Income and Wealth, 43, 335-355.
- Aaberge, R. and Zhu, Y. (2001) "The Pattern of Household Savings during a Hyperinflation: The Case of Urban China in the Late 1980s", Review of Income and Wealth, 47, 181-202.
- Bian, Y and Li, L. (2012) "The Chinese General Social Survey (2003-8). Sample Designs and Data Evaluation", Chinese Sociological Review, 45, 70 – 97.
- Brandt, L. and C. A. Holz (2006) "Spatial Price Differences in China: Estimates and Implications," Economic Development and Cultural Change, 55 (1), 43-86.
- Bramall, C. (2001) "The Quality of China's Household Income Surveys", *The China Quarterly*, 167, pp. 689-705.
- Canberra Group (2011) Handbook on Household Income Statistics, Second Edition, Geneva: United Nations.
- Chan, K.W. (2012) "Migration and Development in China: Trends, Geography and Current Issues", Migration and Development, 1 (2) 187 – 205.
- Chen, S. and M. Ravallion (1996) "Data in Transition: Assessing Rural Living Standards in Southern China" China Economic Review, 7(1), 23-55.
- Chi, W and Li, B. (2008) "Glass Ceiling or Sticky Floor? Examining the Gender Earnings Differential Across the Earnings Distribution in Urban China, 1987 – 2004", Journal of Comparative Economics, 36 (2), 243-263.
- Coady, D. and Wang, L. (2000) "Incentives, Allocation and Labor-Market Reform during Transition: The Case of Urban China, 1986-1990", Applied Economics, 32, 511-526.
- Cook, S. and Keeley, J. (2007) "Micro-data Scoping Study – China" Report submitted to Economic and Social Research Council U.K. Institute of Development Studies and International Institute for Environment and Development, available at http://www.esrc.ac.uk/_images/China_tcm8-5137.pdf
- Eichen, M. and M. Zhang (1993) "Annex: The 1988 Household Sample Survey: Data Description and Availability," in K. Griffin and R. Zhao, eds., The Distribution of Income in China, New York: St. Martin's Press, 331-346.
- Fang, C., Zhang, X., Fan, S. (2002) "Emergence of Urban Poverty and Inequality in China: Evidence from Household Surveys", China Economic Review, 13, 430-443.
- Giles, J. Park, A. and Zhang, J. (2005) "What is China's True Unemployment Rate?", China Economic Review, 16, 149-170.
- Gibson, J., Huang, J and Rozelle, S. (2003) "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey", Review of Income and Wealth, 49, 53-68.
- Gustafsson, B. and Li S. (2006) "Surveys – Three Ways of Obtaining Household Income Data", pages 129 – 152 in Heimer, M. and Thogersen, S. (Eds) Doing Fieldwork in China, Copenhagen: NiAS Press. Chinese version published in 2012 by Chongqing University Press (pages 56 – 78).

- Kung, J. and Lee, Y. (2001) “So What If There is Income Inequality? The Distributive Consequences of Nonfarm Employment in Rural China”, Economic Development and Cultural Change, 30, 395-414.
- Li, S., C. Luo, Z. Wei, and X. Yue (2008), “Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Description,” in B. Gustafsson, S. Li, and T. Sicular, eds., Inequality and Public Policy in China, New York: Cambridge University Press, 337-353.
- Li, H., Zhang, J, Sin, L.T. and Zhao, Y (2006) ”Relative Earnings of Husbands and Wives in Urban China”, China Economic Review, 17 (4), 412-431.
- Li, S., Sato, H. and Sicular, T. (eds) (2013) Rising Inequality in China: Challenges to a Harmonious Society, New York: Cambridge University Press. (Chinese edition published as Zhongguo Shouruchaju Biandong Fenxi: Zhongguo Jumin Shourufenpei Yanjiu IV, Beijing: Renmin Chibanshe [People’s Press], 2013)
- Li S. and C. Luo (2010) “Re-estimating the Income Gap between Urban and Rural Households in China,” in M. K. Whyte, ed., *One Country, Two Societies: Rural-Urban Inequality in Contemporary China*, Cambridge, MA: Harvard University Press, 105-121.
- Li S. and C. Luo (2011) “Zhongguo shouru chaju jiujiing you duoda? [How Unequal is China?]” Jingji yanjiu [Economic Reserach], 2011-4, 68-78.
- Liu, Q. (2012) “Unemployment and Labor Force Participation in Urban China”, China Economic Review, 23, 18-33.
- Luo, C., X. Yue, and S. Li (2011) “Dui Wang Xiaolu Huiseshouru Gusuan de Zhiyi [A critique on Wang Xiaolu’s Estimation of the Gray income]”, Bijiao, 52, 146-158.
- Luo, C, Li, S., Sicular, T. Quheng, D.,, and Yue, X.(2013) “The 2007 household surveys: Sampling, methods and data description” in Li, S., Sato, H. and Sicular, T. (eds) Rising Inequality in China. Challenges to a Harmonious Society, New York: Cambridge University Press.
- Ma, J (2012) “Dali tuijinchengxiangzhuhudiaoachayitihuagaige [The promotion of reforms toward urban-rural integrated household survey]”, ZhongguoTonji, 2012-3, electric version, http://www.stats.gov.cn/tjsujia/zgtj/t20120331_402796201.htm[accessed in September 11, 2013]
- McCulloch, N. and Calandrino, M. (2003) “Vulnerability and Chronic Poverty in Rural Sichuan”, World Development, 31, 611-628.
- Meng, X. (2012) “Labour Market Outcomes and Returns in China”, Journal of Economic Perspectives, 26, 75-102.
- Meng, X, Sheng, K. and Xue, S. (2013) “Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988-2009”, Journal of Comparative Economics, 41, 227-244.
- National Bureau of Statistics (2011) China Yearbook of Household Surveys 2011, Beijing: China Statistical Press.
- National Bureau of Statistics (2012) “2012 quanguo nongmingong jiance diaocha baogao” [The National Rural-urban Migrant Monitoring Survey 2012], published electrically at the official web site of NBS. http://www.stats.gov.cn/tjfx/jdfx/t20130527_402899251.htm. [accessed in September 11, 2013]
- tNational Health and Family Planning Commission (2010) “Guanyu jiaqiang quanyuan liudong renkou tongji, kaizhan liudong renkou dongtai jiance gongzuo de zhidao yijian [Direction on improvement of comprehensive statistics of migrant population and promotion of the monitoring of

population mobility dynamics]”, Official Website of the National Health and Family Planning Commission,
<http://www.moh.gov.cn/ldrks/s3573/201306/a32678f86b2b4c35bdb1dafcca1547b3.shtml> [accessed in September 12, 2013]

Policy Research Office of the Central Committee of the Communist Party of China and the Ministry of Agriculture (2001) *Quanguo Nongcun Shehui Jingjin Dianxing Diaocha Shuju Huibian* [National Rural Socio-economic Survey Data Collection], Beijing: Zhongguo Nongye Chubanshe [China Agricultural Press].

Rural Survey Department of NBS ed. (2011) *Zhongguo Nongcun Pinkun Jiance Baogao 2010* [Poverty Monitoring Report of Rural China 2010], Beijing: Zhongguo Tongji Chubanshe [China Statistics Press].

Sato, H., T. Sicular, and X. Yue (2013) “Housing Ownership, Incomes, and Inequality in China, 2002-2007” in Li, S., Sato, H. and Sicular, T. (eds) *Rising Inequality in China: Challenges to a Harmonious Society*, New York: Cambridge University Press.

Sicular, T., Yue, X, Gustafsson, B and Li, S (2007) “The Urban – Rural Income Gap and Inequality in China”, *Review of Income and Wealth*, 53 (1) 93 – 126.

Stiglitz, J., Sen, A. and Fitoussi, J-P (2009) Report by the Commission on the Measurement of Economic performance and Social Progress, available at
http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf

Tao, R., M. Liu, et al. (2011) “Grain Procurement, Tax Instrument and Peasant Burdens during China's Rural Transition”, *Journal of Contemporary China*, 20 (71), pp. 659-677.
Treiman, Donald, Y. Lu, and Y. Qi (2009) “New Approaches to Demographic Data Collection,” California Center for Population Research On-Line Working Paper Series, CCPR-2009-022.

Tsui, K. (1998) “Trends and Inequalities of Rural Welfare in China: Evidence from Rural Households in Guangdong and Sichuan”, *Journal of Comparative Economics*, 26, 783-804.

Wang, X. (2010) “Huiseshouru yu Guominshourufenpei [Grey Income and National Income Distribution]”, *Bijiao*, 48, 1-29.

Xie, Yu, Zhang Xiaobo, Li Jianxin, Yu Xuejun, Ren Qiang. (2013) *Zhongguo Minsheng Fazhan Baogao* [China Report on People's Livelihood Development] Beijing: Beijing Daxue Chubanshe [Peking University Press].

Yang, D. T. (1999) “Urban-Biased Politics and Rising Income Inequality in China”, *American Economic Review*, 89, 306-310.

Yue, X, S. Li, P. Wang, and B. Guan (2007) *Toushi Zhongguo Nongcun Pinkun* [Perspective on Rural Poverty in China] Beijing: Jingji Kexue Chubanshe [Economic Science Press].

Zhang, J., Zhao, Y., Park, A. and Song, X (2005) “Economic Returns to Schooling in Urban China 1988 to 2001”, *Journal of Comparative Economics*, 33 (4), 730-752.

Table 1 Researcher-initiated databases on economic conditions of Chinese households

Projects/surveys	China Household Income Project	Rural-Urban Migration in Indonesia and China	China Health and Nutrition Survey	Chinese General Social Survey	Chinese Family Panel Studies
Abbreviations	CHIP	RUMiCI	CHNS	CGSS	CFPS
Design	Repeated cross-sections, with revised modules of questionnaires each wave	Panel	Panel (rotating panel), with revised modules of questionnaires each wave	Repeated cross-sections, with new modules of questionnaires each wave	Panel, with a new wave of interviews each year
Reference years	1988, 1995, 2002, 2007 +	2007 +	1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 +	2003, 2004, 2005, 2006, 2008 +	2010+
Sample specifications	Rural and urban samples: National probability sample of households (subsample of NBS's official annual household survey) Migrant sample 2002: selected from NBS's preparatory census data for annual urban household survey, including only long-term, stable migrants Migrant sample 2007: survey team's own sampling through using equal area sampling using street maps of 15 cities, including both long-term stable migrants and temporary migrants.	Samples of 2007 wave are subsamples of CHIP 2007. Rural and urban samples: National probability sample of households (subsample of NBS's official annual household survey) Migrant sample: survey team's own sampling through using equal area sampling using street maps of 15 cities.	Multistage semi-probability sample	National probability sample of people age 18-69	National probability sample (excluding ethnic minority autonomous districts in western region)

Sample size	See Table 2	2007: approximately 5 000 urban households and 8 000 rural households, and 5 000 migrant households.	Approximately 4 400 households, 19 000 individuals	2003:5 894 (urban only); 2004+ approximately 10 000	Approximately 16 000
Questionnaire					
Urban individual	Yes	Yes	Yes	Yes	Yes
Urban household	Yes	Yes	Yes	Yes	Yes
Urban community	(Yes, 2007 in correspondence with migrant samples)	Yes	Yes	No	Yes
Rural individual	Yes	Yes	Yes	Yes	Yes
Rural household	Yes	Yes	Yes	Yes	Yes
Rural community	Yes (from 2002)	Yes	Yes	No	Yes
Migrant individual	Yes (from 2002)	Yes	No	No	No
Migrant household	Yes (from 2002)	Yes	No	No	No

Sources: Eichen and Zhang (1993), Li et al. (2008), Luo et al. (2013), Li et al. (2013), Riskin et al.(2001), Treiman et al. (2009), ISSSS (Institute of Social Science Survey, Peking University, <http://www.iss.edu.cn/index.php?catid=134&action=index>) , CSSOD (Chinese Social Survey Open Database, Renmin University of China, <http://www.cssod.org/index.php>), RUMiCI (Australian National University, rumici.anu.edu.au), CHNS (The Carolina Population Center, the University of North Carolina at Chapel Hill, www.cpc.unc.edu/china)

Table 2 Sample structure of the CHIP surveys

	1988			1995		
Household category	Urban	Rural	Rural-urban migrant	Urban	Rural	Rural-urban migrant
Individuals	31 827	51 352	--	21 694	34 739	--
Households	9 009	10 258	--	6 931	7 998	--
Provinces	10	28	--	11	19	--
	Beijing, Liaoning, Jiangsu, Guangdong, Shanxi, Anhui, Henan, Hubei, Yunnan, Gansu	Beijing, Tianjin, Shanghai, Hebei, Liaoning, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan, Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan, Inner Mongolia, Guangxi, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia		Beijing, Liaoning, Jiangsu, Guangdong, Shanxi, Anhui, Henan, Hubei, Sichuan, Yunnan, Gansu	Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shandong, Guangdong, Shanxi, Jilin, Anhui, Jiangxi, Henan, Hubei, Hunan, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu	
	2002			2007		
Household category	Urban	Rural	Rural-urban migrant	Urban	Rural	Rural-urban migrant
Individuals	21 696	34 719	5 327	29 262	51 847	8404
Households	6 934	7 998	2 005	10 000	13 000	4 978
Provinces	12	22	12	16	16	9
	Beijing, Liaoning, Jiangsu, Guangdong, Shanxi, Anhui, Henan, Hubei, Sichuan, Chongqing, Yunnan, Gansu	Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Shandong, Guangdong, Shanxi, Jilin, Anhui, Jiangxi, Henan, Hubei, Hunan, Guangxi, Sichuan, Chongqing, Guizhou, Yunnan, Shaanxi, Gansu	Beijing, Liaoning, Jiangsu, Guangdong, Shanxi, Anhui, Henan, Hubei, Sichuan, Chongqing, Yunnan, Gansu	Beijing, Shanghai, Liaoning, Jiangsu, Zhejiang, Fujian, Guangdong, Shanxi, Anhui, Henan, Hubei, Hunan, Sichuan, Chongqing, Yunnan, Gansu	Beijing, Hebei, Liaoning, Jiangsu, Zhejiang, Fujian, Guangdong, Shanxi, Anhui, Henan, Hubei, Hunan, Sichuan, Chongqing, Yunnan, Gansu	Shanghai, Jiangsu, Zhejiang, Guangdong, Anhui, Henan, Hubei, Sichuan, Chongqing

Sources: Eichen and Zhang (1993), Li et al. (2008), Li et al. (2013a), Luo et al. (2013), Riskin et al.(2001).

