China Economic Review 30 (2014) 419–431

Contents lists available at ScienceDirect

China Economic Review

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

Björn GUSTAFSSON [a,b,*], Shi LI [b,c], Hiroshi SATO [d]

[a] Department of Social Work, University of Gothenburg, P.O. Box 720, SE 405 30 Gothenburg, Sweden
[b] Institute for the Study of Labor (IZA), Bonn, Germany
[c] School of Economics and Business Administration, Beijing Normal University, 19 XinJie Kou WaiDajie, Beijing 100875, China
[d] Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan

ARTICLE INFO

Article history:
Received 16 September 2013
Received in revised form 20 May 2014
Accepted 21 May 2014
Available online 28 May 2014

Keywords:
Income
Inequality
Poverty
Microdata
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 until recently had a dual system of household surveys — one rural and one urban. This has had 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 large databases aimed at mapping earnings, household income and poverty, and we present seven of these in the paper.

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 the concept and variable in focus, 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 types 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 on web pages. 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 interested in conducting empirical studies on earnings, the distribution of household income and poverty in China did not have access to microdata. Since then, the situation has changed radically, and a large number of studies based on such data can now be found in the literature. This has been made possible by researchers being

* Corresponding author at: Department of Social Work, University of Gothenburg, P.O. Box 720, SE 405 30 Gothenburg, Sweden. Tel.: +46 31 786 1890; fax: +46 31 786 1888.
E-mail addresses: Bjorn.Gustafsson@socwork.gu.se (B. Gustafsson), lishi@bnu.edu.cn (S. LI), sato.zuoteng@r.hit-u.ac.jp (H. Sato).

http://dx.doi.org/10.1016/j.chieco.2014.05.012
1043-951X/

given the opportunity to work with microdata. The National Bureau of Statistics (hereinafter referred to as the NBS) has collected the data. Other data are the fruits of researchers' initiatives of collecting surveys covering large parts of China. The first academic studies on the topics we cover here were based on single cross-sections of data. When the data collection was repeated, it became 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. Examples include studies aiming at understanding the duration of poverty and reasons for moving into and out of poverty. Panel data also put the analyst in a better position when trying to find the causes and effects of differences in earnings, income, and poverty. Another task of this paper is to introduce databases with microdata 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, household income and poverty. Section 3 discusses other official data sources. The topic for Section 4 is microdata that are the outcome of researchers' activities. The paper ends with concluding comments.

2. Household income data collected by NBS[2]

Most countries have official statistical systems that produce and publish information on the populations in which they operate. In the People's Republic of China, the NBS (previously known as the State Statistical Bureau, SSB) is responsible for countrywide reporting and has counterparts in each of the 22 provinces, 4 municipalities and 5 autonomous regions. There are also statistical authorities at the sub-province level. The province level statistical bureau is responsible not only 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. We will start with some broad remarks on what characterizes household data collected by the NBS. After that, and in somewhat more detail, we will discuss the system of collecting data from households on income, a system that has recently changed.

For researchers, the strategy of working with data already collected by the NBS can be advantageous in many respects. Statistical bureaus typically have a long-term commitment to their work, leading to their being experts at all stages of the data gathering process, which thus gives them 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 labor 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 appear to be handled on a case by case basis. Initially, access was typically given to only one or a small number of regions, making it difficult to judge to what extent results could be generalized to large parts of China.[3] However, more recently there are examples of research groups that have been successful in gaining access to microdata from the NBS covering large parts of China and for a substantial number of years.[4]

How well do the NBS and its regional counterparts succeed in producing impartial, high quality and easily accessed information? These are ideals for statistical bureaus in any country, and on these broad questions we will only offer a few comments. We 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 that defines the task for the statistical bureau and 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 for some time is that they have high ambitions at every step of the process. Unfortunately, NBS has 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.[5]

China differs from most other countries, however, in that institutional arrangements for rural and urban areas are rather different. The *hukou* system was introduced in the late 1950s as an integrated part of the planning system, and it not only registers a person as rural or urban but also has large implications for his or her well-being. Many public policy measures have long prioritized urban persons who are also holders of an urban *hukou*. People living in the cities live a different life than their peers in the rural areas.[6] This is the background for why China has long had a Rural Household Survey covering rural areas and an Urban Household Survey for the urban areas rather than 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 collect information to produce the national

[2] This section builds partly on Gustafsson and Li (2006).
[3] 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, Zhang, and Fan (2002).
[4] See for example Zhang, Zhao, Park, and Song (2005), Li, Zhang, Sin, and Zhao (2006), Chi and Li (2008), Meng (2012) and Meng, Sheng, and Xue (2013).
[5] See Canberra Group (2011).
[6] See, for example, Chan (2009) or Whyte (2010).

statistics. When doing this, the statistical bureaus have used different forms for rural and urban areas, and definitions of certain key variables have differed between urban and rural areas. However, because ever larger numbers of people holding a rural *hukou* live in urban areas, such a separation of the people into two categories has become increasingly problematic. As we will discuss further, this has motivated recent changes in the system of collecting household data.

One consequence of the differences in data collection between rural and urban areas is that up to now, it has been difficult to find official statistical descriptions that refer to the entire Chinese population. While the NBS has published results on income inequality in rural areas and urban areas separately, the agency for many years did not publish information on the extent of income inequality in China as a whole. This was 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.[7] However, as we will further discuss later in this section, the NBS recently published a time series of Gini coefficients for China as a whole.

Furthermore, the old household income survey system has meant that people with a rural *hukou* permanently living in urban areas, the long-term rural–urban migrants, were more unlikely to be covered by either rural or urban surveys. Because restrictions regarding rural people living in urban areas were strong during the planning stage, this might have been a relatively minor problem at the time. However, because China started to move toward a market economy, ever larger numbers of rural residents have moved into the cities, where some intend to stay temporarily, but others 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 most likely overestimate average household income among persons actually living in urban areas. Furthermore, income inequality among persons who actually live in the cities might be biased. It is a great challenge for the statistical agencies now making efforts to include migrants in their household surveys, as by definition the migrants cannot be identified in official registers.

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 characterized 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 accommodations similar to urban residents, while others live, for example, on construction sites. Because migration is typically temporary, it is largely difficult to collect panel data on migrants. However, such problems have not hindered researchers from collecting data on earnings and income among migrants in China.[8]

Having two systems for collecting income information, one for rural households and another for urban households, gives rise to a particular type of problem. While the designations of rural and urban locations in China are fixed at one point in time, they are subject to change. As people move into the cities and buildings are constructed on what was previously farmland, rural locations can receive urban status. Because 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 necessarily be the case regarding long-run changes.

It is noteworthy that the NBS has recently released current and retrospective estimations of Gini coefficients of per capita household income at the national level based on the official urban and rural household surveys. The National Gini coefficients for per capita disposable income officially announced by the NBS were 0.473 in 2013 and 0.474 in 2012 (NBS, 2013; NBS, 2014). Retrospective estimations of Gini coefficients for previous years were also recently published (NBS, 2013).[9] The current and retrospective estimation of national Gini coefficients by the NBS is a very important reference of inequality in China because the NBS has tried to integrate the urban–rural dual survey system with the new "urban–rural integrated" sampling framework employed at the end of 2012 (discussed below) by utilizing information from the preliminary data and population census data (Wang, 2013). When estimating the national Gini coefficients, the NBS also used the data from personal income tax records to correct downward bias of estimation.

We will now discuss in more detail how households were previously sampled in the rural and urban household surveys. Generally, the strategy by the NBS of sampling households and individuals was to apply a multi-stage probability sampling approach, striving for the information to be representative at the province level. Some Chinese provinces have the population size of a country, and statistical bureaus in such provinces draw additional samples for producing more detailed information, which is published in the statistical yearbook of the province. In rural areas, counties were chosen after categorizing all counties in a province by average income. Choosing villages in each county followed the same method. In each village, the target was to survey 10 households. In 2010, for example, not fewer than 74,000 households from 869 counties were chosen (NBS, 2011). A household chosen for the rural survey was included for a period of typically five years after which it was rotated out. The rural households were asked not only to provide detailed income information by bookkeeping but also to keep detailed records of their expenditures. While the rural survey assembled information on the demographic characteristics and the education of each household member, it also asked each household member questions on earnings that can only be attributed to each of the household members. Many thousands of assistant enumerators were involved in helping the households keep good records.

Household samples for the Urban Household Survey were drawn from a large sampling frame of households having an urban *hukou*. A two-stage stratified systematic random sampling scheme was applied. In the first stage, cities and counties were classified into five categories by population size in each province. They were then grouped into six administrative regions. In each

[7] See, for example, Sicular, Yue, Gustafsson, and Li (2007) and Li and Luo (2010).

[8] Gustafsson and Li (2006) describe six early surveys. In Section 4, we describe RUMiC, which since 2007 has studied rural to urban migrants in China.

[9] Retrospective figures are 0.479 for 2003, 0.473 for 2004, 0.485 for 2005, 0.487 for 2006, 0.484 for 2007, 0.491 for 2008, 0.490 for 2009, 0.481 for 2010, and 0.477 for 2011. Xie and Zhou (2014) report estimates of the Gini-coefficient for 2010–2012 based on some researcher-initiated surveys presented in Section 4 that are all higher. It is a task for future research to analyze the reasons for such differences.

administrative region, the cities and counties were arranged by average wages of staff and workers in the urban area. Based on this classification, cities and counties were selected by a systematic sampling scheme. At the second stage, sample households were selected at the city or county level. For 2010, for example, this resulted in 66,000 households in 474 cities being selected by the NBS (NBS, 2011).

Households selected for the Urban Household Survey typically remained for a period of three years, after which they were rotated out. As in rural areas, the households were regularly visited by an enumerator who asked questions and assisted in bookkeeping. As in urban areas, the respondent and his or her household had to record all income and expenditures in a very detailed way. For example, questions on earnings were put to each person in the household and the answers constituted the basis for tables published by the NBS. However, as is true for the rural survey, the urban survey did 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.[10]

Before turning to how the household surveys have been reformed, we will discuss the type of information that is collected by the NBS in the household income surveys (previously as well as presently). During the planning stage, 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 or access to very low-price (housing for example). Such compensation was typically not included in the definition of household income applied by the 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 (Li & Zhao, 2003). 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 home ownership should constitute a component of household income (although methods for its implementation differ). However, such a strategy is not taken by the NBS.

In defining household income, the NBS has not included the value of benefits in kind received from employers nor imputed rent from owner-occupied housing when compiling the information the agency publishes. However, researchers at times have tried 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 the NBS publishes. This is the case regarding not only on income and income inequality at one point in time but also on their changes over time. There is also the issue of comparability across time in the published series. The income concept adopted by the 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 of Chinese households is not real but is due to changes toward an improved definition of disposable income.

It is this type 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 cut-off. The poverty line for rural China was long set at a very low level, for example, at 627 Yuan per person per year in 2002. Starting in 2000, the 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 2300 Yuan (in 2010 constant prices) per person per year was introduced in 2011. This poverty line is close to a 1.75 dollar per day poverty line in PPP, meaning it is in the range of the two world poverty lines (set at 1.25 USD or 2.00 USD in PPP per person and day) the World Bank often use today when reporting on the development of poverty worldwide.[11] Based on this poverty line, in 2011 the number of poor in rural China was 122 million, equivalent to 12.7% of the rural population. In contrast, in its monitoring of urban poverty, the 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.

The coexistence of different household survey schemes between urban and rural populations has increasingly been inappropriate for understanding the actual socioeconomic conditions in China. In addition to the vast flow of rural–urban migrants, many regions have carried out *hukou* system reform, unifying urban and rural *hukou* registration. In such regions, all household members have been newly registered into the resident *hukou* (*jumin hukou*). Against this background, the NBS has recently started a significant alteration in the structure of household income surveys; this is called the reform of the household survey for "urban–rural integration" (*chengxiang zhuhu diaocha yitihua gaige*). The new household survey scheme was announced in 2012 and has been in effect since 2013. Ma Jiantang, the director of the 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 inter-sector 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 sample frame by building primary 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, the distribution of city/county surveys is to be adjusted to reflect recent changes in the geographical structure of the population caused by the vast labor migration. Second, statistical indicators, which have not been comparable for urban and rural surveys, are to be unified. For example, "disposable income" ("*kezhipei shouru*") instead of "net income" (*chunshouru*) is to be used as the basic indicator for rural household income.[12] Indicators of transfer income/expenditure are to be re-examined to more accurately reflect the impact of social security and other public policies. The incorporation of imputed rental income of

[10] Much of our description is also adequate for how the household income systems worked during the 1980s and 1990s. For more detailed information on the situation for those decades, see Chen and Ravallion (1996:26-31) and Bramall (2001).

[11] However, note that while NBS bases its estimate on rural poverty on household income, the World Bank typically makes its assessments based on household consumption.

[12] Specifically, public transfer expenditures and interest payments of rural households that are not deducted in "net income" should be deducted in disposable income.

owner-occupied housing is treated as another important issue. In addition, the 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, because relevant data collected through the "urban–rural integrated" scheme have not yet been published, it is unclear to what degree the new scheme will change the overall picture of China's income distribution. Perhaps the most important issue will be the coverage of migrants. Considering the high mobility of migrants and the difficulty in incorporating single migrating households into a household survey system, it might be necessary to conduct a complementary survey scheme for migrants (or for single households generally) in the future. Another issue is to what degree the longitudinal comparability of published data is secured after 2013.

Having discussed the progress in the household income system in China, which is accomplished by urban–rural integration, we will comment here on problems that the reform will hardly be able to solve. Common to statistical authorities in other countries, the NBS faces the problem of a number of households not being willing to participate in their surveys; see, for example, Gibson, Huang, and Rozelle (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.[13] Another issue of relevance is 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 the NBS, some researchers are doing this using a spatial index developed by Brandt and Holz (2006).

The present international discussion on how statistical agencies can best measure economic performance and social progress has been greatly inspired by the report of Joseph Stiglitz, Amartya Sen and Jean-Paul Fitoussi commissioned by French president Nicholas Sarkozy and published in 2009 (Stiglitz, Sen, & Fitoussi, 2009). We cannot judge whether this report has had or will have implications for the statistics the 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 the NBS conducted a study for 2008 on how households spend their time in a manner allowing for international comparisons.[14]

3. Other household data collected by the NBS or other public organizations[15]

In this section, we will discuss some other household data collected by the NBS and other public organizations that have a somewhat different focus than the household income surveys. Much of China's poverty-alleviation strategy has for some time been directed toward officially designated rural poor areas. This provides the background for why, since 1997, the NBS has monitored rural poverty using the Rural Poverty Monitoring Survey (*nongcun pinkun jiance diaocha*). This survey focuses on all nationally designated poor counties (592 counties) and covers 54,000 sample households distributed throughout 5400 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.[16]

To address the increasing mobility of the population, the NBS launched the National Rural–Urban Migrant Monitoring Survey (*quanguo nongmingong jiance diaocha*) at the end of 2008. This survey is a quarterly-based sampling survey of the rural labor force. The survey's purpose is to capture out-migration flow at the place of origin. Its sampling frame includes more than 7500 villages in 899 counties in 31 province level administrative units, and it covers approximately 200,000 rural laborers. The survey is conducted among the same households selected by the regular rural household survey. Aggregated data containing the 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).

Among other data with relevance for studying people's well-being, since the mid-1980s the 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 was the latest such survey, and it includes more detailed information related to demographic, geographic, economic, and housing information than the censuses provide. This large sample survey is useful for studying migration flows, for example, but the income information is limited to monthly earnings (NBS, 2007).

Several countries have conducted regular labor force surveys for many years, from which statistics on employment 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 the local NBS, carried out the China Urban Labour Survey (CULS) for the years 2001, 2005 and 2010.[17] Recently, the Department of Sociology at Sun Yat-sen

[13] There has been considerable discussion in recent years about underreporting of the income of urban high-income groups in China, sometimes referred to as "gray income"; see Luo, Yue, and Li (2011) and Wang (2010).

[14] See http://australian-time-users-group.org/assets/docs/chinese-tus08.pdf.

[15] Cook and Keeley (2007) provide information on an additional number of data sources that are not covered here.

[16] For a detailed introduction to the Rural Poverty Monitoring Survey and empirical analyses of rural poverty using the survey, see Yue, Li, Wang, and Guan (2007).

[17] These surveys cover each of the years' five provincial capital cities (Shanghai, Wuhan, Shenyang, Fuzhou and Xi'an). For 2005 (but not subsequently), seven municipal cities were added (Wuxi, Yichang, Benxi, Zhuhai, Shenzhen, Baoji and Daqing). Guangzhou was added to the 2010 survey. Separate samples of persons with urban residence status and with migrant status were derived using a two-stage procedure: first by area, then by the selection of households to be interviewed. For further information, see Giles, Park, and Zhang (2005) and Gao and Smyth (2012).

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

Projects/surveys	China Household Income Project	Rural-Urban Migration in China	China Health and Nutrition Survey	Chinese General Social Survey	Chinese Family Panel Studies	China Health and Retirement Longitudinal Study	China Household Finance Survey
Abbreviations	CHIP	RUMiC	CHNS	CGSS	CFPS	CHARLS	CHFS
Design	Repeated cross-sections, with revised modules of questionnaires each wave	Panel, with a new wave of interviews each year	Panel (rotating panel), with revised modules of questionnaires each wave	Repeated cross-sections, with new modules of questionnaires each wave; panel from 2010 wave	Panel, with a new wave of interviews each year	Panel, with a new wave of interviews every two years	A cross-section has been completed.
Reference years	1988, 1995, 2002, 2007 +	2007 +	1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 +	2003, 2004, 2005, 2006, 2008, 2010 +	2010 +	2011–2012, 2013 + (pilot survey and resurvey 2008, 2012)	2011
Sampling frame	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 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 equal area sampling using street maps of 15 cities.	Multistage semi-probability sample	National probability sample of people age 18 and over	National probability sample (excluding ethnic minority autonomous districts in western region)	National probability sample of people age 45 and over	Multistage semi-probability sample
Sample size	See Table 2	2007: approximately 5, 000 urban households and 8, 000 rural	Approximately 4, 400 households, 19,000 individuals	2003: 5, 894 (urban only); 2004 + approximately 10,000	Approximately 16,000 households	Approximately 17,500 individuals of 10,000 households	Approximately 8,400 households, 29,500 individuals

		households, and 5, 000 migrant households.		individuals 2010 + approximately 10,000 individuals			
Questionnaire							
Urban individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban household	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urban community	(Yes, 2007 in correspondence with migrant samples)	Yes	Yes	No	Yes	No	No
Rural individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rural household	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rural community	Yes (from 2002)	Yes	Yes	Yes	Yes	No	No
Migrant individual	Yes (from 2002)	Yes	No	No	No	No	No
Migrant household	Yes (from 2002)	Yes	No	No	No	No	No
Method of collecting income data	Household bookkeeping and interview (urban and rural samples), interview (migrant samples)	Household bookkeeping and interview (urban and rural samples), interview (migrant samples)	Interview	Interview	Interview	Interview	Interview (main interview and quarterly-basis complementary interviews)
Main host Institution	China Institute for Income Distribution (CID), Beijing Normal University (BNU) http://www.ciidbnu.org/	RUMiCI project, the Australian National University http://rse.anu. edu.au/RUMiCI/ and CID, BNU	The Carolina Population Center, the University of North Carolina at Chapel Hill http://www.cpc.unc.edu/projects/china	National Survey Research Center, Renmin University of China (RUC) http://www.chinagss.org/	Institute of Social Science Survey, Peking University http://www.isss.edu.cn/cfps/	Institute of Social Science Survey, Peking University http://charls.ccer.edu.cn/zh-CN	The Survey and Research Center for China Household Finance (CHFS), Southwestern University of Finance and Economics (SWUFE). http://www.chfsdata.org/
Related data archives	Inter-university Consortium for Political and Social Research (ICSR) www.icpsr.umich.edu/ The Universities Service Centre for China Studies (USC), the Chinese University of Hong Kong (CUHK) http://www.usc.cuhk.edu.hk/Eng/Default.aspx	International Data Service Center (IDSC), IZA idsc.iza.org		Chinese Social Survey Open Database (CSSOD), RUC http://www.cssod.org/index.php USC, CUHK			

Sources: Eichen and Zhang (1993), Griffin and Zhao (1993), Gustafsson, Li, and Sicular (2008), Li et al. (2008), Li, Sato, and Sicular (2013), Luo et al. (2013), Riskin, Zhao, and Li (2001), Treiman et al. (2009), and official websites of main host institutions listed above.

University launched the China Labor Force Dynamics Survey (CLDS) to monitor labor market dynamics among Chinese adults. The survey includes households in most provinces of mainland China.[18]

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*), which has been conducted by the Ministry of Agriculture since 1986.[19] This survey is a longitudinal survey of villages as well as rural households, and it currently covers 23,000 households living in 360 villages distributed in 30 province level administrative units, excluding Tibet. The sampling of the survey proceeds as follows. First, counties belonging to high-, middle-, and lower-income categories are selected. Second, "representative" villages are selected in each sample county. In the next step, 40–120 households are selected in each village depending on village size. In addition to regular annual surveys on households and villages collecting basic information such as land utilization, labor inputs, agricultural production, income, and expenditures, other surveys are intermittently carried out that focus on more specific issues. Aggregated data of regular annual surveys have been published (see, for example, the Policy Research Office of the Central Committee of the Communist Party of China & the Ministry of Agriculture, 2001). Microdata of the surveys are also utilized in the literature on rural household income, poverty, and other issues of the rural economy.[20] Unfortunately, the availability of the data to researchers is very limited.

Other public agencies 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 that focuses on reproductive health and economic well-being.[21]

4. Researcher initiated databases[22]

While data collected by the NBS is widely used by policy makers, international agencies and many researchers, it also has limitations. Some of these limitations are discussed above. Here, we can stress that individual researchers and research groups cannot count on having access to microdata from the NBS. To the limitations already discussed, one should add that in many cases the NBS does not collect information matching the concepts and circumstances researchers are interested in. This provides the background for why many researchers have involved themselves in the data-collecting process. There are many examples of researchers collecting data for one geographical area at one point in time.[23] Limiting the fieldwork to one location and one point in time makes the sampling strategy and data collection manageable. A limitation with such a strategy, however, is the difficulty with generalizing the results to a wider territory; another is that such data typically do not make it possible to follow and analyze 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 seven 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, the method of collecting income and the main host institution, and in certain cases, related data archives. Much of the data caches are cross-sectional. However, in some cases, retrospective questions are also posed, and there are examples of a prospective panel design. We will first present each of the seven surveys and then compare them.

Of the seven surveys, the first were the China Household Income Project (CHIP) and the China Health and Nutrition Survey (CHNS), both of which started at the end of the 1980s. CHIP was initiated by a group of both international scholars and scholars who at the time were based at the Institute of Economics, Chinese Academy of Social Science, Beijing. Today, CHIP 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. Microdata from both are available to researchers, and a large number of studies using one or the other of the 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 the 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 the NBS's larger surveys and cover many, but not all, province level units; see Table 2. The coverage of province level units is smaller in CHNS and does not include, 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 have been collected.

[18] See http://css.sysu.edu.cn/Data. One study that maps labor force participation and unemployment in urban China using the CHIP-data presented in the next section is Liu (2012).

[19] From the early 1990s, the survey has been 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 Centre for Rural Economy (RCRE) at the Ministry of Agriculture.

[20] For recent literature utilizing the merits of panel data characteristics, see, for example, Tao, Liu, Su, and Lu (2011). See Bramall (2001) for comparisons of income data of the NBS rural household survey and the fixed observation point survey.

[21] 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).

[22] An earlier version of this section was presented at the 2012 Asian Historical Economics Conference (September 13–15, 2012, Hitotsubashi University, Tokyo).

[23] For example, Gustafsson and Li (2006) survey 25 articles published in academic journals from 1997 to 2003 that have used such data. Subsequently, many articles using this approach to data access have been published.

Being a multi-purpose survey containing rich information, CHNS has had broad use outside the circle of economists. Many CHIP users have worked with concepts of income that are broader than that adopted by the NBS and more in line with what is used in international studies. This broader use has been accomplished 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, there is income information for the years 1988, 1995, 2002 and 2007. A new wave for 2013 is in preparation as this text is being written. The first three waves also include information on household wealth. As shown in Table 2, all waves of CHIP consist of separate surveys on rural and urban households. The 2002 and 2007 waves 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, Luo, Wei, and Yue (2008) and Luo, Li, Sicular, Deng, and Yue (2013). The utilization of microdata from the CHIP surveys can be applied for from ICPSR and the China Institute of Income Distribution at BNU.[24] Microdata from CHIP 2002 have recently been made available from the LIS Cross National Data Centre in Luxembourg, where variable definitions, as far as possible, have been harmonized with similar surveys made for other countries in order to facilitate cross-country comparisons of earnings, income and poverty.[25]

The panel survey Rural–Urban Migration in China (and Indonesia — RUMiC/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 workers in the city.[26] The panel survey for China initiated by the Australian National University and Beijing Normal University has the first survey for 2008. It 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 the NBS) and one migrant survey (conducted by a professional survey organization).[27] Along with standard questions on household characteristics, special questions on education, income and migration history are included in the migrant survey.

The first wave of the Chinese General Social Survey (CGSS) concerns 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 the labor market and social security as well as topic modules that vary across surveys. Bian and Li (2012) provide further details on the background and sample design. The CGSS is basically a sociological survey, devoting much attention to subjective aspects of well-being, including questions on the perception of income inequality.

Recently, two nationwide panel surveys were launched by the Institute of Social Science Survey (ISSS) at Peking University. The first survey is the China Health and Retirement Longitudinal Study (CHARLS), and the second is the China Family Panel Study (CFPS). The CHARLS aims to collect nationally representative data of middle-aged and elderly individuals and families. After the pilot survey in Zhejiang and Gansu in 2008, the first nationwide baseline survey of CHARLS was conducted in 2011–2012 and the second in 2013. These surveys cover approximately 17,500 individuals of 10,000 households distributed in 150 counties/districts of 30 province level administrative units.[28] In addition to its panel data characteristics, the survey is unique in its focus on individuals aged 45 and over and their families. An interesting aspect is that similar surveys have been carried out in other East Asian countries that share the same problem of a rapidly aging population.[29]

The CFPS was launched in 2010 after preliminary studies conducted in 2008–2009. The study covers approximately 16,000 households living in 25 province level units. CFPS has taken advantage of its panel data characteristics and unique questionnaire design that combines main questionnaires (separate questionnaires for adults and children) with additional questionnaires focusing on specific population groups. The second wave of the survey was conducted in 2012, and the data from the first two surveys are available for public use.[30]

The China Household Finance Survey (CHFS), launched in 2011, is unique in its detailed information on household wealth, including financial assets and debts, housing assets, and assets for household production and business activities (Gan et al., 2013). In addition to household wealth, the survey also collects information on income and expenditure, including public and private insurances. Another distinct feature of the CHFS is quarterly-based complementary interviews on employment, assets, income, and expenditure that are designed in a two-year cycle of the main survey. The 2011 CHFS covered approximately 8400 households consisting of 29,500 individuals.

We will now compare the seven researcher's initialized surveys described in Table 1. Some have collected information at the community level (village, urban neighborhood) simultaneously with household data. Because 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, Lu, & Qi, 2009), it will be useful to employ a multilevel research framework by combining microdata and community-level/small-region data. The CFPS employs a multilevel survey design and

[24] See, 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/.

[25] See http://www.lisdatacenter.org/news-and-events/first-chinese-dataset-in-lis/.

[26] Here, we do not cover the parallel survey made in Indonesia. For further information, including the Indonesian survey, see http://rse.anu.edu.au/rumici/. Akguc, Giuletti, and Zimmermann (2013) provide much information on RUMiC, and some waves can be downloaded from the Institute for the Study of Labor, see http://idsc.iza.org/?page=27&id=58.

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

[28] For studies based on CHARLS (pilot version), see, for example, Smith, Shen, Strauss, Yang, and Zhao (2012).

[29] Similar surveys are the Japanese Study of Aging and Retirement (JSTAR) and the Korean Longitudinal Study of Aging (KLoSA). See the relevant web sites: http://www.rieti.go.jp/en/projects/jstar/; http://www.kli.re.kr/klosa/en/about/introduce.jsp.

[30] The results and descriptive statistics from the surveys are presented in Xie, Zhang, Li, Yu, and Ren (2013).

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	8, 404
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, Xinjiang	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), Griffin and Zhao (1993), Gustafsson et al. (2008), Li et al. (2008), Li et al. (2013), Luo et al. (2013), and Riskin et al. (2001).

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. For data provided by an official agency, the rural fixed observation-point survey (see Section 3) applies a multilevel framework.

Regarding survey design and sampling frame, CHIP and CFPS have the broadest spatial coverage in regard to sampling frame and are therefore, from this perspective, best suited to estimate nationwide income distribution. RUMiC is designed to represent regions of origin/destination of rural–urban migrants. The sampling frame of CGSS aims to represent the adult population aged 18 and over, and income is treated as a control variable rather than a focus variable. CHARLS focuses on individuals over age 45 and their families. CHNS has a narrower spatial coverage. However, as opposed to the other surveys, it consists of a long-range panel study and has collected detailed data on health, nutrition, and environmental hygiene. When using CHIP and CFPS for income distribution studies, users should consider whether to use appropriate population weights and adjust regional price differences (regional PPP). Although the first wave of CHFS provides broad coverage of provincial level administrative units, there might be bias caused by the relatively small sample size and some non-randomness in sampling when it is used to estimate the nationwide income distribution (Yue & Li, 2013).

For categories of households covered, the inclusion of rural–urban migrants is the key issue. CHIP has an advantage in that it includes migrant households in the sampling frame relatively early (from the 2002 wave on). The 2007 waves of CHIP and RUMiC complement each other by using a common sampling design, which covers fifteen cities distributed in nine provincial level units.[31] The sampling frame of migrant households of the 2007 CHIP and RUMiC surveys is based on a census survey using city maps, enabling it to cover various types of migrants, including those who have an unstable working status (see RUMiCI, 2007 for the detailed sampling procedure). The migrant samples of the 2002 wave of CHIP are made up of those who have a relatively stable working and living status, and a comparison of the two waves should be conducted carefully.[32]

[31] The cities covered are Guangzhou, Dongguan, Shenzhen, Zhengzhou, Luoyang, Hefei, Bengbu, Chongqing, Chengdu, Shanghai, Nanjing, Wuxi, Hangzhou, Ningbo, and Wuhan (RUMiCI, 2007).

[32] See Chapter 2 of Li et al. (2013) for comparisons of the distribution of migrant household income and national income distribution with/without including migrant households between 2002 and 2007.

Income information is collected differently in the seven surveys. Some use a single question, while others ask for particular income sources and rely on bookkeeping.[33] The latter method can be expected to provide the best information, and it is adopted by CHIP for its rural and urban samples meaning that the data collection takes advantage of a subsample of the NBS household survey, where information on major income components is derived from bookkeeping by the respondents that have been regularly monitored by enumerators of the NBS. Additional income components are collected by supplementary interviews. The same approach is used for the urban and rural household samples of RUMiC. In contrast to these two surveys, income data in CHNS, CFPS, CHARLS, and CHFS are collected by interview only. It should also be noted that the income questions of CFPS, CHARLS, and CHFS (main interview) are designed to be answered in ranges instead of absolute values to guard against "do not know" and refusal.

Some income components are more difficult to capture in a survey than others, such as the imputed rent of owner-occupied housing, in-kind income, transfer income, and property income. We will discuss each of these components, starting with the first. Along with the privatization of urban housing and the very rapid increase in housing prices, the imputed rent of owner-occupied housing has increased rapidly in China and now has a larger impact on overall income inequality (Sato, Sicular, & Yue, 2013). However, there are problems, not unique to China, in estimating the imputed rents of owner-occupied housing. The inputs in such calculations are typically information on the market rent and/or market value of housing, both of which can be significantly biased when coming from a respondent's estimates. The housing mortgage amount and the cost of housing ownership are also difficult to collect by interview. It can be noted that the data quality on the market rent and market value of housing as collected in CHIP is relatively high because they reflect information on the housing market at the city level captured by the local NBS bureaus.

The inclusion of in-kind income, especially the value of goods/services for own consumption, for rural households and urban self-employed households, is another income component facing a measurement problem. It is difficult to systematically cover quantities of in-kind income from an interview. Another issue is which prices to set for the goods/services households both produce and consume. Here, the NBS household survey has the advantage, as the amounts of agricultural products are captured by bookkeeping, and values are adopted by certain prices set by the NBS.[34]

There are also issues regarding how to measure transfer income. The probably most important of these is how to include income from social security and social assistance. Public transfer incomes such as pension, medical insurance, minimum living standard guarantee (Dibao), and other programs have become increasingly important both for urban and rural households. In addition, for rural households, subsidies for agricultural production (food grain production, agricultural production materials, purchasing of agricultural machines, and other local level subsidies), the subsidy for the Sloping Land Conversion Program, and other subsidies have become important income components.[35] It is especially difficult to capture payments such as reimbursements of medical expenses from public medical insurance programs. A careful questionnaire design is needed in order to monitor out-of-pocket expenses and their reimbursements. Finally, on the issue of income components, because market allocation has become increasingly important, it has become essential to try to capture income components such as interests and dividends as well as capital gains when selling housing and stocks.

5. Conclusions

In this paper, we have presented official data on household income as well as seven researcher initiated databases and commented on their characteristics. We have found that, to a large extent, the research initiated databases focus on various aspects and are thus complementary to each other. They also complement the data the NBS collects. We conclude this paper by summarizing several basic issues that both official surveys and researcher initiated surveys need to address.

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

Second, along with increases in income level and social mobility, probable biases due to non-response and underreporting of income (especially by the high-income group), which are universal issues in household surveys in other countries, will become increasingly important for Chinese household data. Improvements in the operation and sampling processes are needed for the NBS household survey. The NBS has already tried to improve the operation process, for example, by introducing web-based bookkeeping. Regarding possible downward bias caused by the underreporting of the 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 & Luo, 2011).

Third, some holes are evident with respect to the types of household statistics that are being regularly produced and published for China. One is that there is no official labor force survey that regularly publishes information on urban employment and unemployment. Another weak area is that data on wealth are inadequate in both the NBS household survey and the existing recent researcher initiated surveys. Wealth inequality has been increasingly important for the analyses of household well-being.

[33] See, for example, Micklewright and Schnepf (2010) for an analysis of the bias in income data collected by a single question.

[34] There has been some criticism that the price applied to agricultural products by the NBS is lower than the market price, and therefore, rural household income tends to be understated until the 1990s (Bramall, 2001). However, such possible bias can be considered to have become less important along with the reforms of the purchasing system of agricultural products. See also Benjamin, Brandt, and Giles (2005).

[35] See chapters 1 and 2 of Li et al. (2013) for an overview and description of changes in public transfers in the 2000s.

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 most likely has very strong implications for the intergenerational transmission of inequality.

References

Aaberge, R., & 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., & 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.

Akguc, M., Giuletti, C., & Zimmermann, K. (2013). The RUMiC longitudinal survey: Fostering research on labor markets in China. *Bonn: Institute for the Study of Labor (IZA), Discussion Paper No. 7860*.

Benjamin, D., Brandt, L., & Giles, J. (2005). The evolution of income inequality in rural China. *Economic Development and Cultural Change*, *53*(4), 769–824.

Bian, Y., & Li, L. (2012). The Chinese General Social Survey (2003–8). Sample designs and data evaluation. *Chinese Sociological Review*, *45*, 70–97.

Bramall, C. (2001). The quality of China's household income surveys. *The China Quarterly*, *167*, 689–705.

Brandt, L., & Holz, C. A. (2006). Spatial price differences in China: Estimates and implications. *Economic Development and Cultural Change*, *55*(1), 43–86.

Chan, K. W. (2009). The Chinese hukou system at 50. *Eurasian Geography and Economics*, *50*, 197–221.

Chan, K. W. (2012). Migration and development in China: Trends, geography and current issues. *Migration and Development*, *1*(2), 187–205.

Chen, S., & Ravallion, M. (1996). Data in transition: Assessing rural living standards in southern China. *China Economic Review*, *7*(1), 23–55.

Chi, W., & 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., & 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., & 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., & Zhang, M. (1993). Annex: The 1988 Household Sample Survey: Data description and availability. In K. Griffin, & R. Zhao (Eds.), *The Distribution of Income in China* (pp. 331–346). New York: St. Martin's Press.

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.

Gan, L., Yin, Z., Jia, N., Xu, S., Ma, S., & Zheng, Lu. (2013). *Data you need to know about China: Research report of China household finance survey 2012.* Berlin and Heidelberg: Springer.

Gao, W., & Smyth, R. (2012). *Returns to schooling in urban China, 2001–2010: Evidence from three waves of the China urban labor survey, Discussion Paper 50/12.* Australia: Monash University (available at: http://www.buseco.monash.edu.au/eco/research/papers/2012/5012returnsgaosmyth.pdf).

Gibson, J., Huang, J., & 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.

Giles, J., Park, A., & Zhang, J. (2005). What is China's true unemployment rate? *China Economic Review*, *16*, 149–170.

Griffin, K., & Zhao, R. (Eds.). (1993). *The distribution of income in China*. New York: St. Martin's Press (Chinese edition published as Zhongguo Jumin Shouru Fenpei Yanjiu, Beijing: Zhongguo Shehuikexue Chubanshe [China Social Science Press], 1994).

Group, Canberra (2011). *Handbook on household income statistics* (Second ed.). Geneva: United Nations.

Gustafsson, B., & Li, S. (2006). Surveys — Three ways of obtaining household income data. In M. Heimer, & S. Thogersen (Eds.), *Doing Fieldwork in China* (pp. 129–152). Copenhagen: NiAS Press.

Gustafsson, B., Li, S., & Sicular, T. (Eds.). (2008). *Inequality and public policy in China*. New York: Cambridge University Press (Chinese edition published as Zhongguo Jumin Shouru Fenpei Yanjiu III, Beijing: Beijing Shifan Daxue Chubanshe [Beijing Normal University Press], 2008).

Kung, J., & 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, H., Zhang, J., Sin, L. T., & Zhao, Y. (2006). Relative earnings of husbands and wives in urban China. *China Economic Review*, *17*(4), 412–431.

Li, S., & Luo, C. (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* (pp. 105–121). Cambridge, MA: Harvard University Press.

Li, S., & Luo, C. (2011). Zhongguo Shouruchaju Jiujing you Duoda? [How actually unequal is China?]. *Jingji Yanjiu [Economic Reserach], 2011–4*, 68–78.

Li, S., Luo, C., Wei, Z., & Yue, X. (2008). Appendix: the 1995 and 2002 household surveys: Sampling methods and data description. In B. Gustafsson, S. Li, & T. Sicular (Eds.), *Inequality and Public Policy in China* (pp. 337–353). New York: Cambridge University Press.

Li, S., Sato, H., & 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 Chubanshe [People's Press], 2013).

Li, S., & Zhao, Y. (2003). The decline of in-kind wage payments in urban China. *Journal of Chinese Economic and Business Studies*, *1*(2), 245–258.

Liu, Q. (2012). Unemployment and labor force participation in urban China. *China Economic Review*, *23*, 18–33.

Luo, C., Li, S., Sicular, T., Deng, Q., & Yue, X. (2013). The 2007 household surveys: Sampling, methods and data description. In S. Li, H. Sato, & T. Sicular (Eds.), *Rising Inequality in China. Challenges to a Harmonious Society*. New York: Cambridge University Press.

Luo, C., Yue, X., & Li, S. (2011). Dui Wang Xiaolu Huiseshouru Gusuan de Zhiyi [A critique on Wang Xiaolu's Estimation of the Gray income]. *Bijiao [Comparative Studies], 52*, 146–158.

Ma, J. (2012). Dali Tuijin Chengxiang Zhuhudiaocha Yitihua Gaige [The promotion of reforms toward urban–rural integrated household survey], Zhongguo Tonji [Chinese Statistics], 2012–3. (electric version). http://www.stats.gov.cn/tjshujia/zgtj/t20120331_402796201.htm (accessed in September 11, 2013).

McCulloch, N., & 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., & Xue, S. (2013). Economic reform, education expansion, and earnings inequality for urban males in China, 1988–2009. *Journal of Comparative Economics*, *41*, 227–244.

Micklewright, J., & Schnepf, S. V. (2010). How reliable are income data collected with a single question? *Journal of the Royal Statistical Society, Series A*, *173*(2), 409–429.

National Bureau of Statistics (NBS) (2007). *2005 nian Quanguo 1% Renkou Chouyangdiaocha Ziliao [The 2005 One-Percent Population Sampling Survey].* Beijing: China Statistical Press.

National Bureau of Statistics (NBS) (2011). *Zhongguo Zhuhudiaocha Nianjian 2011 [China Yearbook of Household Surveys 2011].* Beijing: Zhongguo Tongji Chubanshe [China Statistical Press].

National Bureau of Statistics (NBS) (2012). 2012 Quanguo Nongmingong Jiancediaocha Baogao [The National Rural–urban Migrant Monitoring Survey 2012]. (published electrically at the official website of NBS). http://www.stats.gov.cn/tjfx/jdfx/t20130527_402899251.htm (accessed in September 11, 2013).

National Bureau of Statistics (NBS) (2013). Ma Jiantang jiu 2012 nian Guominjingji Yunxing Qingkuang Da jizhe Wen [Press meeting by Ma Jiantangon the national economic conditions in 2012]. (published electrically at the official website of NBS). http://www.stats.gov.cn/tjgz/tjdt/201301/t20130118_17719.html (accessed in April 6, 2014).

National Bureau of Statistics (NBS) (2014). Guojia Tongjiju Juzhang Ma Jiantang jiu 2013 nian Quannian Guominjingji Yunxing Qingkuang Da Jizhe Wen [Press meeting by Ma Jiantang, Director of NBS, on the national economic conditions in 2013]. (published electrically at the official website of NBS). http://www.stats.gov.cn/tjgz/tjdt/201401/t20140120_502414.html (accessed in April 6, 2014).

National Health and Family Planning Commission (2010). Guanyu Jiaqiang Quanyuan Liudongrenkou Tongji, Kaizhan Liudongrenkou 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, & the Ministry of Agriculture (2001). *Quanguo Nongcun Shehui Jingji Dianxing Diaocha Shuju Huibian [National Rural Socio-economic Survey Data Collection]*. Beijing: Zhongguo Nongye Chubanshe [China Agricultural Press].
Riskin, K., Zhao, R., & Li, S. (Eds.). (2001). *China's retreat from equality: Income distribution and economic transition*. Armonk, New York: M. E. Sharpe (Chinese edition published as Zhongguo Jumin Shouru Fenpei Zaiyanjiu, Beijing: Zhongguo Caizheng Jingji Chubanshe [China Finance and Economics Press], 1999).
RUMiCI (2007). Rural–urban migrant project census manual. http://rse.anu.edu.au/rumici/pdf/Census%20manual_China_English08.pdf (accessed in march 25, 2014).
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., Sicular, T., & Yue, X. (2013). Housing ownership, incomes, and inequality in China, 2002–2007. In S. Li, H. Sato, & T. Sicular (Eds.), *Rising Inequality in China: Challenges to a Harmonious Society*. New York: Cambridge University Press.
Sicular, T., Yue, X., Gustafsson, B., & Li, S. (2007). The urban — rural income gap and inequality in China. *Review of Income and Wealth*, *53*(1), 93–126.
Smith, J. P., Shen, Y., Strauss, J., Yang, Z., & Zhao, Y. (2012). The effects of childhood health on adult health and SES in China. *Economic Development and Cultural Change*, *61*(1), 127–156.
Stiglitz, J., Sen, A., & 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., Liu, M., Su, F., & Lu, X. (2011). Grain procurement, tax instrument and peasant burdens during China's rural transition. *Journal of Contemporary China*, *20*(71), 659–677.
Treiman, D., Lu, Y., & Qi, Y. (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 Guominshouru Fenpei [Grey Income and National Income Distribution]. *Bijiao*, *48*, 1–29.
Wang, P. (2013). Guanyu Woguo Juminshouru Jini Xishu Cesuan de Jige Wenti [Several issues on the estimation of Gini coefficient of personal income in China]. (published electrically at the official web site of NBS). http://www.stats.gov.cn/ztjc/ztfx/grdd/201302/t20130201_59099.html (accessed in April 6, 2014).
Whyte, M. K. (2010). *One country, two societies. Rural — urban inequality in contemporary China*. Cambridge and London: Cambridge University Press.
Xie, Yu, Zhang, X., Li, J., Yu, X., & Ren, Q. (2013). *Zhongguo Minsheng Fazhan Baogao [China Report on People's Livelihood Development]*. Beijing: Beijing Daxue Chubanshe [Peking University Press].
Xie, Y., & Zhou, X. (2014). Income inequality in today's China. *Proceedings of the National Academy of Sciences of the United States* (Published ahead of Print April 28, 2014).
Yang, D. T. (1999). Urban-biased politics and rising income inequality in China. *American Economic Review*, *89*, 306–310.
Yue, X., & Li, S. (2013). Guanyu Jini Xishu Zhenglun de Huigu [A Review on the Gini Coefficient Debate]. http://www.ciidbnu.org/news/201307/20130708161954706.html (accessed in February 5, 2014).
Yue, X., Li, S., Wang, P., & Guan, B. (2007). *Toushi Zhongguo Nongcun Pinkun [Perspective on Rural Poverty in China]*. Beijing: Jingji Kexue Chubanshe [Economic Science Press].
Zhang, J., Zhao, Y., Park, A., & Song, X. (2005). Economic returns to schooling in urban China 1988 to 2001. *Journal of Comparative Economics*, *33*(4), 730–752.