东亚研究 | 동아시아학 | 東亜研究

CEAS INSIGHTS 1/2023

REVIEW OF LARGE-SCALE ASSESSMENTS FOR COMPARATIVE STUDIES: CHINESE DATASETS

Hua Li, Ka Lin, Jia'ai Tian

INTRODUCTION

Due to China's planned economy, there is a tradition of collecting statistical data in the country, but most of these are economic data. For social research, Lin Nan's survey about the quality of life, conducted in Tianjin and completed in 1987 and then completed in Shanghai in 1989, stimulated the growth of sociological studies about social indicators in contemporary China. From the late 1990s onwards, social surveys on people's living conditions have been widely conducted. For these surveys, Chinese scholars have constructed databases to be used for different purposes. The themes of these databases are comprehensive and are based on many aspects of social life. The applied survey indicators include economic ones on income and consumption but also in culture, values, lifestyle, family structure, and social class.

Many of these surveys require several rounds in order to consistently update the information. These resources include the Chinese General Social Survey (CGSS), China Health and Retirement Longitudinal Study (CHARLS) and China Family Panel Studies (CFPS). The data from these surveys provide pieces of evidence for researchers to discuss social change and the evolution of people's lifestyle in Chinese society and, thus, provide the data for policy analysis and evidence of issues for policymakers. Some of the data resources are also helpful for developing international comparisons. In the questionnaire design, Chinese researchers may adopt similar survey questions to those in other countries in order to make the data comparable. For instance, CHARLS adopts a number of variables that were included in the Survey of Health, Ageing and Retirement in Europe (SHARE), the Japanese Study of Aging and Retirement, and the English Longitudinal Survey on Aging (ELSA) in the UK. Those databases facilitate cross-country comparative studies around certain subjects.

The structure of this overview is organised into two sections; one is the survey databases for general social issues, and the other is databases for specific subjects (or topics). The data of general surveys stay at the system level and are about social conditions with the purpose of drawing a comprehensive picture of people's lives. We also look at the databases in the categories of thematic subjects such as ageing, health, and religion. Most of these reviewed databases are updated periodically for the need of longitudinal studies. On the basis of this review, we will discuss the strengths and weaknesses of these databases in meeting the needs of academic research.

GENERAL SOCIAL SURVEY DATASETS

CHINESE GENERAL SOCIAL SURVEY (CGSS)

The CGSS survey was initiated in 2003 by the China Survey and Data Centre of Renmin University of China. This CGSS program is one of the earliest examples of a comprehensive social survey completed with a national scope in China. The CGSS survey project is supported by the state's 985 Fund and the Scientific Research Fund of the Renmin University of China. The main purpose of this survey project is to reflect changes in social structures and individual households in urban and rural areas. For this purpose, the CGSS survey investigates indicators of people's lifestyles, living conditions and activities at the community, family, and individual levels.

In the first round of this survey, completed in 2003, data from over 10,000 households in various provinces, municipalities and autonomous regions of mainland China were collected. Later, this survey has been conducted annually between 2004 and 2022. The CGSS survey used three modes of sampling in 2003–2006, 2008 and 2010. These three modes applied the multistage and probability proportional to size (PPS) sampling methods based on different frameworks, selected variables, and dissimilar stages of the sampling

process. The collected data of the survey is regarded as being of high quality and now has nationwide influence.

The questions included in the survey are constructed around three-dimensional factors, namely, social structure, quality of life, and the linkage between the two. The questionnaire consists of three modules: the core module, the thematic module and the extension module. In the core module, the survey questions are answered in full by participants and are relatively similar to the survey questions of previous years. The thematic module should also be answered by all participants but is only included once every five years. The extension module includes 1/3 or 1/4 of the participants, randomly selected from the respondents. The core module and the thematic module are functionally designed to describe and explain the traits of social change, while the extension module is used mainly to meet the need for cross-country comparative studies.

The data are widely used in social research and support a large number of research articles and scholarly data analysis work. For instance, a search for research articles based on the CGSS database through the China National Knowledge Infrastructure (CNKI) – the major resource of electronic publications in China – reveals that there have been 2,321 Chinese articles and 590 dissertations produced by the end of 2022. Furthermore, this CGSS program has been included in the research sources of the International Social Survey Cooperation Organization since 2006, and it is a platform for international exchange and cooperation in social survey studies.

The access to the datasets: http://www.cnsda.org/index.php

Related articles:
[1] Wang, J., Liang, C., & Li, K. (2020). Impact of internet use on elderly health: empirical study based on Chinese general social survey (CGSS) data. In Healthcare, MDPI, 8(4), 482.
[2] Xu, Z., Si, W., Song, H., Yao, L., Xiang, K., & Cheng, Z. (2022). Empirical Analysis of Population Urbanization and Residents' Life Satisfaction—Based on 2017 CGSS. Sustainability, 14(13), 7580.

CHINESE SOCIAL SURVEY (CSS)

The CSS is a sample survey project initiated in 2005 by the Institute of Sociology, Chinese Academy of Social Sciences. The purpose of this survey is to monitor social change nationally in regard to employment, family, social life, and social activity; thus, it provides information for social science research and government decision-making practices. The CSS survey is conducted periodically. The survey responses are collected from 31 provinces and autonomous regions, and among these regions, 151 cities and counties and 604 villages or urban neighbourhood committees. The sample size varied from 7,000 to over 10,000 families in each round, and the participants are mostly aged 18 to 69.

The CSS survey questionnaire is designed within a three-part framework: a basic module, a replacement module, and a hotspot module. The questions in the basic module remain unchanged in different rounds of surveys, with the indicators of personal data, labour and employment, family structure, and economic status. In the replacement module, the factors to be examined include social class, status mobility, social security, leisure and consumption, and social values. These indicators are repeated periodically. The hotspot module reflects changes in social focus and timing, the interests of different social groups, people's livelihood, urbanisation, and other topics related to general public concerns.

At the operational level, the CSS survey relies on universities and research institutions to build survey teams and train the team members as the survey investigators. This interview simulation training includes formulating on-site teamwork methods and has efficient logistical support. The project team members will anonymise the data to ensure the observations are neutral with no negative influence on the survey respondents. An updated version of the questionnaires was produced by CSS 2017 with the adoption of social quality theory into the logic of the questionnaires. This development enriches the content of the survey questions and, thanks to this new data presentation, raises the effectiveness of this survey in terms of analysing social and policy issues.

The CSS database enables the Chinese Academy of Social Sciences' Institute of Sociology to produce the Society of China Analysis and Forecast annually. The data also help researchers to develop social studies; as CNKI illustrated, there were 772 journal articles published on CSS data in Chinese journals from 2008 to 2019. In particular, since 2017, the CSS surveys have included the factors of social quality measurements in the research design, and the questions have referred to a four-dimensional analytical examination of social quality regarding socioeconomic security, social cohesion, social inclusion, and social empowerment. This extension of the logic of the survey makes the database especially useful for the study of social conditions and social quality conditions in society and communities.

The access to the datasets: http://www.cnsda.org/index.php

Related articles:
[1] Li, P. & Cui Y. (2022). Promoting the goal and path of common prosperity in stages starting from the countryside. Xuehai (in Chinese), 2022 (1), 51–59.
[2] Ren L. (2022). Middle-income groups and the identification of the social status in the middle class — A discussion from the approach of social quality theory. Journal of Huazhong University of Science and Technology (Social Science Edition) (in Chinese), 36 (4), 92–101.

DATASETS FOR AGEING AND GERONTOLOGY STUDIES

CHINA HEALTH AND RETIREMENT LONGITUDINAL STUDY (CHARLS)

CHARLS is operated by the National Development Research Institute, the Chinese Social Science Survey Centre, and the Youth League Committee at Peking University. The program is funded by the National Foundation with the aim of collecting a set of micro-data about families and individuals in the middle-aged and elderly age groups. The participants of the survey are aged 45 and above. The CHARLS national baseline survey was designed and carried out in 2011 and covered 150 cities and 450 villages and urban residential communities and about 17,000 respondents from 10,000 households. These survey informants are tracked every two to three years, and the data are available for researchers. Three waves of the survey were completed in 2011, 2013, and 2015, and the fourth wave was completed in 2018. The sample size covered 12,400 households, with 19,000 individual respondents.

The indicators of the CHARLS questionnaire concern personal information, family structure and economic support, health status, physical measurements, medical service utilisation and medical insurance, work, retirement and pensions, income, consumption, assets, and basic community conditions. This information is comprehensive, referring to many dimensions of ageing studies. The project adopts a multistage sampling process and further develops the electronic drawing software technology (CHALRS-GIS) by applying the mapping method to analyse the data of the village-level samples. The data are collected through three channels. The first channel is home visits by investigators, the second is community surveys, and the third is policy surveys.

With these designed data collection orientations, CHARLS has features of a multi-channel collection of survey information. For instance, in 2014, CHARLS collected the data on the life course of the elderly and used surveys to construct this database. This resulted in the first data record about life course surveys of the elderly in contemporary China. In 2016, the focus of the CHARLS survey concentrated on the history of individual life thus conducted interviews with the elderly about their reflections on life and work in rural China in the 1950s. These data resulted in a database of people's lives in the period of the newly established People's Republic of China.

CHARLS data are useful for international comparisons. In the design of its questionnaire, CHARLS made references to a number of international survey programs on ageing, including the US Health and Retirement Survey (HRS), ELSA, and SHARE. Thus, in CHARLS, many database indicators are comparable with international databases. This makes the CHARLS database is easy to use in international comparison studies. Accordingly, this program receives support from both domestic and international resources, such as China's National Foundation of Nature Sciences, Peking University, US National Institutes of Health, the World Bank, and the Chinese Medical Foundation.

The access to the datasets: http://charls.pku.edu.cn/en/

Related articles:

[1] Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: the China health and retirement longitudinal study (CHARLS). International Journal of Epidemiology, 43(1), 61–68.
[2] Strauss, J., Lei, X., Park, A., Shen, Y., Smith, J. P., Yang, Z., & Zhao, Y. (2010). Health outcomes and socio-economic status among the elderly in China: Evidence from the CHARLS Pilot. Journal of Population Ageing, 3(3), 111–142.

CHINA LONGITUDINAL AGING SOCIAL SURVEY (CLASS)

CLASS is operated by the China Survey and Data Centre (NSRC) of the Renmin University of China. It completed primary surveys in 2011 and 2012. In 2014, a nationwide baseline survey was completed with a total of 11,511 individuals from 28 provinces and autonomous regions. In 2016 and 2018, national surveys were conducted. At the community level, a total of 462 urban communities or villages responded. The survey indicators included the categories of family relationships, economic status, and social support for the elderly. These survey data present a general picture of older people's living conditions and socioeconomic backgrounds and also, reveal the risks and challenges of their lives during the ageing process.

The method of investigation is an on-site household interview. The indoor interview is conducted with families with people aged over 60, and the principle of selection is one household per selected interviewee. For the community-level surveys, the sampling selection follows the stratified multistage probability sampling method. The communities are selected in accordance with maps, and in the selected villages and urban communities, no more than 25 households are included for each community. The data are collected by the survey investigators of the Chinese Social Survey Network (CSSN), and the survey data are open for users for academic purposes. These data could be useful in the evaluation of the actual effects of the ageing process and the social policy measures applied to improve the quality of life for the elderly.

The access to the datasets: http://class.ruc.edu.cn/English/Home.htm

Related articles:

[1] Feng, Z., Phillips, D. R., & Jones, K. (2018). A geographical multivariable multilevel analysis of social exclusion among older people in China: Evidence from the China Longitudinal Aging Social Survey ageing study. The Geographical Journal, 184(4), 413–428.
[2] Wang, Z., Yang, H., Zheng, P., Liu, B., Guo, Z., Geng, S., & Hong, S. (2020). Life negative events and depressive symptoms: the China longitudinal ageing social survey. BMC Public Health, 20(1), 1–6.

CHINESE LONGITUDINAL HEALTHY LONGEVITY SURVEY (CLHLS)

The CLHLS survey program is organised by the Centre for Healthy Aging and Development and the National Development Research Institute, Peking University. The baseline survey was completed in 1998, and follow-up surveys were carried out in 2000, 2002, 2005, 2008–2009, 2011–2012, 2014, and 2017–2018. The covered area of the survey sample was 23 provinces and autonomous regions. The survey targeted groups of people aged 65 and above. For those aged 35–64, the survey used a method of random sampling, with nearly half of the total cities and counties randomly selected. In the round of 2017–2018, 15,874 people aged 65+ were interviewed, and the survey collected information on 2,226 elderly people who had died between 2014 and 2018.

The survey objectives are of two target groups: the elderly and their family members. The survey questions include indicators in regard to living conditions, socioeconomic backgrounds, family structures, sources of income, and economic status. The questions also include subjective indicators such as the self-assessment of health, quality of life, and cognitive function. Moreover, indicators of caring, health, services, and activities are also essential, and these indicators include daily activities, lifestyle, disease treatment, and medical expenses. In addition, the information gathered about the deceased elderly includes details such as the time of death and the cause of death.

The program is supported by various organisations, including the National Social Science Foundation of China in 2001, the United Nations Population Fund in 2002, the National Natural Science Foundation of China and Hong Kong Research Grants Council in 2004, and the Institute of Population Studies at the Max Planck Institute of Germany. This database is widely used in various disciplines. According to the registration data for users up to March 2020, 8,019 researchers (excluding their students and members of their research teams) have officially registered to use the database and have had free access to the data. Based on these data, scholars have published 17 monographs, 356 papers in SCI journals and Social Sciences Citation Index academic journals under anonymous international review, 455 papers in domestic journals, 35 doctoral theses and 104 master's theses that had passed the thesis defence, and 58 policy advisory reports.

The access to the datasets: https://opendata.pku.edu.cn/dataverse/CHADS

Related articles:
[1] Shi, Z., Zhang, T., Byles, J., Martin, S., Avery, J. C., & Taylor, A. W. (2015). Food habits, lifestyle factors and mortality among oldest old Chinese: The Chinese Longitudinal Healthy Longevity Survey (CLHLS). Nutrients, 7(9), 7562–7579.
[2] Gu, D. (2008). General data quality assessment of the CLHLS. Healthy Longevity in China, 39–60.

DATASETS FOR FAMILY STUDIES

CHINA FAMILY PANEL STUDIES (CFPS)

The CFPS survey focuses on the economic and non-economic welfare of Chinese residents. In 2008 and 2009, CFPS carried out initial interviews and follow-up interviews, respectively, in Beijing, Shanghai, and Guangdong. In 2010, CFPS formally implemented baseline surveys in 25 provinces, cities and autonomous regions across the country and, finally, completed interviews with 14,960 households and 42,590 individuals. The operation of the survey program is the responsibility of the Institution of Social Science Surveys at Peking University, and it uses a computational system to help complete the interviews. The program receives support from the National Nature Science Foundation of China and the research fund of Peking University.

The CFPS survey questionnaire consists of four subject areas: community, family, adult individuals, and children. The survey questions include various research topics with indicators of economic activity, educational attainment, family relationships and dynamics, population migration, and physical and mental health. Regarding children, all family members are defined in the baseline survey, and their future biological or adopted children are defined as CFPS gene members. By testing these indicators, researchers can disclose the all-around aspects of family life of their associated groups, helping to evaluate the quality of life using a substantial data basis.

The survey uses a multistage probability sample extracted by the implicit stratification method according to circular equidistant sampling. Each sample is extracted in three stages. The first two stages of the sampling selection use the data divided by the official administration units. In the third stage, using the maps of the residential areas in villages or communities, the households to be visited are selected by random selection. By choosing the targeted groups of surveys in this manner, a sample group is selected, and the coverage is then extended in order to ensure that 25 households are included in each selected unit.

The access to the datasets: https://www.isss.pku.edu.cn/cfps/en/index.htm

Related articles:
[1] Xie, Y., & Hu, J. (2014). An introduction to the China family panel studies (CFPS). Chinese Sociological Review, 47(1), 3–29.
[2] Xie, Y., & Lu, P. (2015). The sampling design of the China family panel studies (CFPS). Chinese Journal of Sociology, 1(4), 471–484.

CHINA HOUSEHOLD FINANCE SURVEY (CHFS)

The CHFS is a nationwide sample survey project conducted by the China Household Finance Survey and Research Centre, Southwestern University of Finance and Economics.

From 2011 to 2019, five surveys were completed (in 2011, 2013, 2015, 2017, and 2019), and the sixth round of investigation began in 2021. The samples were selected nationally; for instance, in 2011, the samples were collected from 367 counties (at the district and county levels) and 1,481 communities in 29 provinces. The survey covered 40,011 households, with 127,012 individuals being representative of national, provincial, and sub-provincial cities. The purpose of this survey series is to collect relevant information on the financial situation of individual households, and its focus is on family income at the micro-level.

The questionnaire is divided into four parts: demographic characteristics, assets and liabilities, insurance and security, and expenditure and income. Within this frame, the chosen indicators include housing assets and financial wealth, debt and credit constraints, income and consumption, social security and insurance, intergenerational transfer payments, demographic characteristics, employment, and payment habits. Thus, the survey looks at the economic and financial behaviours of individuals and households. This survey uses a method of case selection with three-stage PPS sampling in order to ensure the randomness and representativeness of the samples. This method ensures the availability of high-quality data on household finances for academic researchers and government decision-makers.

The access to the datasets: https://chfs.swufe.edu.cn/dczx.htm

Related articles:
[1] Chen, X., & Ji, X. (2017). The effect of house price on stock market participation in China: Evidence from the CHFS microdata. Emerging Markets Finance and Trade, 53(5), 1030–1044.
[2] Zhang, D. (2016). Understanding China from a household's perspective: Studies based on the China household finance survey (CHFS). Emerging Markets Finance and Trade, 52(8), 1725–1727.

CHINESE HOUSEHOLD INCOME PROJECT (CHIP)

CHIP is a household survey of family income conducted by the China Income Distribution Research Institute of Beijing Normal University. This survey has a long history and started its first round in 1989. Then, rounds of surveys were published in 1996, 2003, 2008 and 2014, and the surveys are named CHIP 1988, CHIP 1995, CHIP 2002, CHIP 2007 and CHIP 2013, respectively. Given the growing relevance of rural-to-urban migration, the 2002 survey added a section on the mobile population. The 2002 CHIP survey included three subgroups of samples. The 2007 survey adopted the same method and consisted of three parts: urban households, rural households, and the floating population.

In the round of 2019, the sixth round of the CHIP survey, the data on income and expenditure came from the year 2018, and the survey was thus named CHIP 2018. It should be noted that the data of CHIP 2018 were adopted from the large sample pool of the 2018 national population survey conducted by the National Statistics Bureau of China. These surveys were jointly organised by Chinese and foreign researchers as part of the 'Research on Income and Inequality in China'. This structure reflected the urban–rural divide in China and the increasing number of rural individuals migrating to urban areas over the past two decades. It covered 160,000 households in 31 provinces (municipalities and autonomous regions).

All CHIP survey data include information on urban and rural households, e.g., information on basic living and employment in an individual household, background information on income and expenditure, and various thematic issues. Through the stratification of sampling, the CHIP survey obtains the samples in line with the East, Central, and West regions (currently, the samples cover 15 provinces) based on a systematic sampling method. The collected information is concentrated on family and individual life. The survey includes the demographic data of individuals living in urban areas and information on urban households. The core contents of the survey are depicted by individual-level, household-level, and village-level economic variables. It also contains variables for surveying social networks and social activities in urban and rural areas.

The access to the datasets: https://www.icpsr.umich.edu/web/pages/

Related articles:
[1] Schettino, F., Gabriele, A., & Khan, H. A. (2021). Polarization and the middle class in China: A non-parametric evaluation using CHNS and CHIP data. Structural Change and Economic Dynamics, 57, 251–264.
[2] Huang, J., Jin, M., Deng, S., Guo, B., Zou, L., & Sherraden, M. (2013). Asset poverty in urban China: A study using the 2002 Chinese household income project. Journal of Social Policy, 42(4), 763–781.

DATASET FOR RELIGION STUDIES: CHINA RELIGION SURVEY (CRS)

The CRS is a significant survey project conducted by the NSRC of the Renmin University of China. The survey project was operated from 2013 to 2015 and focused on religious sites and their religious commitment in different cities. Respondents included officials from religious authorities and heads of religious sites. The aim of this program is to develop records and explanations about changes in the conditions of religion in China, with evidence from the collected data and comprehensive reflection on the development of religions in the period of social transformation. This information can assist national representatives with relevant academic research and religious policy formulation. The survey design is neither a fully local one nor a foreign one, but the questionnaire is required to include both international and local contexts.

This analysis used a multistage PPS sampling method based on districts and counties. There were 243 districts or counties involved, with about 5,000 religious individuals from 31 provinces and autonomous regions selected as survey participants. The size of the sample covered 1/10 of China's counties, giving special consideration to ethnic minority areas with special religious committees. Surveys were also completed in five Chinese megacities, with interviews on-site and in religious places. The changes in religious sites were recorded in the surveys. This information is valuable since there has been a lack of surveys about religious groups in China, which has led to a need for sustainable data for conducting studies on the roles and relations of religious organisations.

The access to the datasets is now unavailable online.

Related article:
[1] Wei, D. (2015). Renmin University of China "China Religion Survey Report" released. http://tzb.ruc.edu.cn/xwbd2/7cd9ad220ba7448fa99e9c46378690dd.htm. (in Chinese)

DATASET FOR HUMAN RESOURCE STUDIES: CHINA LABOUR-FORCE DYNAMICS SURVEY (CLDS)

The CLDS is carried out by the Social Science Survey Centre of Sun Yat-sen University. It is a biennial follow-up survey of Chinese urban and rural residents on labour-related issues. CLDS completed the pilot survey in Guangdong Province in 2011, the national baseline survey in 2012, and follow-up surveys in 2014 and 2016. The survey samples covered 29 provinces and cities in China, with a sample size of 401 villages, 14,214 households, and 23,594 individuals. These data included representatives from eastern, central and western China. It used a rotating sample tracking method to reflect China's rapid social change. Its sample was tracked for four rounds (2011, 2012, 2014, and 2016).

The CLDS approach adopts a multistage and multilevel probability sampling method proportional to the size of the labour force. The database contains longitudinal and cross-sectional data at the levels of individual labourers, families, and communities. CLDS targets the working-age population aged 15–64 and focuses on the economic status of labourers and status changes, the conditions of education, employment, labour rights, occupational mobility, occupational protection, and, more broadly, health, satisfaction, and happiness. It is an interdisciplinary project with indicators concerning factors such as political, economic, and social factors of development through labour relations and demographic structure, family property and consumption, family production, and land use. The database can provide qualified survey data as empirical support for theoretical and policy research.

The access to the datasets: http://css.sysu.edu.cn/

Related articles:
[1] Wang, F., & Zhang, C. (2020). Housing differentiation and subjective social status of Chinese urban homeowners: evidence from CLDS. Housing Studies, 36(4), 567–591.
[2] Chen, G., Qiu, F., Dai, X., Lan, H., & Song, J. (2022). Research on the influence of informal employment on residents' happiness in China: Empirical analysis based on CLDS data. International Journal of Environmental Research and Public Health, 19(15), 9085.

DATASET FOR EDUCATION STUDIES: THE CHINA EDUCATION PANEL SURVEY (CEPS)

The CEPS program is implemented by the China Survey and Data Centre of the Renmin University of China. The survey was carried out on a school basis, with 112 schools and 438 classes randomly selected for the survey nationally from 28 cities (at the county or district level of administration) across the country. The students of the selected classes were the participants in the surveys, and thus a total of about 20,000 students were engaged in the survey. The survey started in the academic year of 2013–2014 in the school calendar, which was set as the baseline, taking the 7th-grade and the 9th-grade cohorts as the targeted groups. The objectives of the survey included the sampling of students and their parents (or guardians), course teachers, and school managers.

The questions under inquiry included personal information of individuals, household registration by local citizens or migrants, the experience of growth, physical and mental health, the interaction between the parents and children, the experience of school learning, relationships with teachers or classmates, the orientation of social behaviours, educational expectations, family members, parents' basic information, habits and after-school activities, and community environments. The CEPS survey conducts comprehensive cognitive ability tests and basic personality tests and organises health and physical examinations for students. It intends to reveal the structure and output of an individual's education and explores the process of education and its impact on the life course of individuals. Moreover, the CEPS surveys propose to continue even after the respondents have graduated from junior high school, and the entire cycle of the surveys could take 30 years to complete.

The access to the datasets: http://ceps.ruc.edu.cn/English/Home.htm

Related articles:
[1] Guo, Y., Chen, Q., Zhai, S., & Pei, C. (2020). Does private tutoring improve student learning in China? Evidence from the China Education Panel Survey. Asia & the Pacific Policy Studies, 7(3), 322–343.

[2] Tani, M., Xu, L., & Zhu, Y. (2021). The impact of an un (der) funded inclusive education policy: Evidence from the 2013 China education panel survey. Journal of Economic Behavior & Organization, 190, 768–784.

DISCUSSION AND IMPLICATION

A survey database is a tool or instrument used in social research. In China, in spite of the relatively short period of revival from the 1990s until now, survey skills and framework designs regarding indicator systems have gradually matured. Many indicator systems are constructed based on international standards, and many survey frameworks on certain topics follow the general consensus achieved by scholars in the fields. In this two-decade period, the number of survey databases has increased, and they provide many high-quality data for social research. As observed from the data of this study, the number of databases is substantial, and many of these surveys have great influence. This is also because of the need to evaluate social conditions and to explore the trend of social change, which provide the dynamic in the development of social surveys. By using these survey data, particularly the CHARLS and CFPS surveys, the technique of quantitative studies is improved, and the basis for international comparison becomes more solid.

However, this database review brings to our attention various problems in the operation of these databases. The first issue is the sustainability of the databases. In this study, we included the major resources and also the most influential research databases, and many of them last for several rounds of surveys. However, compared to the large research need and large project funds invested in surveys, we may argue that the sustainability of these projects under operation is still a challenge for social surveys, as sustainable databases are still too few.

The second issue is the skilled design of longitudinal research data. In several cases, we observed difficulties in conducting longitudinal surveys because the standard of data selection and the needs of the survey may change. These issues impose a difficulty in conducting comparative studies from a life-course perspective and reveal the changes or trends for future development.

The third issue to be addressed is weak theoretical support. Among these databases, most of them follow the theories of life quality and social quality in the tradition of family surveys and means-tested social assistance. Some of these databases are constructed with the social indicators of psychological and health systems, but there are many surveys and questionnaires targeted at social problems and policy analysis without a theoretical backup. Thus, strengthening the theoretical analysis of the social indicators system should be a way to explore the paths of database construction. This issue can cause another problem, the fragmentation of the designed indicators and the collected data. The weakness in the theoretical work would inevitably cause fragmentation of the indicator systems and thus also undermine the value of the collected data. Additionally, this problem has a negative

impact on developing comparative studies with international databases, as the standard and criteria of the indicators are not comparable even with similar (or the same) indicators.

Moreover, there are also some technical issues with social survey databases. For example, many databases lack usage instructions, and both CFPS and CHARLS data assign weights to samples, but the use of these weights is not well explained, and many users are not sure how to use the databases. For those databases to be comprehensive, they may lose their focus. This feature is often related to the weak theoretical base of the survey in terms of research design. This phenomenon encourages the creation of thematic databases focused on various specific topics. For instance, at the current stage, we have many data on the elderly but very few on children, and thus, we need to add childcare as a theme in social surveys. However, these databases on specific issues have another problem, as many variables in the social context and demographic features may be missed.

Therefore, for developing the survey data system in China, we need to update the quality of the theories and skills at the stage of survey design and pursue stable support to ensure the sustainability of database operation. Thus, we need to strengthen the technical processing and make the data comparable. In addition, these databases need to be used in comparative studies and our analysis of these data needs to be developed. When designing the research questionnaire, it is also necessary to enhance the direction of international comparisons. This requires collaboration in social research between China and other countries. It is also necessary to increase exchanges and discussions along similar themes.

REFERENCES

Chen, G., Qiu, F., Dai, X., Lan, H., & Song, J. (2022). Research on the influence of informal employment on residents' happiness in China: Empirical analysis based on CLDS data. International Journal of Environmental Research and Public Health, 19(15), 9085.
Chen, X., & Ji, X. (2017). The effect of house price on stock market participation in China: Evidence from the CHFS microdata. Emerging Markets Finance and Trade, 53(5), 1030–1044.
Feng, Z., Phillips, D. R., & Jones, K. (2018). A geographical multivariable multilevel analysis of social exclusion among older people in China: Evidence from the China Longitudinal Aging Social Survey ageing study. The Geographical Journal, 184(4), 413–428.
Gu, D. (2008). General data quality assessment of the CLHLS. Healthy Longevity in China, 39–60.
Guo, Y., Chen, Q., Zhai, S., & Pei, C. (2020). Does private tutoring improve student learning in China?

Evidence from the China Education Panel Survey. Asia & the Pacific Policy Studies, 7(3), 322–343.

Huang, J., Jin, M., Deng, S., Guo, B., Zou, L., & Sherraden, M. (2013). Asset poverty in urban China: A study using the 2002 Chinese household income project. Journal of Social Policy, 42(4), 763–781.

Li, P. & Cui Y. (2022). Promoting the goal and path of common prosperity in stages starting from the countryside. Xuehai, (01), 51–59. (in Chinese)

Lin, N. & Lu, H. (1989). Discussion on the Structural Model of Social Indicators and Quality of Life—A Study on the Life of Urban Residents in Shanghai. Social Sciences in China, (4), 75–97. (in Chinese)

Lin, N. Wang, L. Pan, Y. & Yuan, G. (1987). The structure and indicators of quality of life——Analysis of data from the survey of thousands of households in Tianjin in 1985. Sociological Study, (6), 73–89. (in Chinese)

Ren L. (2022). Middle-income groups and middle-class social status identification—A discussion based on the theory of social quality. Journal of Huazhong University of Science and Technology (Social Science Edition), 36(04), 92–101. (in Chinese)

Schettino, F., Gabriele, A., & Khan, H. A. (2021). Polarization and the middle class in China: A non-parametric evaluation using CHNS and CHIP data. Structural Change and Economic Dynamics, 57, 251–264.

Shi, Z., Zhang, T., Byles, J., Martin, S., Avery, J. C., & Taylor, A. W. (2015). Food habits, lifestyle factors and mortality among oldest old Chinese: The Chinese Longitudinal Healthy Longevity Survey (CLHLS). Nutrients, 7(9), 7562–7579.

Strauss, J., Lei, X., Park, A., Shen, Y., Smith, J. P., Yang, Z., & Zhao, Y. (2010). Health outcomes and socio-economic status among the elderly in China: Evidence from the CHARLS Pilot. Journal of Population Ageing, 3(3), 111–142.

Tani, M., Xu, L., & Zhu, Y. (2021). The impact of an un (der) funded inclusive education policy: Evidence from the 2013 China education panel survey. Journal of Economic Behavior & Organization, 190, 768–784.

Wang, F., & Zhang, C. (2020). Housing differentiation and subjective social status of Chinese urban homeowners: evidence from CLDS. Housing Studies, 36(4), 567–591.

Wang, J., Liang, C., & Li, K. (2020, November). Impact of internet use on elderly health: empirical study based on Chinese general social survey (CGSS) data. In Healthcare (Vol. 8, No. 4, p. 482). MDPI.

Wang, Z., Yang, H., Zheng, P., Liu, B., Guo, Z., Geng, S., & Hong, S. (2020). Life negative events and depressive symptoms: the China longitudinal ageing social survey. BMC Public Health, 20(1), 1–6.

Wei, D. (2015). Renmin University of China "China Religion Survey Report" released. http://tzb.ruc.edu.cn/xwbd2/7cd9ad220ba7448fa99e9c46378690dd.htm. (in Chinese)

Xie, Y., & Hu, J. (2014). An introduction to the China family panel studies (CFPS). Chinese Sociological Review, 47(1), 3–29.

Xie, Y., & Lu, P. (2015). The sampling design of the China family panel studies (CFPS). Chinese Journal of Sociology, 1(4), 471–484.

Xu, Z., Si, W., Song, H., Yao, L., Xiang, K., & Cheng, Z. (2022). Empirical Analysis of Population Urbanization and Residents' Life Satisfaction—Based on 2017 CGSS. Sustainability, 14(13), 7580.

Zhang, D. (2016). Understanding China from a household's perspective: Studies based on the China household finance survey (CHFS). Emerging Markets Finance and Trade, 52(8), 1725–1727.

Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: the China health and retirement longitudinal study (CHARLS). International Journal of Epidemiology, 43(1), 61–68.