Published in final edited form as:

Econ Polit Stud. 2014 ; 2(2): 151–177. doi:10.1080/20954816.2014.11673849.

Rural Panel Surveys in Developing Countries: A Selective Review

Lingxin Hao^{*}, Weidong Wang^{**}, and Guihua Xie^{***}

^{*}Department of Sociology, Johns Hopkins University

^{**}Department of Sociology, Renmin University of China

^{***}Department of Sociology, Renmin University of China

Abstract

Rural panel surveys are the most appropriate means to provide data for studying the unprecedented rapid migration and urbanization currently taking place in China and other developing countries. To maximize heterogeneity in urbanization and development over the last three decades, we selected rural panel surveys from five Asian countries, including India, Indonesia, Nepal, Thailand, and China. This paper provides a selective review, focusing on their panel survey methodology, which is organized based on our proposed four basics of panel surveys – representativeness, retrospect-prospect, multilevel tracking, and temporality. We analyzed the strengths and weaknesses of the selected panel surveys to provide directions for designing future rural panel studies in China and elsewhere in the developing world.

Keywords

panel survey; survey methodology; rural development; developing country

Introduction

Urbanization in the developing world has accelerated in the new millennium (World Bank 2000). Both the rate and size of urbanized rural population growth are unprecedented. As rapid technological and political changes are transforming the global economy, the world's urban population has reached 50%, amounting to 3 billion (Cohen 2006). In China, foreign direct investment and land grabbing have uprooted much of the rural workforce and created the largest single migration to urban and peri-urban areas in world history (Hao 2012). Debates have focused on the lagged infrastructure development in megacities where rural migrants are concentrated or the underdevelopment of urban conditions in small- and medium-sized cities that also absorb scores of migrants, which center on the urban poor (Cohen 2006; Henderson 2002). Less addressed have been the impacts of rural migration and urbanization on the development and living conditions of rural communities (Lin 2003). Correspondingly less effort has been made to collect panel data necessary for studying rural development as a consequence of outmigration and urbanization.

The large exodus of the rural labor force has generated questions about the conventional wisdom of absorbing surplus labor locally and transforming rural areas via small-scale township and village enterprises (Bai 2003; Gu and Li 1998; Wen 2000). Findings regarding the net impact of migration and remittances on rural development are mixed. Although a 1997 survey of 13 counties in China found that some returned migrants used migration as a means of capital accumulation (Wang et al. 2003), analysis of the 1995 survey data from Hebei and Liaoning in China showed a negative net impact of migration and remittances on agricultural production (Rozelle et al. 1999). The deteriorating irrigation infrastructure because of a lack of state and collective fund for maintenance has created a pressing need for agricultural loans, but rural families have few collaterals and credits. How might the household production economy handle economic risks and credit market failure? How might remittances promote, rather than dampen, rural economic and social development? How might rural production modernize and attract young, educated, skilled rural labor force? Answers to these questions require information on the development process actually happening in villages, but most social surveys in China thus far have focused on urban areas. Questions about rural communities and families have been supplemental, and peripheral at the best.

The purpose of this paper is to offer a selective review of extant rural panel surveys under the contemporary rapid urbanization and rural outmigration. This review will focus on the under-researched rural development and provide guidelines for designing national rural panel studies in China and elsewhere in the developing world. We draw experiences from panel studies throughout the world that collected data on the dynamics at individual, family, household, and village levels. Rural panel surveys serve a number of objectives. First, they facilitate researchers to uncover the causal processes and causal reciprocity between macro factors and micro behaviors. Second they can be used to test theories on rural-to-urban migration and the impact of outmigration on rural development. Third, their results can provide insights to policy makers to reduce the negative impact of migration and urbanization and enhance the wellbeing of rural population.

In the developed world, large-scale migration and urbanization occurred before longitudinal survey methodology was systematically developed, and thus we are left with no developed-country examples of rural panel surveys under urbanization and rural outmigration. In the developing world, migration and urbanization have speeded up in the past half a century, corresponding to the period when longitudinal survey methodology was developed. Because China and other developing countries share many features of rural outmigration and urbanization, rural panel surveys in developing countries can offer valuable experience. This review considers rural panel surveys in five Asian developing countries – China, India, Indonesia, Nepal, and Thailand, aiming to identify their strengths and weaknesses to point directions for designing new rural panel surveys in China and everywhere in developing countries.

The organization of the paper is as follows. We start with addressing the importance of a rural-oriented perspective in research on rural development under rapid urbanization and rural outmigration in Section 2. Section 3 proposes four basics of panel surveys to guide our selection of panel surveys and our review. Sections 4 – 8 are devoted to reviewing the rural

panel surveys in the five Asian developing countries, each of which starts with an overview and then details one or two basics of panel surveys. Section 9 compares the strengths of the reviewed rural panel surveys and concludes with future directions for the design of new rural panel surveys.

A Rural-Oriented Perspective

Much of the urbanization and rural outmigration literature has been urban-oriented with a focus on urban destinations (e.g., Liang and Ma 2004). Although the urban-oriented perspective surely provides useful information on migrants' adaptation in destinations, we need knowledge of the entire migration process starting from the causal forces of the rural origin for outmigration to the consequences of migration for rural development. Thus a rural-oriented perspective is fundamental. Since the establishment of the household responsibility system in 1978, the initial marketization in rural areas was stalled because of the lack of credit and future markets development (Zhang 2003). The rural labor force has been exiting rural areas, responding to the upward pressures of demands for cheap labor from foreign direct investment. The burgeoning outmigration and subsequent return migration when migrants get old not only affect urban areas but more profoundly impact the development of the origin through changing labor-force compositions and resource flows. Migration also changes the motivations, aspirations, and lifestyles of rural people, a normative and cultural change generated by remittances and sustained by migrant networks. Under the rapidly changing macro-economic and labor-market conditions, rural social institutions are also in transition; for example, the stem family and old-age support institutions are weakening and the rural education and healthcare systems are stagnating (Zhang 2003). In the last decade, however, the law on arable land, the governmental investment in infrastructure of villages, and the rise of social insurance of healthcare and pensions may alleviate the deteriorating living conditions of rural areas. An urban-oriented perspective is likely to overlook the multifaceted social change happening in rural areas. In contrast, a rural-oriented perspective can capture the dynamics and swiftness of social change in the rural context to address the important issue of rural development.

We devote our review to panel surveys that follow micro social processes as they unfold under broader changing social contexts. The earliest large-scale, nationally representative panel studies is the Panel Study of Income Dynamics (PSID) in the United State, which has followed over 18,000 individuals in 5,000 households since 1968 and is still ongoing.¹ Referenced as a gold standard, the PSID is one of very few panel studies with data linking individuals across three generations during the postindustrial era in the United States. Advantages of panel data are multiple, and the most important one is that they make it possible to identify micro process (e.g., individual migration histories) under macro process (e.g., national and local economic restructuring and differential capital flows). Panel surveys are the only valid survey methodology to study causal relationships and causal processes because they contain micro processes under macro structures. Since PSID, the panel survey methodology has been used in developed countries to study a wide variety of micro social process. For example, the ongoing National Longitudinal Survey of Youth of the United

¹The PSID website is <http://psidonline.isr.umich.edu/>.

States has followed a nationally representative sample of 12,686 youths since 1979 (NLSY79)² and a later cohort of 8,984 youths since 1997 (NLSY97)³. These two panel surveys provide long event-history data for studies of the labor-market experiences of youth cohorts under economic restructuring in the United States, based on which important policies have been made. A rural panel survey in the United States is the Iowa Youth and Family Project⁴, conducted in 1989–1992 with a sample of 451 families in rural Iowa, to investigate the human consequences of the Farm Crisis. Since the United States has undergone urbanization and the agricultural economy differs from China's rural economy, this survey has little relevance to our purpose.

For our purpose of guiding the design of rural panel surveys to understand the linkage among rapid urbanization, sizable rural outmigration, and stalled rural development, we selected rural panel surveys conducted in Asian developing countries. Regional similarities in the global economic and political system suggest the usefulness of a focus on Asian developing countries such that the global position is kept relatively stable. Within Asia, we maximized political, economic, and social heterogeneity when selecting five countries – India, Indonesia, Nepal, Thailand, and China. For example, India and Indonesia are former colonies while Thailand and Nepal are still under monarchy. While a democratic political system has long been established in India, the king is the head of the present-day Thailand state. We also maximized heterogeneity in demographics, development levels and urbanization paces, summarized in Table 1. Drawing from the human development report (World Bank 2012), we show the major indicators in 1980 and 2012 to capture the diverse development and urbanization. In 2012, the five countries' human development index ranked below the 100th place with China taking the 101th, Thailand the 103th, Indonesia the 121th, India the 136th, and Nepal the 157th. China and India are the most populated countries but India's population grows much faster than China's. The population growth is also rapid in Indonesia, Nepal, and Thailand. Given the large population base, China's urbanization is very fast, from 19.4% in 1980 to 51.9% in 2012, contributed by enormous rural-urban migration. In contrast, the urbanization pace has been slow in India and Thailand, which is even slower than Nepal. The gross national income per capita (GNIPC) in 1980 and 2012 indicates economic development growth and the current level, which are the greatest for China and smallest for Nepal. This ranking is true for educational development measured by adults' years of schooling, as well as Carbon dioxide emissions per capital capturing industrialization and air pollution. To capture technology, landline and mobile phone coverage is used to show the fastest diffusion in China and the highest level in Thailand. To capture international migration, we show remittances sent from abroad as the percentage of GDP. The only remittance-dependent country is Nepal, at 21.66% in 2012. Finally the social dimension is indicated by gender inequality in reproductive health, labor market, and empowerment, which suggest that Chinese women have advanced much ahead of their counterparts in other four countries, especially Indian women.

²The NLSY79 website is <http://www.bls.gov/nls/nlsy79.htm>.

³The NLSY97 website is <http://www.bls.gov/nls/nlsy97.htm>.

⁴The Iowa Youth and Family Project website is <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/26721>.

Given the diverse political, economic, social, developmental, and urbanizing dynamics in the five selected Asian countries, their rural panel surveys lend great relevance to our purpose of developing a systematic origin-oriented perspective toward urbanization, outmigration, and rural development. The rapidity of urbanization and fast-paced rural outmigration in these selected countries makes panel studies a highly effective method to study the causes and consequences of migration and urbanization occurring within a person's panel duration.

Methods

Under rapid urbanization and rural outmigration, many social processes are not stationary but vibrantly dynamic. Repeated cross-sectional data at different points in time may illuminate cohort trends but does not reveal micro processes and micro-meso-macro interactions. For example, it is hard to conduct empirical work on whether and how foreign direct investment flows drive rural outmigration decision and action. Because the rural population is exposed to and is a part of such a process, studies must follow rural individuals and their social contexts over time. Panel studies can exactly fulfill this need because they portray dynamics at the individual level (Angrist and Krueger 1999; Abadie 2005). Our review of rural panel studies starts with introducing a selected rural panel study in the five selected Asian countries and then discusses one or two basics that characterize the panel survey under review.

Panel surveys maintain a sample of individuals to collect data on measures of interest repeatedly over time (Stafford 2010). Panel data are collected from micro units on two or more occasions, commonly referred to as waves. The first wave is called the baseline and the subsequent waves are referred to as follow-ups. The baseline survey collects current and retrospective information of the initial condition and time-constant variables that do not need to repeat in follow-up waves. Both baseline and follow-ups ask a set of core questions, which are stable in terms of questions and wording across waves to allow temporal measures of continuity and change.

We propose four basics of panel surveys: representativeness, retrospect-prospect, multilevel tracking, and temporality. First, representativeness indicates that panel surveys maintain a representative, sufficiently large sample at the baseline as well as reasonably small attrition at the follow-ups. Some panel surveys may include a mechanism to refresh the sample to offer cross-sectional representativeness of every wave. Refreshing panel studies provide both repeated cross-sectional representative data at each wave and within-individual changes (Bailar 1989). Second, retrospect-prospect highlights whether retrospective data are collected in tandem with prospective data to allow earlier initial time and longer temporality. Third, units of observation in panel designs often go beyond the micro units, which we term multilevel-tracking. Higher-level units include community in household-based survey, family, school, and neighborhood for school-based panel surveys, and firms and local governments in workplace-based surveys. Fourth, temporality, the length of the time spanning the panel varies greatly from a year to over a lifetime from birth. Intervals from wave to wave could be equal or unequal.

These four basics of panel surveys complement each other. Temporality and multilevel-tracking of panel studies are necessary for understanding distal causes at the societal or micro level and life course consequences of individuals and cohorts through multilevel and cross-level mechanisms during different phases of life. Given historical path-dependence, retrospect-prospect allows a longer view of the phenomenon of interest, with a special attention to include periods before and after a salient social change or a policy shift. Representativeness ensures that the conclusion can be made for the study population and provide evidence to test theories. These four basics of panel surveys together enable researchers to study causes, consequences and causal mechanisms at and cross macro, meso, and micro levels.

A National Panel Survey in India

National panel studies in developing countries with vast rural areas are still in its infancy. This section reviews the India Human Development Survey (IHDS)⁵ and focuses discussion on its innovative sampling design to guarantee a high level of representativeness and a long panel.

Overview

The IHDS is an omnibus, nationally representative panel survey of 215,154 individuals in 41,554 households, including 26,734 rural households in 1,503 villages and 14,820 urban households in 971 urban neighborhoods in all states and union territories in India except for an offshore state with 0.5% of the total population. Figure 1(a) shows the widespread of area coverage of the IHDS rural sample. The first wave of the IHDS was conducted from November 2004 to October 2005, and its second wave (a follow-up of all Wave-1 households) was fielded in 2011–2012. Villages and urban blocks (comprising of 150–200 households) were the primary sampling unit (PSU) from which the households were selected. The IHDS conducted one-hour interviews to two respondents in each household covered health, education, employment, economic status, marriage, fertility, gender relations, and social capital. Children aged 8–11 completed short reading, writing, and arithmetic tests. The interviews collected detailed current and retrospective information on family structure, fertility, economic status, education/training, transfers, migration, and many other topics. The survey also collected community-level data, including information on villages, schools, and medical facilities. While the IHDS exhibits many strengths according to our four basics of panel surveys, we choose to highlight its representativeness below.

An Example of Representativeness

The IHDS can be an example of high representativeness given its innovative design. The first wave of the IHDS includes a random sample from the Human Development Profile of India (HDPI), which was a cross-sectional rural survey conducted in 1993–1994. This existing random sample is combined with a new random sample selected from the 2004 population. A similar sampling design was used by the PSID. An important advantage of

⁵The IHDS website is <http://ihds.umd.edu>.

IHDS' sampling design is to produce a 2-time-point panel over 11 years for a part of the sample at Wave 1. This makes it possible to examine micro-level social processes under the condition of the social change from 1993–2004. Such a design is more costly than starting a panel with one new sample, because it is challenging to design and implement a follow-up of a previous sample from another survey. The benefit outweighed the cost, however, given the path-dependence of ongoing social change on the social change in the previous decade.

The previous HDPI is a cross-sectional, large-scale rural survey 1993–1994 (Shariff 1999). It used a stratified probability sample of 33,230 households, located in 16 major states, 195 districts, and 1,765 villages. While not designed as a panel study, the HDPI provided important information for re-contacting the original households, including the sampled village names, a list of all households and the household head's name, and the key socio-demographic characteristics of the households. This set of information proved to be crucial for the IHDS to take advantage of the HDPI and re-interview one half of the HDPI households.

Extra care and additional procedures were required to maintain the integrity of probability sampling and national representativeness. Including re-interviewing from the previous HDPI sample increases threats to representativeness, because (1) the HDPI did not cover all 34 states of India in 2004, (2) an elapsed time of 11 years saw much rural-to-urban transition of the originally sampled villages, and (3) re-contacting the original respondents were difficult given mortality, migration, and household splits during the long period. To address these challenges, the IHDS sample consists of four components (re-interviews, replacements, refreshers, and extensions), which together make up a probability sample representing the whole population of India in 2004 (Desai et al. 2010). Below we discuss these components one by one.

The first component consists of re-interviews. Because the previous HDPI sampling design was a stratified probability sample where village characteristics such as caste and religion were stratifying variables, randomly selecting 50% of the original households would be expensive given the need to revisit every village. A less costly alternative is to maintain the original stratified sampling design. The HDPI villages were randomly listed within districts under each stratum, and households were sampled within strata according to the village ranking. For example, within Stratum 1, households were sampled starting from the first village to the village at the middle of the randomly listed villages. The resulting sample includes “total” villages where all originally sampled households in a village were included and “partial” villages where part of the originally sampled households was included. All split households remaining in the village would be interviewed. Out of the 11,153 root households 2,440 households were split and living in the same village and were interviewed. This procedure illustrates both rigor and efficiency of the IHDS re-interviewing to uphold representativeness.

A second component, replacements, was at the household and village levels given the multilevel design of IHDS. At the household level, the estimated re-contact rate was 80% because of migration, mortality, and other reasons. The “replacement” households were to be selected within districts and strata, usually taken from the “partial” villages so that some

of these villages became “total” villages. At the village level, some villages were not identified because of alternative spellings, duplicate village names, and alternative village names. A principle was to find “replacement” villages for the lost-to-follow-up villages within the same district and stratum.

A third component includes refreshers. In each district where re-interviews were conducted, two new villages were randomly selected using the probability proportional to size (PPS). In each refreshed village 20 randomly selected households were interviewed. The refreshed sample was compared with the re-interviewed sample to check whether the distribution of variables such as caste, religion, education, and economic status would be similar.

A fourth component was extensions, taken from the states and union territories not covered by HDPI. New villages were selected using PPS, and 20 households were chosen from each refreshed village. The four components (re-interviews, replacements, refreshers, and extension) together form a probability sample of households that is nationally representative. Appropriate sample weights were created to enable inferences to the study population.

A major advantage of this 4-component design is that it yields a long panel that is capable of capturing micro-level social processes under the 1993–2010 social change. Another advantage is that it yields a nationally representative sample at the time of survey based on both the panel sample and the fresh cross-sectional sample. The IHDS emphasizes nationally representativeness with a virtually complete coverage of the nation. Its scale and scope are large by any standards. In spite of this, as with any large-scale, nationally representative survey, the IHDS cannot detail locality-specific contexts, nor can it address the social networks in which respondents are embedded. Rural panel surveys in other countries will offer experience on these respects.

A National Panel Survey in Indonesia

A similarly long panel study is the Indonesia Family Life Survey (IFLS)⁶. It has 4 waves with varying intervals from 1993–2007. The sample includes 22,000 individuals in 7,224 households. This study has strengths in representativeness, retrospect-prospective, and multilevel tracking.

Overview

The IFLS is a continuing longitudinal socioeconomic and health survey. Four waves of data from 1993–2007 are publically available. The baseline survey (IFLS1) selected 13 out of 26 provinces to maximize representation of the population and capture the cultural and socioeconomic diversity of Indonesia. The sampling design stratified provinces then randomly sampled within provinces. The IFLS1 sample of over 22,000 individuals in 7,224 households represents 83% of the population. In the second wave (IFLS2), 94.4% of the Wave-1 respondents were re-contacted and re-interviewed in 1997. A partial follow-up survey (IFLS2+) was conducted in 1998 with 25% of the sample to measure the immediate impact of Indonesia’s economic and political crisis. The third wave (IFLS3) was fielded in

⁶The IFLS website is <http://www.rand.org/labor/FLS/IFLS>.

2000 with a re-contact rate of 95.3% and a re-interview rate of 91% from the IFLS1 sample. The fourth wave (IFLS4) was fielded in 2007. The IFLS is multi-purpose, collecting data on household consumption, individual labor-market outcomes, health outcomes, and health-program availability and quality at the community level.

An Example of Retrospect-Prospect

The IFLS has a unique strength in retrospect-prospect at multiple levels. A benefit of gathering retrospective data beyond prospective data is to lengthen the temporality of panel surveys. While most panel surveys ask certain retrospective questions, such as educational history and migration history, very few panel surveys address retrospect at different levels as IFLS did. It is well known that changing community conditions shape the changes of individual behaviors, but household panel surveys seldom collect detailed community histories. IFLS is one of a few exceptions. In addition to prospective data collected at multiple points in time, IFLS provides retrospective information to form long longitudinal data at the community, facility, household, and individual levels. These data cover aspects of the physical and social environment, infrastructure, employment opportunities, food prices, access to health and educational facilities, and the quality and prices of services at those facilities. In addition, IFLS1 collected retrospective and current information on most topics at the micro level as well. IFLS2 updated the same information during 1993–1997. Retrospective and prospective micro-level data on schooling, family, migration, and work experience can be linked with community and facility longitudinal data to address many important questions regarding the impact of economic crises, governmental policies and social, economic, and environmental changes on the life chances of the respondents.

The IFLS retrospective data were collected from knowledgeable sources. IFLS interviewed official village/township leader and staff, and community leaders to collect longitudinal data on means of transportation, communications, sanitation infrastructure, agriculture and industry, history of the community, credit opportunities, community development activities, schools, and health facilities. In addition, IFLS interviewed staff representatives to collect facility longitudinal data on the prices and availability of healthcare services and the school, teacher, and student characteristics of elementary, junior-high, and senior-high schools. These collected data were supplemented with statistical data and standardized test scores of students' math and language. IFLS' retrospect at multiple levels naturally gives rise to a longer length of multilevel tracking to afford an opportunity to understand the interactive dynamics at the individual, household, and community levels.

A Local Rural Panel Survey in Nepal

We selected the Chitwan Valley Family Study (CVFS) in Nepal⁷ to review. Localized with thick contextual conditions seen from the local residents' eye, the CVFS is unique for multilevel tracking in an authentic sense.

⁷The CVFS website is <http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/4538>.

Overview

The CVFS is a multi-topic rural panel survey in Nepal conducted at three time points from 1996 to 2007. The study was designed to investigate the influence of changing social contexts on family formation, marriage, childbearing, and contraceptive use. The research focused on the extent to which how changes in the community produced changes in family formation and how individual behaviors changed along the life course. The CVFS interviewed all residents aged 15–59 and their spouses in the selected neighborhoods. An overall response rate of 97% generated 5,271 completed interviews. The latest development related to CVFS is a survey of out-migrants from Chitwan Valley to six countries in the Gulf regions (Girmere et al. 2013), pointing to the long-term benefit of a well-designed, well-implemented rural panel surveys.

An Example of Multilevel Tracking

The choice of the survey site in the Chitwan Valley of central Nepal was based on the fact that both the spread of mass education and a transition to fertility limitation had occurred within the lifetimes of the valley's residents. The study used an integrated application of survey and ethnographic methods to obtain detailed measures of social change occurring in the Chitwan Valley. The sample neighborhoods represented the neighborhoods in Western Chitwan, including each of the five major ethnic groups living in the area. All neighborhoods were sorted into three strata according to near, medium, and far distance from an urban center. The neighborhood sample was drawn in two stages. Stage 1 drew a sample of 30 settlements (each included 15–1000 households) from the three strata using 1991 census data as a sampling frame. Stage 2 sampled 171 neighborhoods, each of which had approximately 5 to 15 households. Whenever neighborhoods are concerned, geographic boundaries are always at issue. Usually physical boundaries of neighborhoods are determined by natural boundaries, such as farmlands, roads, streams, and irrigation canals. The CVFS, however, asked residents about their context without reference to boundaries. By avoiding the discussion of physical boundaries, residents could report any aspect of their context they considered relevant to their day-to-day lives, no matter how distant.

Neighborhood histories focused on educational social service facilities in the Chitwan Valley context. Schools were places of instruction for children of all ages. Health clinics were any places of care and healing, such as doctors' offices, hospitals, and health posts. Employers were any places that employed 10 or more people for pay. Cinemas were movie theaters or halls where movies were shown. Bus stops were any places where people could obtain a ride for pay on a vehicle. A combination of ethnographic and survey research methods was used to gather histories of 171 neighborhoods, 142 schools, 118 health services, and 20 bus routes.

A unique feature of the CVFS is its use of the neighborhood history calendars to collect retrospective community-level changes (Axinn et al. 1997). A benefit of neighborhood calendars was that they improved respondents' recalls. Calendars allowed respondents to see the sequence of events in their lives, which helped them recall the timing of events and related events. For example, to set timing cues, calendars record national, regional and neighborhood events that are salient to local residents but less relevant to the study.

Neighborhood events include changes in agricultural services (mills, cooperatives, and dairies), local women and youth groups, and the introduction of public banks. The calendar format facilitates respondents to use salient events to recall less salient events on the one hand and interviewers to cross check multiple interviews with a same respondent on the other. In this way the CVFS identified nearly 50 neighborhood characteristics, essential for explaining changing marriage timing, childbearing behavior, and contraceptive use.

In three visits to each neighborhood, calendar data were collected using unstructured or semi-structured interviews, group interviews, and archival cross checks. These mixed methods minimized response error while allowing the interviewer to interact naturally with respondents. After unstructured interviews with 2–4 informants of different ages, genders, and educational levels, group interviews encouraged respondents to generate collective agreements about the dates of important neighborhood changes. Efforts were made to encourage diverse participants. For instance, women were encouraged to join the discussion with their husbands, older uneducated residents were persuaded to join discussions with young educated residents, and residents of Buddhist ethnicities were asked to join discussions with residents of Hindu ethnicities. The collected information was cross-checked with local and regional public records about public schools, health-service facilities, electrification, and implementation of economic-development programs.

The major advantage of the CVFS neighborhood calendar method is that neighborhood events were based on neighborhood residents' collective memory. To face the challenge of gathering the collective memory of neighborhood-specific significant events, the CVFS used mixed methods in multiple visits to neighborhoods. Apparently such a thorough approach can be applied only to local surveys. With this unique feature, the CVFS has contributed greatly to advancing our understanding about how contextual changes as what the inhabitants conceive shape changes of individual and family behaviors.

A Local Rural Panel Survey in Thailand

Networks are important but hard to collect in survey. The Nang Rong socio-demographic survey⁸ has much on this respect to offer. The cross-level linkages between individuals at the micro level and networks at the meso level are unique in multilevel tracking.

Overview

The Nang Rong district (1300 square kilometers), a relatively poor, migration-prone farming area, is located in northeast Thailand. The Nang Rong survey is a rural panel survey of 34,035 people living in 5,860 households in 51 villages, conducted in 1984, 1994, and 2000. Nang Rong was the site for the Community Based Integrated Rural Development project (CBIRD), a Thai government initiative seeking to improve sanitary conditions and agricultural skills for this population of primarily subsistence rice farmers. The survey was intended to assess CBIRD in rapid social and environmental change and their impact on the lives of Nang Rong people including those who had migrated from Nang Rong.

⁸The Nang Rong website is www.cpc.unc.edu/projects/nangrong/.

The baseline survey collected the community profile of all Nang Rong villages, a complete household census in 51 villages on fertility and contraceptive behavior, migration history, and social networks, and an emigrant survey on migration history and adaptation. The survey instruments collected data on community profile, household survey, and migrant follow-up. First, the community questionnaire collected profiles of all 310 Nang Rong Villages. A group discussion method with selected village informants (village headman and assistants, village health volunteers, community development volunteers, housewives club officers) was employed to collect village information. Second, a household questionnaire collected census of all households in the 51 sample villages. Respondents for the household census were either the household head or the head's spouse. Third, the migrant follow-up component collected data from 1,781 emigrants from 22 of the 51 original villages to four urban destinations. The 22 villages were selected randomly within strata created by cross-classifying general location and distance from major paved roads. About 70 percent of the migrants were followed and interviewed. Land-use and land-cover data were derived from satellite spectral classifications. All levels of data were integrated in a Geographic Information System (GIS).

The 2000 survey continued to collect social and spatial data on population, environment, and land use from more than 50,000 individuals at the community, household, and individual levels. The community survey profiled all 346 Nang Rong villages in 2000. The household survey included a complete census of all persons in all households in the 51 baseline villages, which were split into 92 villages. An annual life history was collected for those aged 18–41. The migrant follow-up tracked more than 3,400 migrants from 22 villages to four urban destinations and to rural villages within the Nang Rong district. Furthermore, an enhanced geo-spatial component was incorporated to collect locational GPS data.

An Example of Multilevel Tracking

The Nang Rong panel survey is remarkable for its integrated spatial and environmental data. The Nang Rong database combines and integrates information from pixels, plots, lines, and patches to watersheds and landscapes. It covers a range of spatial, social, and temporal scales of people, places, and environments. The design and collection of relevant data and its integration into a wide-ranging, flexible, and spatially explicit GIS database is one of the major accomplishments of the Nang Rong surveys. The Nang Rong survey is both retrospective and prospective, covering a crucial period in the recent history of the country.

The linkage of households to their land use is a major advance in associating people, places, and environments. Using cadastral maps encoded into GIS from the Thailand office, as well as information from discussions with knowledgeable villagers, the boundaries of land parcels were drawn in village territories linked to households, thereby creating a spatially explicit connection of households and their used land during a three-year period. The linked data include a three-year crop history on these parcels of land and the current users/owners of neighboring parcels. Other data sources include aerial and satellite images from 1954 to the present. A GIS was developed to include a digital elevation model, a road network made by digitizing road types from 1984 base maps, and a hydrographic data layer augmented by satellite views of small lakes. The GIS provides data input, storage, management, retrieval,

analysis, visualization, and statistical and cartographic output. When incorporated into the GIS, data are stored in layers that can be analyzed separately or together, allowing users to gain a deeper understanding of multifaceted issues. The GIS also allows for a synthesis of various data types, facilitating analyses that investigate the complex interplay and feedbacks between spatial and socioeconomic data.

The Nang Rong surveys set up a feasible example of linking people and places to allow the feedback between spatial and socioeconomic data. Different from the temporal links of neighborhood events in the CVFS, the Nang Rong surveys used spatial environments in temporal orders to contextualize residents' behavior, as well as to reveal the consequences for spatial environments resulting from residents' changing behaviors.

Another equally important advantage of the Nang Rong surveys is its social network method that bridges micro and meso levels. Using a census approach in the village of origin, complete social network data were collected on multiple relations and at multiple levels of observation (individual, household, and village). Social network data were also collected from emigrants, who can be explicitly linked to their households and villages of origin. Network data were gathered in conjunction with follow-ups in 1994 and 2000.

The social networks method of Nang Rong survey addresses a number of issues in rural panel surveys. First, all longitudinal data sets lose some panel members in follow-ups, and this attrition is unlikely to be random. Household and community positions in social networks can be used to analyze how they influence the success in locating respondents, especially migrants. Second, remittances from migrants to their origin households and villages have long been of interest to migration researchers. A migrant's position and the family's position in social networks can be used to advance the understanding of the role remittances in rural development. In addition, the social networks method asks about the flow of money and goods from the place of origin to the migrant at the destination. Also, since the migrant's origin household may have changed, the network method collects the flow of resources between the migrant and all households that might have resulted from the fissuring of the original household. The consistency between givers and receivers could also be checked. Third, by distinguishing between recent and more established households in their social networks positions, the social networks method could address how existing social networks affect new residents.

Overall, the social networks method allows researchers to bring fresh insights to social demography, the life-course framework, and survey methodology. At the same time, the available social demography and life-course data allow researchers to better understand the causes and consequences of social networks. The major benefit is the feedbacks between social networks and the life course. As for neighborhood calendars and GIS, social networks data collection requires greater amounts of time and energy, which are appropriate for local panel surveys.

Rural Panel Surveys in China

This section introduces rural panel surveys in China, including a national rural fixed-site observational study, a comparative study of rural migrants in China and Indonesia, and a local panel study of rural children and families. After a brief introduction to each survey, we comment on their major strengths and weaknesses based on the four basics of panel surveys.

The National Rural Fixed Observational Site Survey (全国农村固定观察点) (NRFOS) is a large-scale, national rural panel study from 1984⁹, designed by the Agricultural Ministry and implemented by each of the 31 provinces. Purposive sampling was used to select 10 administrative villages per province, and within each village households were either systematically selected or all included. Information on villages' rural-to-urban transition, village mergers/splits, and household splits over the 25-year period was not collected. Questionnaires and household diaries of income and consumption were used to collect household data. Despite its household-level data, this panel database was used only for aggregate statistics. The survey focused on agricultural production with little information on non-agricultural production, outmigration, return migration, or the transformation of rural institutions such as labor market, family, school, and health-care services. Although long in temporality, the NRFOS is weak in, representativeness, retrospect-prospect, and multilevel tracking.

The ongoing Rural-Urban Migration in China and Indonesia (RUMICI) plans to collect data through five-year longitudinal surveys in China and Indonesia from 2008.¹⁰ The purpose is to study the impact of migration on income mobility, poverty alleviation, education, health and nutrition of children of migrants, the assimilation of migrants in cities, and the impact of migration on the rural communities they exited and the urban communities they entered. The sampling design for the China component selected four major destination provinces (Guangdong, Shanghai, Jiangsu, and Zhejiang), which received around 60% of total migrants, and six major sending provinces (Anhui, Henan, Hubei, Sichuan, Chongqing, Hebei), which sent more than 50% of total rural migrants. Migrants were selected from the workplace listing databases, using simple random sampling. Replacements were taken using the nearest-birth-date method. The survey pays special attention to changes in city boundaries and industrial structure. Different questionnaires are used to collect data from the household head or spouse, the parent or guardian of children, and individual adults and children. Given migrants' frequent geographic mobility (changing every 4–5 months on average), various tracking methods were used, such as sending a present every three months and holding a lottery during major holidays. Despite these efforts, the attrition rates among migrants were very high, 60% from 2008 to 2009 and 50% from 2009 to 2010. The attrition rates for rural households, however, were low at 5%. Two problems of the RUMICI threaten its representativeness: the sampling frame of migrants from workplace may miss many informally employed migrants and the attrition rates are high.

⁹The website of NRFOS is <http://ynszxc.gov.cn/szxc/CountyModel/ShowDocument.aspx?DepartmentId=1596&id=2494124>.

¹⁰The website of RUMICI is <http://rumici.anu.edu.au>.

The Gansu Survey of Children and Families (GSCF) is a longitudinal, multilevel study of rural children's welfare, including education, health, and psychosocial development.¹¹ Gansu is a Western province with relatively low educational and economic development. Figure 1(b) shows the area coverage of GSCF. Data were collected in 2000, 2004, 2007, and 2009. The study focused on (1) children's academic achievement, educational attitudes, behaviors, experiences, psychosocial development, and physical health; (2) attitudes and practices of children, families, and teachers about parenting and schooling; (3) the home, community, and school mechanisms by which poverty affects children's welfare; and (4) children's human-capital acquisition and subsequent labor outcomes, including migration. The GSCF used a 4-stage (county, township, village, and children) stratified probability sampling design. First, 20 counties were selected from 83 counties in Gansu using systematic sampling along the per-capita income ranked villages and households with probability proportionate to size (PPS). Second, out of the total 447 townships in the sampled 20 counties, 42 were selected with PPS. Third, 100 villages were selected from 700 villages within the sampled 42 townships with PPS. The fourth stage used the systematic random sampling method to select 30 children aged 9–12 from each of the 100 sampled villages (10 children were used as replacements). This produced the final sample of 2,000 children. The units of observation included the focal child, the mother, father, and homeroom teacher, and village leader. A census covered all teachers and principals in primary schools in each sample village. These data were linked to the child micro data. Three standardized tests on Chinese, math, and cognitive ability were given to one half of the sample children to provide objective measures of academic achievements. Three follow-up surveys were completed in 2004, 2007, and 2009. The website of this panel study has not published the re-contacting strategies, annual attrition rates, or modification of the survey contents. The GSCF baseline has high representativeness. Yet, we do not know much about its attritions at the three follow-ups and the thus the representativeness of the longitudinal data is called into questions. The GSCF design is particularly strong in its multilevel tracking, because it includes not only families and villages but also teachers and schools usually not tracked in household-based surveys such as the GSCF.

Summary and Future Directions

After introducing and reviewing the selected rural panel surveys in the five Asian countries, we now compare them according to the four basics of panel surveys: representativeness, retrospect-prospect, multilevel tracking, and temporality. Table 2 maps the strengths of each reviewed survey with the four basics of rural panel surveys.

First, representativeness based on probability sampling is required for making inference to the population. To judge the representativeness of a sample we ask whether the sampling design clearly documents the probability sampling method and procedure in details and whether the design addresses the specific implementation issues in the specific context. Among our selected rural panel surveys, multistage, stratified sampling was applied. Given the rapid urbanization and outmigration, issues of sample replacement, re-contact, and re-interview must be considered and handled with great care and must be appropriate for the

¹¹The website of GSCF is <http://china.pop.upenn.edu/Gansu>.

context. In this regard, IHDS (India) is a good example of national survey, followed by IFLS (Indonesia). For local probability sampling samples, we can consult CVFS (Nepal), Nang Rong (Thailand), and GSCF (China). Examples using non-probability sampling include NRFOS (China) at the national level and RUMICI (China and Indonesia) at the local level. A fundamental limitation of non-probability samples is that all conclusions are confined to the sample itself. Without inferences, it is difficult to test theories and offer policy recommendations. This is not to say that non-probability samples are useless. Insights from non-probability surveys can help modify existing conceptual models, which then can be tested with probability samples. Thus, despite their inefficiency, when probability sampling is not feasible and purposive sampling makes theoretical sense, they can contribute to the development of social-science research.

A panel survey of a probability sample is not representative until it has a high response rate in the first wave and low attrition rates (the proportion of baseline respondents lost to follow-up). Because data of multiple waves are used to construct panel data and attrition is usually non-random, a panel sample with a high attrition rate is no longer random even though the wave-1 sample is a probability sample with a high response rate. The panel component of IHDS (India) has a relatively low attrition rate of 18% after an 11-year interval. In addition, the IHDS design includes a component to check the randomness of the panel component with a new random sample of villages. The IFLS (Indonesia) also has low attrition rates. While the attrition rates for the urban and rural households of the RUMICI are relatively low (5% after a 1-year interval), the attrition rates for the migrant sample of RUMICI are high (60% for the first follow-up and 50% for the second follow-up) despite many efforts of re-contacting. This case indicates a serious challenge in panel surveys of migrants. Despite its importance, information about minimizing attrition by design and the actual attrition rates are not always readily available for panel surveys. For example, we were unable to find this information at the website of CVFS (Nepal), Nang Rong (Thailand), or GSCF (China). The attrition information about RUMICI was taken from a presentation of the PI. We emphasize that striving for low attrition rates and making actual attrition rates publically available should be important tasks for any panel surveys.

An issue related to representativeness is sample size, the absolute number of the unit of analysis (individual, family, or household). Sufficiently large sample sizes are necessary for high statistical power of analyses using the sample. When the sample is a national sample, we must consider regional and areal differences and population heterogeneity. Thus, for a national sample, it is not the total sample size but subsample sizes for different regions and areas and social groups that should be considered. While a precise calculation of test power must refer to the variables involved in a substantive analysis, the absolute sample size can give us a general idea of whether there will be a statistical power problem. Without sufficient statistical power, a probability sample would not be representative. The IHDS (India) and IFLS (Indonesia) are two examples of a large national probability sample size. The NRFOS (China) also has a large national sample, but it is a non-probability sample. Nang Rong (Thailand) has a large size of local sample whereas GSCF (China) has a relatively small sample size.

While some panel surveys draw representative sample at the baseline, others represent the study population at each wave. The need for refreshing arises from systematic attritions when self-selected respondents refuse to be followed up or new elements enter the study population after the baseline (e.g., migrants). This requires refreshing the sample at each follow-up survey. The IHDS (India) and IFLS (Indonesia) employed the refreshing method such that not only the baseline but also the follow-up waves are nationally representative.

A second basic of retrospect-prospect is a useful feature that can effectively lengthen the duration of a panel survey. Data are collected in two major modes: retrospective (things that happened in the past) and prospective (things that are happening when interviewed). The prospective data collection through the same set of questions is a defining part of all panel surveys. Two types of retrospective data collections are used. At the baseline wave of a panel survey, retrospective data can be collected by various means, such as community leader interviews in IFLS (Indonesia) and RUMICI (China), neighborhood history calendars in CVFS (Nepal), community GIS in Nang Rong (Thailand). In follow-up waves of a panel survey, retrospective data between waves can be collected as done in IFLS (Indonesia). For national panel surveys, IFLS (Indonesia) offers an excellent example to employ retrospective data collection methods at multiple levels of communities, facilities, households, and individuals. The CVFS (Nepal) and the Nang Rong (Thailand) are two good examples for local panel surveys to follow. We stress the importance of multiple levels when collecting retrospective data. The practice of leaving out community histories in retrospective data collection seriously reduces the use of routinely collected individual histories of marriage, fertility, education, and migration, because behavioral changes in these domains must be understood under the immediate changing community context.

A third basic of multilevel tracking focuses on levels of observation. Multilevel observations provide data to examine micro social activities under macro social conditions, and panel data provide the history at both macro and micro levels. These data offer opportunities to test theoretical models involving macro-to-micro and micro-to-macro transitions (Coleman 1990). Observations at the neighborhood, household, and individual levels were used in IHDS (India), IFLS (Indonesia), CVFS (Nepal), Nang Rong (Thailand), and RUMICI (China) and observations at the school, teacher, family, and child levels were used in GSCF (China).

A final basic of panel surveys is the temporality of data collection. A longer panel with waves at many time points would be ideal, such as the gold standard of the PSID (USA), an annual panel since 1968. Because of a late start, developing countries have not had a long panel as PSID. The IFLS (Indonesia) started in 1993 with a long-panel design and now four waves of high-quality panel data are available. The IHDS (India) of 2004–2012 had an innovative design to sample from a previous national probability sample in 1993, effectively lengthening the duration by 11 years. The reviewed rural panel surveys are of unequal intervals, presenting additional challenge to panel data analysis that assumes equal intervals.

Overall, our selective review of rural panel surveys in five Asian developing countries provides a systematic understanding of the important methodological issues commonly facing designers of rural panel surveys in developing countries. With rapid migration and

urbanization in the developing world, more national and local rural panel surveys are urgently needed. The continuous advancement of panel survey methodology will ensure that such designs are productive and efficient. The overview and special strengths of the rural panel surveys reviewed in this paper provide us with strengths to follow and weaknesses to avoid. These will lead to designs of rural panel surveys that are theoretically sound, conceptually clear, methodologically rigorous, and empirical feasible.

References

- Abadie A. Semiparametric difference-in-difference estimators. *Review of Economic Studies*. 2005; 72:1–19.
- Angrist, JD., Krueger, AB. Empirical strategies in labor economics. In: Ashenfelter, Card, D., editors. *Handbook of labor economics*. Vol. 3A. The Netherlands: Amsterdam; 1999. p. 1277-1366.
- Axinn WG, Barber JS, Ghimire DJ. The Neighborhood History Calendar: A Data Collection Method Designed for Dynamic Multilevel Modeling. *Sociological Methodology*. 1997; 27:355–392. [PubMed: 12348199]
- Nansheng, Bai. Urbanization in China (). *Management World*. 2003; 11:78–97.
- Bailar, BA. Information needs, surveys, and measurement errors. In: Kasprzyk, D., Duncan, G., Kalton, G., Singh, MP., editors. *Panel Surveys*. New York: Wiley; 1989. p. 1-24.
- Cohen, Barney. Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society*. 2006; 28:63–80.
- Coleman, James S. *Foundations of Social Theory*. Cambridge: Harvard University Press; 1990.
- Central China Normal University. Chinese Rural Survey: Ten Year's Observation on 100 Villages. Rural Problem Research Center; 2008. (中国农村调查: 百村十年观察 (2008年卷)) <http://www.docin.com/p-102402535.html#>, retrieved 12/19/2010
- Desai, Sonalde, Dubey, Amaresh, Joshi, BL., Sen, Mitali, Shariff, Abusaleh, Vanneman, Reeve. India Human Development Survey: Design and Data Quality. 2009. IHDS Technical Paper 1
- Gu, Shengzu, Li, Zhengyou. Institutional Analysis on Bottom-to-Top Urbanization in China (). *Social Science in China*. 1998; 2:60–90.
- Hannum, Emily, Kong. *Gansu Survey of Children and Families Papers*. University of Pennsylvania; 2007. Educational Resources and Impediments in Rural Gansu, China. retrieved 6/5/2013 http://repository.upenn.edu/cgi/viewcontent.cgi?article=1010&context=gansu_papers
- Hao, Lingxin. Cumulative Causation of Rural Migration and Initial Peri-Urbanization in China. *Chinese Sociological Review*. 2012; 44(3):6–33. [PubMed: 25525564]
- Henderson, Vernon. Urbanization in developing countries. *The World Bank Research Observer*. 2002; 17(1):89–112.
- Liang, Zai, Ma, Zhongdong D. China's Floating Population: New Evidence from the 2000 Census. *Population and Development Review*. 2004; 30:467–478.
- Yifu, Lin. The 'Rural, Agricultural and Farmer's Problems' and the Rural Development in the Future (). *Issues in Agricultural Economy*. 2003; 1:19–24.
- Scott, Rozelle J., Taylor, Edward, de Brauw, Alan. Migration, Remittances, and Agricultural Productivity in China. *American Economic Review*. 1999; 89(2):287–291.
- Shariff, A. *India: Human Development Report*. New Delhi: National Council of Applied Economic Research; 1999.
- Stafford, Frank P. Panel Surveys: Conducting Surveys Over Time. In: Marsden, PV., Wright, JD., editors. *Handbook of Survey Research*. Vol. 2010. Emerald Group Publishing Limited; 2010. p. 765-793.
- Wang Xiyu, Cui Chuanyi, Zhao, Yang. Jobbing and Homing: Employment Transfer and Rural Growth (). *Management World*. 2003; 7:99–109.
- Tiejun, Wen. Way to China's Urbanization and Its Related Institutions (). *China Opening Herald*. 2000; 5:21–23.

- World Bank. World Development Report 1999/2000: Entering the 21st Century. New York: Oxford University Press for the World Bank; 2000.
- World Bank. World Development Indicators 2012. Washington, D.C: World Bank; 2012. <http://data.worldbank.org><http://data.worldbank.org/data-catalog/world-development-indicators> [Accessed April, 2013]
- Xiaoshan, Zhang. Three Constraining Factors, One Basic Recognition and Two Kinds of Policy in Deepening Rural Reform and Promoting Rural Development (). Chinese Rural Economy. 2003; 1:4–1.

(a) IHDS Rural Sample Coverage (Desai et al. 2009)

(b) GSCF Sample Coverage (Hannum and Kong 2007)

Figure 1.
Area Coverage of IHDS (India) and GSCF (China) Rural Samples

Table 1**Development and Urbanization Heterogeneity of Selected Countries 1980–2012**

Indicator	China		India		Indonesia		Nepal		Thailand	
	1980	2012	1980	2012	1980	2012	1980	2012	1980	2012
HDI rank ^a	na	101	na	136	na	121	na	157	na	103
Population in 1,000,000	983	1,354	700	1,258	29	51	15	31	47	70
% urban population	19.4	51.9	23.1	31.6	22.1	51.5	6.1	17.3	26.8	34.4
GNIPC in international \$ ^b	524	7,945	880	3,284	1,278	4,154	566	1,137	2,199	7,722
Adult schooling years ^c	3.7	7.5	1.9	4.4	3.1	5.8	0.6	3.2	3.7	6.6
CO ₂ emission PC in tons ^d	1.5	5.3	0.5	1.5	0.6	1.7	0.0	0.1	0.8	4.2
Phone per 100 persons ^e	0.2	86.2	0.3	64.3	0.2	107.5	0.1	33.5	0.8	113.6
Remittance as % GDP ^f	0.00	0.89	1.45	3.21	0.00	0.98	0.00	21.66	1.18	0.55
Gender inequality ^g	0.579	0.213	0.636	0.610	0.549	0.494	0.649	0.484	0.458	0.360

Data source: Human Development Report, United Nation Global Development Network <http://hdr.undp.org/en/statistics/>.

^aThe human development index (HDI) is a single statistic to serve as a frame of reference for both social and economic development.

^bGNI per capita in PPP terms (constant 2005 international \$). Aggregate income of an economy generated by its production and its ownership of factors of production, less the incomes paid for the use of factors of production owned by the rest of the world, converted to international dollars using purchasing power parity (PPP) rates, divided by midyear population.

^cMean years of schooling (of adults) (years). Average number of years of education received by people ages 25 and older, converted from education attainment levels using official durations of each level.

^dCarbon dioxide emissions per capita (tons). Human-originated carbon dioxide emissions stemming from the burning of fossil fuels, gas flaring and the production of cement, including carbon dioxide emitted by forest biomass through depletion of forest areas, divided by mid-year population. The later figures were in 2008.

^eFixed and mobile telephone subscribers per 100 people (per100 people). Sum of telephone lines and mobile subscribers, expressed per 100 people.

^fRemittance inflows (Workers' remittances and compensation, total received) (% of GDP). Earnings and material resources transferred by international migrants or refugees to recipients in their country of origin or countries where they formerly resided.

^gGII: Gender Inequality Index, value. A composite measure reflecting inequality in achievements between women and men in three dimensions: reproductive health, empowerment and the labor market. Reported earlier figures were in 2000 or 2005.

Table 2
Methodological Summary of Selected Rural Panel Surveys in Developing Countries

Panel Survey Basics	Example
Representativeness	
Probability sampling	IHDS, IFLS, CVFS, Nang Rong, GSCF
Low attrition rate	IHDS, IFLS,
Large sample	IHDS, IFLS, Nang Rong
Refreshing	IHDS, IFLS
Retrospect-Prospect	IFLS, CVFS, Nang Rong
Multilevel Tracking	
Neighborhood, household, individual	IHDS, IFLS, CVFS, Nang Rong
Spatial	Nang Rong
Networks	Nang Rong
School, teacher, family, child	IFLS, GSCF
Temporality	
Long panel	IHDS (1993–2010); IFLS (1993–2007)
Number of waves	IHDS (3 waves); IFLS (4 waves)
Unequal intervals	IHDS (15.2 years); IFLS (4.3, 7 years)