Chinese Sociological Review

ISSN: 2162-0555 (Print) 2162-0563 (Online) Journal homepage: https://www.tandfonline.com/loi/mcsa20

New Approaches to Demographic Data Collection

Donald J. Treiman , Yao Lu & Yaqiang Qi

To cite this article: Donald J. Treiman , Yao Lu & Yaqiang Qi (2012) New Approaches to Demographic Data Collection, Chinese Sociological Review, 44:3, 56-92, DOI: 10.2753/CSA2162-0555440303

To link to this article: https://doi.org/10.2753/CSA2162-0555440303

Published online: 08 Dec 2014.

Chinese Sociological Review, vol. 44, no. 3, Spring 2012, pp. 56–92.

ISSN 2162–0555/2012 $9.50 + 0.00.
DOI 10.2753/CSA2162-0555440303

New Approaches to Demographic Data Collection

Donald J. Treiman, *University of California at Los Angeles*
Yao Lu, *Columbia University*
Yaqiang Qi, *Renmin University of China*

Abstract: *As population scientists have expanded the range of topics they study, increasingly considering the interrelationship between population phenomena and social, economic, and health conditions, they have expanded the kinds of data collected and have brought to bear new data collection techniques and procedures, often borrowed from other fields. These new approaches to demographic data collection are the concern of this article. We consider three main topics: new developments in sampling procedures, new developments in fieldwork procedures, and new developments in the kind of information collected in demographic and social surveys. We conclude with some comments on data sharing in the social research community and a list of major Chinese surveys publicly available to researchers. Where possible we illustrate our points with Chinese examples.*

In recent years the field of demography has been expanding rapidly and broadly, so that current demographic research now extends well beyond studies of fertility, mortality, and migration, the traditional concerns of the field. These developments have been of two main kinds. First, much recent work in many social science and biomedical disciplines (sociology, economics, geography, public health, and medicine) has adopted a *population-based approach* (Duncan 2008, 763–772). Sometimes this is as simple as basing analysis on data collected from samples of

---

The authors are grateful to Harry B.G. Ganzeboom, Perry Peifeng Hu, Zai Liang, William M. Mason, and Shige Song for very helpful comments, and to Ted Mouw for pointing us to some useful references. Preparation of this article was supported by a grant from the National Science Foundation to UCLA: SES0551279 ("Internal Migration in China: Data Collection and Analysis"). This is a revised version of a paper published in Chinese as "New Approaches to Demographic Data Collection." Pp. 385–439 in *Frontiers of Demography,* edited by Zai Liang. Beijing: Renmin University Press, 2012. An expanded version is available in English as Population Working Paper PWP-CCPR-2009–022. Los Angeles: UCLA, California Center for Population Research. Available at http://papers.ccpr.ucla.edu/.

general populations, for example, studies of the incidence of disease based on probability samples of all residents of geographic areas rather than, say, employees of an enterprise or visitors to a particular clinic. But sometimes demographic concepts are merged with substantive ideas drawn from other disciplines, such as studies of social mobility that take account of fertility differentials by social status (Mare and Maralani 2006). Second, there has been a great deal of *cross-fertilization* within these fields, resulting in the creation of new subspecialities. Two good examples are the recent emergence of *bio-demography,* the incorporation of biological data and research questions into demographic studies; and *spatial demography,* the incorporation of spatial data (geo-coded information and satellite-generated measures and images) into population studies.

As is clear from these two examples, an important locus of these developments has been in the expansion of the kind of data collected and in data collection techniques. These developments should be of interest not only to population scientists, but to many social scientists who do not think of themselves as demographers. In this article, we consider three main topics: new developments in sampling procedures, new developments in fieldwork procedures, and new developments in the kind of information collected in demographic surveys. We conclude with some comments on data sharing in the social research community. Although these developments are worldwide, where possible we refer to Chinese examples.

Sample Designs and Sampling Procedures

This section describes several developments in sample designs and sampling procedures, including the move toward true probability samples of general populations, sample design issues, and new sampling procedures, particularly those designed to sample rare or difficult-to-find populations.

Probability Sampling

The conventional design for sampling general populations is multistage area probability sampling in which a hierarchy of geographical units is identified (e.g., counties, townships, villages/neighborhoods, and households); units at each level are selected at random but (except at the household level) with the probability of selection proportional to the population size of the unit; and then individuals are selected to be interviewed. A national sample of 8,000 people might be created by selecting 100 counties, then two townships within each selected county, then two villages or neighborhoods within each selected township, and then 20 households within each village or neighborhood. Within each selected household, one individual (or more) would be selected at random to be interviewed. Multistage samples often are more complex than just described because they are *stratified* by features of the geographical units, for example, whether the county is in a coastal, central, or western province, or whether it is predominantly urban or rural.

Probability sampling is not a new method. However, it has increasingly come to be recognized in China that neither convenience samples (for example, visitors to a clinic) nor samples of so-called typical places provide an adequate basis for inference to broader populations; see the useful discussion by Duncan (2008, 764–767) showing how such samples can result in misleading inferences. While casual samples still are widely used in health research, because of the difficulty of collecting health information from general populations, there is increasing emphasis even among medical and public health researchers in securing population-based samples because of known biases in hospital- or clinic-based samples. Two studies of the effect of prenatal malnutrition caused by the Great Leap Forward famine on the prevalence of young-adult schizophrenia provide a useful contrast. St Clair et al. (2005) first suggested a link between the famine and schizophrenia, combining hospital records and historical population data for a single county. However, Song et al. (2009), in motivating their study based on data from the 1987 *Chinese National Disability Sample Survey,* a very large national probability sample survey, noted that hospital-based studies in China are vulnerable to the selectivity of hospital admissions, which favors urban high-status people. Because of the selectivity of hospital admissions, generalizations to the general population based on hospital records inevitably will be biased. Indeed, Song and his colleagues found sharp urban-rural differences in the linkage between exposure to the famine and schizophrenia that were entirely missed by St Clair and his colleagues.

Even when Chinese surveys purport to be based on general population samples, all too often the samples are chosen on a casual basis. Even major studies do this. For example, the well-known *China Health and Nutrition Study* (*CHNS*),[1] a joint project of the Carolina Population Center and the Chinese Center for Disease Control and Prevention, which has followed the same people since 1989, is based on semiprobability samples of nine provinces concentrated in Eastern and Central China with some primary sampling units (PSUs) chosen at random; others chosen purposively; and small areas within the PSUs chosen randomly. Thus, although the results from this study are often presented as pertaining to China as a whole, they hardly do so because, with the exception of Guizhou province, the entire Western Region of China, where health and nutrition problems are most severe,[2] is omitted from the sample. Fortunately, many studies conducted by the Ministry of Health and the Family Planning Commission do sample the entire population of China, but many of these studies are not readily available for use by the research community, a point we will return to. True national samples are less common in university-based studies, but they are becoming increasingly common. These are described in detail, with access information, in the Appendix.

Design Issues

The increasing opportunity for Chinese social scientists to carry out sample surveys lends added importance to the need for good sample design. Unfortunately, many Chinese surveys exhibit design flaws that make them less valuable than they could be. Thus, here we offer some points for consideration when designing new sample surveys.

Comparisons Are the Essence of Social Research

Unfortunately, researchers interested in a particular population often make the mistake of collecting data only from members of that population. The problem is that you cannot study a constant. If you want to know how the urban population compares to the rural population, you need a sample of both urban and rural people. Similarly, if you want to study, say, the determinants and consequences of internal migration, you need a sample of both migrants and nonmigrants. This is an extremely simple point, but one that often is ignored in data collection efforts. For example, in 2002–2003, the National Bureau of Statistics carried out the *National Survey of Temporary Migrant Children in China*, a nine-city survey of 6,343 migrant households with 0- to 18-year-old children. The restriction of the sample to the children of temporary migrants makes it impossible to assess the effect of migration on children's lives. Of course, a description of migrant children is possible, but without a comparison to nonmigrant children, it is difficult to interpret the meaning of the descriptive statistics. From an analytic perspective, all that can be done is to make internal comparisons, for example, comparing the conditions and experiences of migrant children residing in different cities (Lu 2007; Zou et al. 2005). But, presumably, the question that motivated the data collection was how migration affects children. Given the design, this question cannot be answered.

Partial Samples Are Vulnerable to Sample Selection Bias

There is an additional problem with partial samples, such as samples of migrants or samples of the urban population only—even descriptive statistics may be biased as a result of sample selection. Consider the claim that in urban China intergenerational social mobility is particularly pronounced relative to what is observed in other nations (Whyte and Parish 1984). The difficulty with such a claim is that in China rural-to-urban migration is often a consequence of upward mobility, the outcome of a process in which the brightest and most ambitious of those from peasant origins do well and go far in school and thus qualify for urban jobs. Thus, *de jure* urban-only samples tend to consist both of those from urban origins, who experience typical patterns of intergenerational mobility, and formal rural-to-urban migrants, that is, those from rural origins who have not only moved to urban areas, but also have acquired urban registration (*hukou*) and thus have experienced extreme upward mobility. Computations of mobility rates that fail to distinguish between these two groups will overstate the true rate of upward mobility in the Chinese population as a whole. See Wu and Treiman (2007, 416–417 and 440–443) for a more detailed discussion of this issue.

National Bureau of Statistics Rural and Urban Surveys Do Not Jointly Cover All of China

Apart from the sample selection bias inherent in analyzing samples selected on the basis of the dependent variable, there is an additional problem with conducting separate surveys of the urban and rural sectors, which is the conventional practice of the National Bureau of Statistics (NBS). The NBS has two separate data collection units, one for rural China and the other for urban China. In both cases, samples are drawn from the lowest administrative unit, *village committees* and *neighborhood committees*. However, not all of China is covered by village committees and neighborhood committees. In particular, commercial buildings, factories, hospitals, universities, and construction sites, as well as roadways and parks, are excluded, yet people live in these places. Thus, sampling only within the boundaries of villages or neighborhoods will exclude a portion of the population. This means that even when it is possible to combine NBS rural and urban surveys, which is seldom because, in general, common questionnaires are not used, the combination will not yield complete coverage of the population of China. Moreover, a 2003–2004 pilot study for a national survey of migration and health encountered situations in urban fringe areas in which long-time residents were regarded as under the purview of the village administration but newly arrived migrants were not (Treiman et al. 2006). In these areas, a part of the de facto resident population would be excluded from the sampling frame by definition.

So what to do? There are two possibilities. First, villages and neighborhoods could be sampled in the usual way, but in addition portions of a township not covered by either villages or neighborhood committees could be sampled. This would require careful estimation of the fraction of the population of a township living in such places to ensure that they are sampled in proportion to that fraction. However, in urban China, there is no longer a consistently defined unit below the township (*jiedao*): *neighborhood committees* have in some instances been replaced by larger *community committees*; high rise buildings often have their own *property owners' committees*; and so on. Thus, given that townships are too large to be enumerated (in 2000, they averaged about 25,000 in population size), a way of subsampling small areas (enumeration districts) within townships would need to be devised. One way to do this would be to sample administrative units down to the township level and then, as an additional stage, to sample smaller areal units within each township, using a version of the spatial sampling procedures described below. For an example of an application of such a procedure, see the documentation for the 2008 *IMHC* survey.

Many Surveys Include Only the Easy-to-Reach Population

Until fairly recently, the conventional way of carrying out general population surveys in China has been to sample from the local population register (*hukou*) and, sometimes, from lists of temporary residents. This is no longer an adequate sampling strategy, at least in urban areas, because there is a great deal of residential mobility in contemporary China, and many people fail to change their registration when they move. Computations from the *IMHC* survey show that 39 percent of those living in cities changed their residence in the past five years and that, of these, 61 percent moved within the same city the last time they moved. Moreover, 82 percent of the within-city movers lack local *hukou*, presumably because their *hukou* is still held by their old neighborhood committee. Informal rural-to-urban migrants also commonly lack local registration. Fully 75 percent of peasants (those with rural *hukou*) who reside in cities lack local *hukou*. These results are consistent with the findings of Landry and Shen (2005) who found, in a 2002 survey in Beijing, that 45 percent of their respondents would have been missed had the survey team relied on registration lists. Because of the growing inadequacies of *hukou* lists, adequate sampling of the Chinese population now requires household enumeration procedures. We will discuss these in the section on fieldwork procedures.

Apart from reliance on registration lists, another shortcoming of many Chinese surveys is their propensity to conflate families and households. Adequate coverage of the population requires that all members of a *dwelling unit* or *household* be enumerated.[3] The exclusion of roomers, servants, and employees from those eligible to be interviewed means that the sample does not adequately cover the population. Sometimes this is done on the ground that such people are so-called temporary members of their households and will be counted in their permanent households. But sample surveys almost never interview members who are living elsewhere, even when they are recorded in household rosters. Because of this, the migrant population is often undercounted. Migrants also tend to be undercounted in panel surveys because few surveys attempt to interview those who have moved away between waves. But this, too, results in inadequate coverage of the population, especially in contemporary China where more than 40 percent of the population live other than where they are registered and 21 percent of the population has moved within the last three years, the typical length of time between waves of panel studies (both computations are from the *IMHC*). A final way in which surveys fail to cover the entire population is a propensity to overlook people living in nonstandard dwellings. We will discuss this point in the section on fieldwork procedures.

Clustering and Stratification

Good sample designs strike a balance between coverage and cost. As already has been noted, general population surveys often are designed as multistage probability samples, with cases clustered within particular small geographic areas. The primary reason for clustering cases is to reduce costs—it is much less expensive to conduct interviews in several households within a limited geographical area than to send interviewers to widely scattered households. However, the typical consequence of clustering is to inflate standard errors.[4]

Sometimes surveys are excessively clustered so that the standard errors become very large or, to state the same point differently, the standard errors become similar to those expected from a much smaller simple random sample. A case in point is a recent cardiovascular study said to be based on a "nationally representative sample of 15,540 Chinese adults aged 35–74 years in 2000–01" (Gu et al. 2005, 1398–1399). Upon close inspection of the sampling description, it turns out to be based on 20 sampling points, a rural township and an urban street committee selected in each of 10 provinces via multistage sampling procedures that did not take account of differential population size. Even ignoring the fact that the provinces "were selected to be representative of the geographic and economic characteristics in their regions," that is, were selected purposively, this sample is so highly clustered that it almost certainly has the precision of a much smaller sample.[5]

Also as previously noted, multistage samples often are stratified by region, type of place, the socioeconomic level of the community, and so on. Stratification is used both to ensure that there is adequate coverage of all segments of the population, sometimes sampling segments at different rates (e.g., oversampling migrants in order to have enough migrants to be able to compare them to nonmigrants), and to reduce the deleterious effect of clustering because stratification generally reduces standard errors. Indeed, it sometimes is possible to dramatically reduce the size of design effects by stratifying on variables strongly correlated with the main analytic variables (for a Chinese example, see Treiman et al. 1996; see also the discussion of the same study in Treiman 2009, 207–212).

Household Versus Person Samples

Many social processes of interest to population scientists pertain to the behaviors and decisions of families and households rather than or in addition to those of individuals. For example, the decision to migrate is often seen as the outcome of family risk-diversification strategies—some grown children are kept home to work on the family farm while others go out for wage work, thus providing two alternative ways to protect against downside risk (Massey et al. 1998; Roberts 1997; Stark 1992). Surveys that take individuals as the sampling units may not be able to capture family structures and dynamics that affect outcomes. One common solution is to collect extensive information about other family and household members; but usually it is necessary to restrict the information collected to socio-demographic items because other family members cannot be expected to be reliable reporters regarding attitudes, values, preferences, and the reasons for making particular decisions, nor regarding retrospective information such as work, residential, medical, and other histories.

For this reason, surveys increasingly are designed to interview all the members of each sampled household. The *CHNS* uses such a design, as do the *Indonesian Family Life Survey* (*IFLS*) and the *Mexican Family Life Survey* (*MxFLS)*. Although relatively expensive, because the interviewer must spend many hours with various members of a household (but not nearly as expensive as samples of the same size where one adult per household is interviewed), such surveys yield rich information on the characteristics and beliefs of each member of the household and also permit direct biometric, anthropometric, and cognitive assessments (see the discussion below of developments regarding the kind of information collected in sample surveys). An additional advantage is that such designs permit the estimation of household fixed-effects models, which provide a useful way of purging analysis of the effect of all unmeasured variables that are constant across members of a household. See Treiman (2009, ch. 15) for a brief introduction to such models and Allison (2005) and Wooldridge (2006, chs. 13 and 14) for more extended explications. Finally, household samples do not suffer from a bias inherent in the most common household interview design in which households are chosen at random within small areas and one eligible person (say, one adult) per household is chosen at random to be interviewed. The problem with such designs is that in large households each individual has a smaller chance of being selected than do individuals in small households. The usual way this bias is corrected is by weighting the data for each person interviewed by the number of eligible respondents in the household. While this is a reasonable approach, it is not optimal because weighted data have larger standard errors than nonweighted data.[6]

The use of household samples becomes problematic when family members live apart, as when some members have "gone out" for work, or when individuals live in large collective households with nonfamily members, as is true of many migrants in China. In such cases, it may be preferable to sample individuals and then also to sample selected relatives, for example, siblings, spouses living apart, or parents. Doing this will, of course, increase the expense of carrying out the surveys, because it sometimes will be necessary to send interviewers to other locales to interview the selected relatives, but the by-now near universal use of mobile phones makes phone surveys with absent family members increasingly feasible. Interviewing absent family members one way or another is a useful way to gain the advantages of household samples in situations in which many families have members living in more than one dwelling unit.

New Sampling Methods

In this section, we consider some relatively new sampling strategies designed for situations in which conventional probability samples of households do not work well. These include spatial sampling, contingent quota sampling, and respondent-driven sampling.

Spatial Sampling

When an adequate sampling frame is not available, a useful alternative is to sample spatial units identified by global positioning system (GPS) coordinates, a method first used in studies of the environment—ocean temperatures, vegetation types, air pollution, etc. In such applications, the area to be studied is subdivided into a set of regular polygons, which are then selected at random. In studying human populations, however, investigators often want to take account of such factors as population density, in order to represent populations of people rather than populations of space, and also other aspects of population heterogeneity. These considerations add complexity to the sampling process, requiring measurements of density and heterogeneity and their incorporation into sampling designs. An interesting application is a respiratory health and demographic survey of households carried out in the Delhi metropolitan area in 2004 (Kumar 2007).

In a different sort of application, Landry and Shen (2005) used spatial sampling as a way of overcoming the limitations of *hukou* lists in contemporary Beijing, especially in the coverage of migrants. Their strategy was to overlay a map of Beijing with a grid showing degrees of latitude and longitude and to randomly sample, but with probability proportional to a separately derived estimate of population size, 50 *square minutes* (trapezoidal areas of approximately 1.4 by 1.8 kilometers at Beijing's latitude) as their PSUs and then, within each PSU, to randomly sample four *square seconds* (areas of approximately 54 meters by 54 meters) as secondary sampling units (SSUs). In each of the selected SSUs, doors were counted to create a list of addresses from which to select households for interviewing. This strategy has clear advantages with respect to the completeness of coverage. Landry and Shen compared their results to data obtained from the Beijing Area Survey, which had sampled the same districts in Beijing, and showed that, as noted earlier, 45 percent of the respondents to their spatial survey would have been missed by the Beijing Area Survey. See Gibson and McKenzie (2007, 222–225) for additional examples of the use of spatial data as a sampling tool.

But the method also has some disadvantages, of which two are substantial. First, the PSUs and SSUs do not conform to any administrative boundaries, making it difficult to utilize administrative data as a supplement to the data collected from respondents. Second, the SSUs selected often result in buildings being divided, partly in and partly out of an SSU, which requires awkward adjustments. A third limitation, reliance on a simple count of doors, can be overcome by changing enumeration procedures, a point we will return to below. Still, it is a method worth considering as an alternative or supplement to more conventional ways of implementing multistage sample designs.

Contingent Quota Sampling

Quota sampling is a way of obtaining sufficient numbers of respondents within sample strata to achieve a specified level of statistical power. It is similar to stratified probability sampling in that a specific number of subjects from different subgroups are recruited. The difference is that subjects are recruited by convenience rather than randomly. The quotas are predetermined (i.e., according to sex and age groups) and interviewers are instructed to fill their quotas as best they can for the randomly selected households. Once the quota for a subgroup is met, the interviewers no longer recruit subjects for that subgroup. More detailed discussions of the method can be found in Moser and Kalton (1971) and Kalton (1983).

While the nonrandom element is a major weakness of this approach and quota samples are essentially convenience samples, several studies suggest that quota sampling often represents a fair balance between technical rigor and the production of useful data (Cumming 1990; Moser and Kalton 1971; Owen et al. 1998). Although it is not as rigorous as probability sampling, it is often chosen as the most cost-effective means of sampling hard-to-reach or hidden populations. One experimental study explicitly evaluates the performance of quota sampling versus probability sampling and finds that in many situations quota sampling can be considered an acceptable alternative to probability sampling when the latter is impossible or difficult to accomplish (Cumming 1990). Quota sampling is especially useful when a sampling frame is not available. Another advantage of quota sampling is that researchers can be specific about the type of subjects desired for the study and can be assured that specific subgroups are represented adequately in the final study sample.

In addition to nonrandomness, quota samples are vulnerable to bias resulting from giving interviewers excessive latitude in the way they meet the defined quotas. For this reason, a modification, contingent quota sampling, often is employed, in which a small geographic area is specified, for example, the square seconds used by Landry and Shen (2005), and interviewers are instructed to secure interviews with specified numbers of people with particular characteristics, for example, men under age 40 or men age 40 and over, in each small area. The implicit assumption of this procedure is that, within small areas, there is no correlation between inclusion in the quota and other attributes of the respondents. This assumption certainly is more reasonable within small geographic areas than over wider, and presumably more heterogeneous, sampling points, but may nonetheless be problematic. This procedure was utilized for part of the *IMHC* and in several other recent surveys for which the data collection was done by the Beijing-based survey research firm, Millward Brown ACSR. Specifically, if an interviewer was unable to gain entrance to a high-rise apartment building (an increasingly common problem in Chinese cities), the interviewer was instructed to stand outside the building and approach residents as they entered the building, alternating males and females. The assumption of this approach is that those who agreed to interviews would be similar to members of the randomly selected households in the same buildings that interviewers were supposed to visit.

Respondent-Driven Sampling—A Modern Approach to "Snowball" Sampling

For certain research topics, it is necessary to collect data on difficult-to-find populations, such as prostitutes, HIV-positive individuals, jazz musicians, and so on. Such populations cannot be sampled in conventional ways because usually there is no complete list of the population to be sampled nor are targeted individuals sufficiently prevalent in the general population, even within small geographic areas, to make area probability sampling feasible. Moreover, when special sites are targeted, they tend to produce biased coverage because the kinds of individuals found at such sites tend to be a selected subset of all eligible individuals. Moreover, when the criterion for inclusion in the sample involves stigmatized behavior, such as prostitution or drug use, individuals may be reluctant to be interviewed.

Until recently, the only available sampling method was "snowball" (also called "chain-referral" or "network-based") sampling. In this method, a few individuals in the group to be sampled, say, prostitutes, are identified and interviewed. In the course of the interview, they are asked to identify other prostitutes, who are in turn interviewed and also are asked to identify still other prostitutes. In this way, the sample is expanded, like a snowball rolling downhill until the target size is reached.

The obvious limitation of this method is that the "sample" is not a probability sample of the targeted population. *Sociometric stars,* that is, those who are central to networks, are much more likely to be named, and they may well differ in their characteristics from more isolated individuals. Moreover, because of the tendency for people to be more likely to interact with others similar to themselves, the choice of initial contacts is likely to affect the distribution of characteristics in the realized sample. Thus, valid generalization from the sample to the population of interest is usually impossible (Erickson 1979).

However, in 1997, Douglas Heckathorn, a sociologist at Cornell University, proposed an important improvement, a method he called "respondent-driven sampling" (RDS). Heckathorn combines snowball sampling with a set of procedures and a statistical model that generates weights for individuals in the derived sample that correct for the fact that some individuals are more likely to be included than others. Initial contacts (usually chosen purposively) are interviewed and are then asked to nominate others, customarily by distributing coupons that are returned to the investigators by those willing to be interviewed. A crucial aspect of the procedure is to collect information from nominators on the number of eligible individuals in their networks, to be able to estimate population coverage. In a series of papers (Heckathorn 2007; Volz and Heckathorn 2008 and the papers cited there), Heckathorn and his colleagues have shown that, when certain (strong) assumptions are met, his method yields asymptotically unbiased population estimates when conventional statistical inference procedures are applied; see also Heckathorn's Web site (www.respondentdrivensampling.org).

There now has been considerable work evaluating RDS procedures, and in particular the assumptions required for correct statistical inference; see, in particular, Gile and Handcock (2010). Much of this work has focused on the impact of nonrandom selection of initial respondents and the effect of two properties of the resulting networks in overcoming bias introduced by initial nonrandom selection: the degree of subgroup segmentation within the target population (that is, the extent to which subsequent nominations approximate random selection) and the number of links in the referral chain (the more, the better).

Because it meets a strongly felt need—to be able to survey hard-to-locate populations—RDS has been widely adopted; Malekinejad et al. (2008) review more than 100 RDS-based studies in more than 30 nations. Merli's research on prostitutes in

Shanghai is a nice application of the method. She and her colleagues interviewed 522 female sex workers in Shanghai in 2007 using RDS procedures (Merli et al. 2010), which enabled them to reach those female sex workers who were not employed in specific venues such as karaoke bars, hotels, and saunas. However, her data did not meet the assumptions specified by Heckathorn for valid statistical inference. This is a common outcome and suggests that RDS must be used cautiously. For additional discussion of the implementation of RDS designs, see Johnston et al. (2008).

Screening

Under certain circumstances it may be possible to screen for members of rare populations in the course of carrying out surveys for other purposes. If, for example, a very large general population survey is mounted, as is done fairly frequently by the Ministry of Health and the Family Planning Commission, it would be possible to insert questions asking whether people have certain rare conditions or diseases. If the answer is affirmative, contact information would be solicited so that people with the disease or condition can be followed up with a specialized questionnaire. For example, the *National Survey of Family Growth (NSFG)* was based on a sample of reproductive age women previously interviewed in the *National Health Interview Survey (NHIS)* (National Center for Health Statistics 1987). Also, Harris et al. (1984) located victims of serious accidents by adding a screening question to a large number of omnibus surveys until a sufficient number was obtained. Of course, the adequacy of this strategy turns on the completeness of the original sample. If the initial sample is truly representative of the population, then individuals with the rare condition also will be representative of the population of such individuals. It also turns on the ability to elicit honest answers from potential respondents. Questions about whether individuals are HIV positive or have had homosexual sex or extramarital affairs simply inserted into questionnaires concerned mainly with other topics are not likely to elicit candid responses.

Other Methods for Sampling Hard-to-Research Populations

Several other methods for sampling hard-to-reach populations have been utilized by researchers. These include targeted sampling, time-and-place sampling, door-to-door screening, capture-recapture methods, and adaptive sampling. Space limitations preclude discussing them here; for a review, see the Chinese-language version of this article or the version available as a *Population Working Paper* (*PWP*), both mentioned in the acknowledgment footnote.

Fieldwork Procedures

In this section, we discuss new approaches to implementing sample surveys, including improved enumeration procedures, the use of incentives to increase response rates, methods for converting refusals, procedures for handling sensitive questions in the absence of privacy, methods for keeping track of respondents between waves in panel studies, and computer-assisted personal interviewing (CAPI) methods for recording responses.

Improved Enumeration Procedures

As noted above, in many nations, and increasingly in China, the conventional way to select households or individuals to be interviewed in sample surveys is first to select small areas, usually in a multistage process and usually with probability proportional to size. Then, within each small area, all the households or all the individuals are enumerated and a list is created from which households or individuals are selected at random to be interviewed. There are basically three ways to carry out the enumeration task (see also Treiman et al. 2006, 84–95).

The simplest and least expensive option is to simply move through the area, listing each address. This method works well in residential neighborhoods consisting entirely of single family houses, such as are common in the United States, although even in such neighborhoods it is likely that so-called mother-in-law apartments, occupied guest houses, etc., will be missed. It also works well in high-rise apartment areas, provided the enumerator is able either to gain entry to the building or to obtain information on the number of units in each building from the management company or from building guards. But, of course, what is generated is a list of households, which then requires a screening step with an informant in the selected households to determine which individual household members are to be interviewed.

A better, but infrequently used, approach is to develop a list of addresses together with information on the composition of each household—minimally, the number of eligible residents, for example, adults. This can be done by engaging as an informant a neighborhood or village committee official who accompanies the interviewer, helping to locate addresses and approaching householders when necessary to obtain information on household composition. The advantage of a local informant is that nonstandard dwelling places (sleeping spaces in shops, lofts over workshops, etc.) and doors-behind-doors (extra rooms in courtyards, etc.) are less likely to be missed. A second advantage is that a random sample of people, rather than households, can be drawn from the enumeration list, even though the exact individuals to be interviewed are not known and must be determined at the dwelling place by listing all the eligible household members and randomly selecting the required number using a Kish table or similar procedure. A disadvantage of this method is that often it is not possible to accurately determine the size or composition of households simply by using informants.

The gold standard is to conduct a minicensus, visiting each household, determining who lives there, and recording for each resident the age, sex, and, where necessary, other identifying characteristics. For example, if a household consists of the employees of a small workshop, there may be several individuals of the same age and sex, in which case they need to be distinguished, for example, the "tall girl from Sichuan." There are several advantages to the minicensus approach: (1) it makes it possible to draw a genuinely random sample of residents of the area without the clustering that is inherent in sampling first households and then individuals within households; (2) the sample can be drawn by the supervisor, making it less possible for interviewers to cheat by substituting immediately available household members for those requiring a return visit; and (3) it improves the likelihood of complete coverage because household members can act as informants not only about their own households but about the presence of subhouseholds (doors-behind-doors) that the interviewer may have missed. But this method is seldom used because it is very expensive.

No matter which option is used, a map of the area needs to be created with addresses listed on the enumeration sheet also identified on the map (Treiman et al. 2006). Without such a map, it often is extremely difficult for interviewers to find the households selected for the survey. This is particularly true in crowded neighborhoods in which apartment units are subdivided and extra rooms are built in courtyards, on rooftops, and in storage areas to rent out to migrants; such separate dwelling units are difficult to identify as such and often lack separate addresses. A useful supplement to a map, which is now technically feasible, but which could arouse suspicion in the Chinese context, is to identify each address by geo-coordinates and a photograph (although the accuracy of GPS receivers needs to be kept in mind).

The Use of Incentives to Improve Response Rates

In principle, the validity of inferences drawn from sample surveys depends on having high response rates (see three special issues of journals devoted to the problem of nonresponse to sample surveys: in 1999, vol. 15, no. 2, of the *Journal of Official Statistics,* edited by Edith de Leeuw; in 2006, vol. 169, no. 3, of the *Journal of the Royal Statistical Society*: *Series A* (*Statistics and Society*), edited by Peter Lynn; and in 2006, vol. 70, no. 5, of the *Public Opinion Quarterly*, edited by Eleanor Singer; see also Groves and Couper (1998)). When some individuals in a probability sample cannot be interviewed (because they cannot be contacted or refuse to participate), the danger arises that statistics computed on the portion of the sample for which data are available are biased. This occurs if those not interviewed differ systematically from those who are interviewed with respect to the variables under consideration. Although substantial nonresponse does not always result in substantial bias, it often does (Groves 2006). Thus, the best protection is to secure interviews from as high a fraction of targeted individuals as possible. Unfortunately, response rates have been declining in recent years, both in China and elsewhere, mainly because of an increase in refusals (de Leeuw and de Heer 2002).[7]

In China, a combination of increased personal freedom, possibly rising levels of mistrust, rapid urbanization (around the world response rates are lower in urban areas), and changes in the housing environment (an increasingly large fraction of the urban population lives in restricted-access buildings), has resulted in a sharp reduction in response rates. Indeed, the only Chinese surveys able to secure high response rates these days are "official" surveys that people feel uncomfortable refusing; in the United States as well, government-sponsored surveys tend to have higher response rates than academic surveys, which tend to have higher response rates than commercial surveys (Groves and Couper 1998).

Because response rates have been declining throughout the world, there has been a fair amount of research on the effect of incentives on response rates. Singer and her colleagues conducted a meta-analysis of 39 U.S. studies using face-to-face interviewing in which the effect of incentives was experimentally assessed (Singer et al. 1999; see also Singer 2002). Their main conclusions were that incentives improve response rates, especially in studies with low response rates (where presumably other motivations for participation are absent); the larger the incentive, the larger the increase in response rates; both monetary and gift incentives help, but monetary incentives are more effective even when the value of the different types of incentives is equated; incentives matter more when the survey is burdensome because the questionnaire is long; and prepaid incentives appear to be more effective than promised rewards for completed interviews, although the limited number of comparisons makes this result less than definitive.

The use of lotteries as incentives is favored by some investigators. However, the evidence on the efficacy of lotteries, all based on mail or self-administered surveys, is mixed (Singer 2002, 6). Moreover, there is some question as to whether lotteries would be effective in contemporary China because of the apparent low level of trust.

Methods for Converting Nonresponses

It long has been understood that repeated call-backs are effective in reducing noncontact nonresponses (Goyder 1985). Of course, remaining in an area for several days or even weeks in an attempt to eventually find someone home can add substantially to the cost of surveys, which is why survey houses, especially commercial survey houses, often are reluctant to do repeated call-backs.

Refusals are more difficult to convert than are noncontacts because the latter is mainly a matter of returning enough times to finally find someone at home. Three strategies are used to convert those who initially refuse: offering incentives, or increasing the value of the incentive if incentives were initially offered; elaborating on the value of the study, often by sending a different interviewer; and employing especially skilled interviewers. Sometimes, because of the costs involved, nonrespondent conversion is attempted only for a subsample of initial refusers, and the results are combined using the principles of *double sampling* (Hansen et al. 1993; Levy and Lemeshow 2008). Starting with its 2004 survey, the U.S.-based *General Social Survey* (*GSS*) has been using double-sampling methods to convert nonresponses, typically sampling half of the so-called temporary nonrespondents as of some cutoff date, usually 10 weeks after the start of interviewing (Davis et al. 2009, 2200). For another example of the use of double-sampling methods to improve coverage, see Jenkins et al. (2004).

An interesting example of the use of skilled interviewers to convert nonresponses is the second (2005) wave of the *MxFLS*, conducted three years after the first wave of data collection. By 2005, about 850 of the approximately 35,000 initial respondents had emigrated to the United States. When other family members were asked for the U.S. addresses of these migrants, many refused, probably because in many cases the migrants were in the United States illegally. However, when the study team sent a specially trained and highly skilled interviewer to revisit the households, he was able to obtain U.S. cell phone numbers and addresses for almost all of the emigrants; then another highly skilled female interviewer successfully contacted and interviewed a large fraction of this sample (Elizabeth Frankenberg, personal communication).

Handling Sensitive Questions in the Absence of Privacy

In many Chinese households, it is difficult to conduct interviews in private. Often space is limited, and even when it is not, it is common and culturally acceptable for family members and also neighbors to gather around to watch the entertainment afforded by the interview. In such circumstances, candid and accurate responses to sensitive questions, such as extramarital affairs, drug use, homosexual activity, or even whether people have had various diseases, whether they are happy, depressed, or have particular attitudes or opinions, are unlikely. People may refuse to answer the question or may give socially desirable responses. Because of this, procedures have been devised to elicit accurate responses in the absence of privacy and accurate estimates of extremely socially undesirable behaviors such as wife-beating, criminal activity, or sexual abuse. These are discussed in the extended Chinese-language and *PWP* versions of this article.

Panel Maintenance: Keeping Track of Respondents Between Waves

In panel surveys, where the same respondents are periodically reinterviewed (often approximately every three years, but sometimes much more frequently), it is crucial to minimize sample attrition because the ability to compare responses over time is limited to cases for which there are interviews at each time point in the comparison. Obviously, if a survey has an excellent response rate in the first wave of interviewing, say, 80 percent, but then loses half of the wave-1 respondents through failure to keep track of those who have moved or to retain the motivation to participate of those who have not, the effective response rate for wave-2 is 40 percent. For this reason, high-quality panel surveys devote considerable effort to *panel maintenance,* which includes periodically recontacting respondents between waves, often with a small gift and a brief phone call or letter reminding them of the importance of the survey, and tracing those who have moved between waves. In order to do this, it is necessary to collect extensive information from each respondent at wave-1, including names, addresses, land-line and cell phone numbers, and also, crucially, the names, addresses, and phone numbers of relatives, friends, and neighbors who are likely to know the whereabouts of respondents if they have moved. Doing this in such a way as to avoid refusals and the provision of deliberately incorrect information by respondents requires considerable interviewer skill. For a useful discussion of the panel maintenance procedures used in the *British Household Panel Survey,* see Laurie et al. (1999).

CAPI Procedures

With the increasingly low-cost availability of portable computing devices, it has become possible to carry out computer-assisted personal interviews in field settings. This is done by programming a portable computing device with the survey protocol. The advantages and disadvantages of this and allied procedures are discussed in the extended Chinese-language and *PWP* versions of this article.

Information Collected

There have been important new developments in the kind of data collected by social scientists concerned with the description and analysis of social structures and processes affecting general populations. Three areas have seen major expansion in recent years: the integration of data on individuals or families with data on characteristics of their social contexts (neighborhoods and communities but also classrooms, schools, hospitals, and other institutions); the integration of survey data with geographic information obtained from satellite-generated data and images; and the addition of biological data to population surveys. There also have been developments in data collection that result in improved measurements.

Contextual Variables and Multilevel Designs

It often has been hypothesized that characteristics of communities affect individual and household outcomes and behaviors (e.g., Brooks-Gunn et al. 1993; Sampson et al. 1999). One strategy for studying such possibilities using data from sample surveys is to construct summary indicators of community characteristics by averaging data from all the respondents residing in each community. However, this strategy is suboptimal both because not all community characteristics are aggregates of individual characteristics and, even for those that can be reasonably represented as aggregates, autocorrelation difficulties are likely to arise when both individual- and aggregate-level variables derived from the same information are included in the same model. For this reason, an increasing number of surveys of households and individuals also include a *community characteristics* component, often obtained by interviewing community officials and staff representatives from schools, health, and other community facilities; compiling records; and using the interviewer (or respondent) as an informant to report on the type and condition of housing and ease of access to facilities, for example, the location of the nearest school, hospital, library, or bank.

The usual goal is to gather information that captures community infrastructures and levels of socioeconomic development. The measures commonly collected include community educational, health, and sanitation facilities, means of transportation, availability of electricity, and characteristics of the local economy. Most often the focus is on community characteristics at the time of the survey, but sometimes historical information is collected as well. From an analytic standpoint, data about the nature of the community in the past are extremely useful since population researchers often want to know about the nature of the social environment at the time respondents were growing up. Thus, for example, in a study of adult outcomes, it would be much more useful to know whether a lower middle school existed in the respondent's community when the respondent was age 14 than whether a lower middle school existed in the community of residence at the time of the survey.

Good examples of sample surveys that systematically collect community characteristics via separate *community questionnaires* include *CHNS*, *IFLS*, and the *Living Standards Measurement Study,* a collection of national and subnational surveys conducted in over 30 developing nations. For good introductions to statistical procedures for simultaneously handling individual- and community-level data, see DiPrete and Forristal (1994) and Mason (2001).

Spatial Data

The increasing, and now widespread, availability of two types of geographic information has made it possible to incorporate spatial measurements of various kinds into population studies. The first are measurements of latitude and longitude via the GPS. The second are satellite-based environmental measurements and images of the earth's surface, which can be tagged with geo-coordinates.

Geo-Coordinates

The GPS is a satellite-based navigation system made up of a network of 24 satellites placed into orbit by the U.S. Department of Defense. The GPS originally was intended for military applications, but in the 1980s, the U.S. government made the system available without cost for civilian use. Each GPS satellite circles the earth twice a day in a very precise orbit and transmits signal information to earth. GPS receivers record the signals from three or more satellites and use triangulation, based on the elapsed time between transmission and receipt, to calculate the user's exact location. Relatively inexpensive GPS receivers (less than US$100) are accurate to within about 15 meters, and various enhancements can substantially increase accuracy. In addition, training of users is necessary to ensure accurate measurements for research purposes (Spencer et al. 2003).

Location and distance are key components of many processes social scientists study, for example, distance to market affects the choices by peasants of what crop to produce; the availability of local schooling is an important component of decisions about whether to keep children in school (Fafchamps and Wahba 2006); and the degree of isolation from transport hubs may affect decisions to migrate. In the past, researchers usually have relied on subjective reports by respondents about distances, but such reports are known to be highly inaccurate, with error often correlated with outcomes of interest (Gibson and McKenzie 2007). Using GPS technology, it is possible to precisely measure the location of each household and each other point of interest (e.g., schools, roads, markets, clinics or other public services, or nearest cities). For useful examples of GPS-based studies of distance as a barrier to services, see Entwisle et al. (1997) and Rosero-Bixby (2004).

In a different kind of application, Rosero-Bixby and Palloni (1998) studied the relationship between population growth and deforestation in Costa Rica, which has experienced an increase in population and a decline in forest cover over a 40-year period. Their paper addresses a central debate in research on population and environment—the extent to which rapid population growth is associated with the massive deforestation currently underway in many tropical nations. They compiled geo-coded data from two population censuses and a series of land cover maps and used these to model the net impact of population growth on the 1973–1983 probability of deforestation in about 31,000 750-square-meter parcels that were covered with forest at the beginning of the period. They showed a strong connection between population potential (the size of the population in surrounding areas) at the beginning of the period and subsequent deforestation.

Other Satellite Data

Satellite-based data, including both images and measurements of properties of the earth's surface, are now available without cost to researchers for essentially all locations in the world, albeit with higher resolution in some places than others. Social scientists have made a variety of uses of such data, often combining satellite measurements with data collected through population surveys. For example, Frankenberg et al. (2005) studied the impact of forest fires that blanketed the Indonesian islands of Kalimantan and Sumatra in 1997 on health outcomes by combining satellite measurements of aerosol levels in the atmosphere with longitudinal data for individuals from the *IFLS*. They showed that individuals who were exposed to haze from the fires experienced greater increases in difficulty with activities of daily living than did people living in nonhaze areas.

In another study, Frankenberg et al. (2008) studied the effect of the December 26, 2004, tsunami on the mental health of residents of the Indonesian island of

Sumatra, where 130,000 people died. Of interest here, they combined longitudinal sample survey data collected before and after the tsunami with community-level assessments of the extent of devastation caused by the tsunami. Devastation was measured by comparing satellite images from December 17 and 29 for each of their 585 study sites. For each site, images for a 0.6-kilometer square centered on the study site were classified into three categories based on the degree to which the pre-tsunami ground cover visible in the first image had been replaced by bare earth in the second image. About 15 percent of the areas were classified as heavily damaged (at least 20 percent of pre-tsunami ground cover replaced by bare earth); about 35 percent were classified as undamaged (no change in ground cover); and the remaining 45 percent were classified as moderately damaged. They then showed that mental health was strongly affected by the level of damage, net of other factors.

A different application of remote sensing technology is to use night-light measurements to study population density. Zhuo and his colleagues (2009) generated population density measurements for 1-kilometer squares covering all of China based on 1998 night-light data. Because these data are geo-coded, they can be integrated with ground-based data from sample surveys. See also the U.S. Department of Energy's global population database, LandScan, which provides annually updated population size estimates for 30-second by 30-second polygons (less than one kilometer square), using a regionally specific model based on provincial-level census data and four geo-spatial input data sets (land cover, roads, degree of slope, and extent of night light) (see Bhaduri et al. 2002; Dobson et al. 2000; and www.ornl.gov/sci/landscan/).

Finally, there has been increasing interest in combining microdata (from surveys and censuses) with remotely sensed data on land cover and land use to study the impact of population on the environment, particularly global warming, with attendant changes in weather patterns, increased vulnerability to natural disasters, and the loss of biodiversity. The collection of papers edited by Fox et al. (2003) provides a good introduction to this topic. See also Liverman et al. (1998), Matthews (2003), Walsh and Welsh (2003), and Walsh et al. (2004).

Biological Data

As social scientists have become increasingly interested in health, and health researchers have increasingly seen the advantage of population-based health studies, there has been a move to integrate questions about social and economic circumstances, health histories, and anthropometric and biometric measurements in general population surveys. One development, in particular, has been the collection of blood samples in order to do bioassays for the presence of various diseases and also of buccal cells to carry out genetic assays. However, the collection of venous blood in field conditions is difficult and expensive. This difficulty has led to the development of procedures for conducting bioassays and genetic assays from dried blood spots, which are much easier to collect and manage (McDade et al. 2007). The *IMHC* and the *China Health and Retirement Longitudinal Study* (*CHARLS*) surveys are examples of Chinese surveys that have collected dried blood spots successfully.

The basic procedure is that a respondent's finger is sterilized and then pricked using a spring-loaded device. The resulting blood droplets (usually four or five) are then collected on filter paper, dried, stored in plasticine bags, and within seven days shipped to the laboratory that will do the bioassays, where they are stored in medical freezers (−20°C) until ready for analysis. In China, by law the person doing the blood collection must be medically trained. Thus, a nurse, doctor, or other health worker must accompany the interviewer to do the blood collection. This, of course, adds considerably to the cost and complexity of the fieldwork procedures; but it has the advantage of improving other biometric and anthropometric measurements because nurses and doctors are used to touching and physically manipulating patients and thus are able to put respondents at ease when they are taking body measurements, such as hip and waist circumference, procedures that may be awkward for interviewers, especially when dealing with the opposite sex.

Apart from blood spots, many population-based studies now include measurements of blood pressure, heart rate, lung capacity, hip, waist, and upper arm circumference, height and weight, and brief physical ability tests (e.g., the timed chair stand: how long it takes for a person to stand up and sit down five times), a grip strength test, and so on. Sometimes other biological specimens, such as buccal cell samples for DNA analysis and urine samples to study sexually transmitted diseases are collected as well. Examples of studies that include biological measurements are the U.S.-based *National Longitudinal Study of Adolescent Health* (*Add Health*), *National Health and Nutrition Examination Survey*, and *Health and Retirement Study, IFLS, MxFLS, IMHC,* and *CHARLS.* All these studies are available for analysis by the research community.

Two particularly interesting examples of the use of biological data in social research are a field experiment on the efficacy of iron supplements in Indonesia (Thomas et al. 2006) and a study of gene-environment interactions in the United States (Guo et al. 2008). Iron-deficiency anemia is a serious problem in many developing nations, increasing the risk of reduced immune response, delayed cognitive and physical development, and fatigue and reduced work capacity. In conjunction with an on-going panel study of the Indonesian population (*IFLS*), Thomas and his colleagues carried out an experiment in Java involving approximately 17,000 participants in which weekly iron supplements were administered to all members of half of the participating households selected at random. Measurements of hemoglobin, a marker for anemia, were taken before the supplements began and again after the first six months of intervention. In addition, changes in employment status and income were assessed. The results were striking. Men who were iron deficient prior to the intervention and who were provided supplements substantially increased their hemoglobin levels, increased the probability of working, reduced their hours of sleep, lost less time to illness, became more able to conduct physically arduous activities, and improved their emotional health. Among the self-employed, hourly earnings, and therefore monthly earnings, increased. Benefits for women were in the same direction but the effects were more muted.

Guo and his colleagues carried out an analysis of data collected as part of *Add Health,* which extracted DNA from buccal cells for a subsample of the 20,000 adolescents in the initial sample. They used data from about 1,100 males for whom measures of both DNA and delinquency were available to study the interaction of genetic and environmental influences on delinquency. They were able to show that three genetic polymorphisms were significant predictors of serious and violent delinquency when added to a social-control model of delinquency. But two of them interacted strongly with family processes, school processes, and friendship networks. Specifically, the risk of delinquency only increased among those who both had the genetic propensity and suffered from weak social control. This is one of the first studies to show genetic-environment interactions in the kinds of behaviors ordinarily studied by social scientists.

Improved Measurement

Here we touch briefly on the value added by longitudinal data and discuss the advantages of multiple measurement of concepts. There also have been useful new developments in scaling and measurement procedures, but space limitations preclude discussing them here. For a discussion of two of these developments, the use of *anchor points* to calibrate self-reports of health status and of the method of *unfolding brackets* to obtain information on income and wealth, see the expanded Chinese-language and *PWP* versions of this article.

Longitudinal Data

We already have referred several times to longitudinal, or panel, data, in which the same individuals are queried at two or more points in time. Such data have many advantages, including in particular their ability to fix people's circumstances, conditions, behavior, and psychological states in time, and to purge from the analysis factors that are constant over time via fixed- or random-effects models, and hence to improve the possibility of being able rigorously to establish causal effects. But, as noted earlier, it is particularly important in panel studies to minimize panel attrition.

Even when it is not possible to mount panel studies, it sometimes is possible to generate over-time measurements by including retrospective questions in questionnaires, although in this case it is important to be sensitive to the possibility of recall error (Smith and Thomas 2003). One useful way to ask such questions is to collect *event history* rosters, for example, asking people to report each time they changed residence, each time they changed jobs, each time they married, and so on. For examples of the use of such rosters, see the questionnaires for the *Life Histories and Social Change in Contemporary China* (*LHSCCC*) and *IMHC* surveys.

Multiple Measurements

It is well known that responses to survey questions, even factual questions such as the level of education completed by the respondent's father, sometimes are erroneous. Even random error has serious consequences for statistical analysis, generally reducing the strength of the association between variables and, if the degree of error differs between variables in a model, distorting the relative strength of different effects. Systematic error, for example, a tendency to overstate the similarity between mothers and fathers or to *yea-say,* giving positive responses regardless of the content of the question, is even worse, seriously biasing statistical estimates. Methods for dealing with measurement error, particularly random error, are well summarized in Alwin's (2007) monograph. Most methods require either multiple measurements or multiple indicators.

Multiple measurements are independent measurements of a single underlying dimension either in the same survey—for example, by asking attitude questions using a 7-point scale near the beginning of a questionnaire and an 11-point *feeling thermometer* near the end of the questionnaire, or by asking about both the years of school completed and the highest level of schooling attained—or, if the phenomenon can be assumed to be constant over time, as father's education usually is, by asking the same question in two or more waves of a panel study. Ganzeboom et al. (1991, 292) discuss the advantages of multiple measurement, and Ganzeboom (2005, 2007) provides examples of the advantage gained by measuring occupations in two different ways. *Multiple indicators* are items that, at least in part, reflect some underlying dimension, for example, asking people whether in the past year they had experienced unfair treatment in a variety of circumstances—by the police, by shopkeepers, and so on—in order to construct a "harassment" scale; see the questionnaire for the *IMHC* study.

Multiple measurements make it possible to compute estimates of measurement reliability, which can then be used to correct multivariate models using errors-in-variables regression (Draper and Smith 1998, 89–91; Kmenta 1997, 352–357; Treiman 2009, 258–261). Multiple indicators also permit utilization of *multitrait-multimethod* models via confirmatory factor analysis procedures (Alwin 2007). Such models are a special case of structural equation models (SEMs), which permit the estimation of systems of equations, often involving unmeasured or latent constructs; for a good introduction to such models, see Bollen (2010).

For variables that must be coded, such as narrative reports on occupational position, an alternative to multiple measurement is multiple coding by independent coders—although building multiple measurements into the questionnaire is preferable. If, say, a set of narrative reports is independently assigned codes in the Chinese Standard Occupational Classification by two different coders, the codes can be converted into status scores (Ganzeboom and Treiman 1996; Ganzeboom et al. 1992) and reliability estimates can then be computed or the two versions can be used in the measurement model portion of an SEM. It is important to recognize that in order to fully exploit the power of either multiple measurement or multiple coding, it is necessary to apply these procedures to all the variables entering into one's model. For example, to study intergenerational occupational mobility, it doesn't help much to have multiple measurements of the respondent's occupation but not of the father's occupation.

Data Sharing

Several times we have mentioned data "in the public domain," that is, freely available for use by the research community. The sharing of data is highly desirable, both because no one researcher or research group will have the imagination or energy to fully exploit the large and very expensive data sets that constitute the database for much of contemporary social research and because independent validation and replication of results is the way science advances and error is corrected (Fienberg 1994). In the United States, the two leading funders of scientific research, the National Science Foundation and the National Institutes of Health, require that all sample surveys and other systematic data collection efforts paid for by them must be deposited in an archive or otherwise made available[8] within a reasonable period after the data are collected (the convention is two to three years, but many studies make their data available as soon as they are analysis-ready). The Organization for Economic Cooperation and Development (OECD), a consortium of about 30 democratic nations, mainly in Europe, has adopted a similar policy (OECD 2007).

Unfortunately, China is a laggard in this respect. To our knowledge, there currently are only three data archives in China: the Chinese Social Survey Open Database (CSSOD) at Renmin University; the China Survey Data Network (CSDN) at the China Center for Economic Research, Peking University; and the Databank for China Studies (DCS) at the Chinese University of Hong Kong. But the holdings of each of these archives are quite limited. Moreover, DCS charges a usage fee. Nor are there strong norms encouraging researchers to share their data. As a result, most of the studies of the Chinese population freely available for research use either were conducted by or in conjunction with foreign researchers or have been made available in foreign archives with English-language documentation, but are not readily available from Chinese sources or with Chinese-language documentation. It is our hope that as China increasingly participates in the international research community, this situation will change.

A happy exception is the *Chinese General Social Survey* (*CGSS*), a Chinese version of the *GSS*. This survey was initiated in 2003 by Yanjie Bian, at the Hong Kong University of Science and Technology, with an explicit commitment to make each wave of data publicly available. Although the 2003 survey surveyed only urban China, each subsequent survey has been a probability sample of essentially the entire population of China. So far, this survey, which from 2008 has been carried out by Renmin University, has been conducted six times, in 2003, 2004, 2005, 2006, 2008, and 2010, with another wave scheduled for 2012. The 2003, 2005, 2006, and 2008 data are publicly available with the 2010 data scheduled for release sometime in 2012.

Other major sample surveys of the Chinese population that are available to researchers are described in the Appendix. Apart from creating research opportunities for scholars lacking the formidable resources of time, energy, and funding required to conduct their own surveys, the availability of multiple national probability sample surveys of the Chinese population has several additional advantages. First, it permits replication of results obtained from a single survey. If two or more surveys based on independent samples yield the same conclusions, we can have far greater confidence in the validity of the results than when they derive from a single survey. Second, it often is possible to combine surveys conducted in different years to permit cross-temporal comparisons (see Treiman 2012 for an example combining the 1996 *LHSCCC* and 2008 *IMHC*). Third, surveys can be combined to increase the sample size and thus the power of the analysis, which is particularly important when relatively small subcategories of the population are studied (see Zhang and Treiman 2011 for an example combining the *IMHC* and the 2008 *CGSS*).

Notes

1. Each survey mentioned in the text is identified by its full name and abbreviation the first time it is mentioned and subsequently by its abbreviation. Publicly available Chinese surveys are described in the Appendix.

2. As one indicator, 22 percent of those living in China's Eastern and Central Regions report themselves to be in excellent health, compared to only 6 percent of those living in China's Western Region. Computations are from the survey of *Internal Migration and Health in China* (*IMHC*).

3. It is convenient to define a household or dwelling unit as "a living space with a doorway opening into a common or public area. Thus, for example, a room rented out to a migrant that opened into a family's living space would be regarded as part of the family household. A room opening into a courtyard or passageway would be regarded as a separate household." (Treiman et al. 2006, 89–91).

4. A second reason for clustering cases is to permit multilevel analysis (DiPrete and Forristal 1994; Mason 2001), studying whether some microprocess, say the relationship between mother's education and fertility, varies across communities (Entwisle and Mason 1985). A powerful analog to multilevel analysis for experimental designs is cluster randomized trials (Hayes and Moulton 2009).

5. The authors report (p. 1400) that they assume a design effect "(the ratio of the variance of a statistic from a complex sample to the variance of the same statistic from a simple random sample of the same size)" of 1.5, but this almost certainly is far smaller than if they had correctly calculated actual design effects (Treiman et al. 1996).

6. An alternative to weighting is to include all the variables used to construct weights in one's estimation equations and then to carry out the estimation without weights. The optimal way to carry out estimation for complex sample designs is an unsettled and rapidly evolving topic in statistical research. For useful recent discussions, see Little (2004); Gelman (2007)

plus five comments and Gelman's rejoinder in the same issue; and Faiella (2010).

7. Valid comparisons of response rates across surveys require a standardized way of measuring nonresponse. This has led to the development of a set of standards by the American Association for Public Opinion Research that has been widely adopted internationally (AAPOR 2011). Unfortunately, as two recent reviews of response rates make clear (Feng 2007; Hao 2007), the AAPOR standard often does not appear to be followed by Chinese investigators.

8. The Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan is the leading U.S. archive for large sample surveys, but there are also many other archives in the United States and elsewhere. For additional information on the holdings and terms of use of social science archives, see www.sociosite.net/databases.php and www.sscnet.ucla.edu/issr/da/Home.Other%20Archives.htm. In addition, major studies increasingly have established their own Web sites with provisions for downloading data; see the URLs listed in the Appendix. Most of the foreign surveys cited in this article have Web sites, which can be found by searching the Web on the survey titles.

References

AAPOR. 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys,* 7th edition. Lenexa, KS: American Association for Public Opinion Research.

Allison, Paul D. 2005. *Fixed Effects Regression Methods for Longitudinal Data Using SAS.* Cary, NC: SAS Institute.

Alwin, Duane F. 2007. *Margins of Error: A Study of Reliability in Survey Measurement.* Hoboken, NJ: Wiley-Interscience.

Bhaduri, Budhendra, Edward Bright, Phillip Coleman, and Jerome Dobson. 2002. "LandScan: Locating People Is What Matters." *Geoinformatics* 5: 34–37.

Bollen, Kenneth A. 2010. *Structural Equations with Latent Variables,* 2d edition. New York: Wiley.

Brooks-Gunn, Jeanne, Greg J. Duncan, Pamela Kato Klebanov, and Naomi Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99: 353–395.

Cumming, R.G. 1990. "Is Probability Sampling Always Better? A Comparison of Results from a Quota and a Probability Sample Survey." *Community Health Studies* 14: 132–137.

Davis, James Allan, Tom W. Smith, and Peter V. Marsden. 2009. *General Social Surveys, 1972—2008: Cumulative Codebook.* Chicago: National Opinion Research Center.

de Leeuw, Edith, and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." Pp. 41–54 in *Survey Nonresponse*, edited by Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J.A. Little. New York: Wiley.

DiPrete, Thomas A., and Jerry D. Forristal. 1994. "Multilevel Models: Methods and Substance." *Annual Review of Sociology* 20: 331–357.

Dobson, Jerome E., Edward A. Bright, Phillip R. Coleman, Richard C. Durfee, and Brian A. Worley. 2000. "LandScan: A Global Population Database for Estimating Populations at Risk." *Photogrammetric Engineering and Remote Sensing* 66: 849–857.

Draper, Norman R., and Harry Smith, Jr. 1998. *Applied Regression Analysis,* 3d edition. New York: Wiley.

Duncan, Greg J. 2008. "When to Promote, and When to Avoid, a Population Perspective." *Demography* 45: 763–784.

Eichen, Marc, and Ming Zhang. 1993. "Annex: The 1988 Household Sample Survey—Data Description and Availability." Pp. 331–346 in *The Distribution of Income in China*, edited by Keith Griffin and Renwei Zhao. New York: St. Martin's Press.
Entwisle, Barbara, and William M. Mason. 1985. "Multilevel Effects of Socioeconomic Development and Family Planning Programs on Children Ever Born." *American Journal of Sociology* 91: 616–649.
Entwisle, Barbara, Ronald R. Rindfuss, Stephen J. Walsh, Tom P. Evans, and Sara R. Curran. 1997. "Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice." *Demography* 34: 171–187.
Erickson, Bonnie H. 1979. "Some Problems of Inference from Chain Data." *Sociological Methodology* 10: 276–302.
Fafchamps, Marcel, and Jackline Wahba. 2006. "Child Labor, Urban Proximity, and Household Composition." *Journal of Development Economics* 79: 374–397.
Faiella, Ivan. January 2010. "The Use of Survey Weights in Regression Analysis." Bank of Italy Temi di Discussione Working Paper No. 739.
Feng, Xiaotian. 2007. "Is High Response Rate Better? Another Understanding on Response Rates for Social Survey." *Sociological Research* 3: 121–135 (in Chinese).
Fienberg, Stephen E. 1994. "Sharing Statistical Data in the Biomedical and Health Sciences: Ethical, Institutional, Legal, and Professional Dimensions." *Annual Review of Public Health* 15: 1–18.
Fox, Jefferson, Ronald R. Rindfuss, Stephen J. Walsh, and Vinod Mishra (eds.). 2003. *People and the Environment: Approaches for Linking Household and Community Surveys to Remote Sensing and GIS*. Norwell, MA: Kluwer Academic.
Frankenberg, Elizabeth, Douglas McKee, and Duncan Thomas. 2005. "Health Consequences of Forest Fires in Indonesia." *Demography* 42: 109–129.
Frankenberg, Elizabeth, Jed Friedman, Thomas Gillespie, Nicholas Ingwersen, Robert Pynoos, Iip Umar Rifai, Bondan Sikoki, Alan Steinberg, Cecep Sumantri, Wayan Suriastini, and Duncan Thomas. 2008. "Mental Health in Sumatra after the Tsunami." *American Journal of Public Health* 98: 1671–1677.
Ganzeboom, Harry B.G. 2005. "On The Cost of Being Crude: A Comparison of Detailed and Coarse Occupational Coding." Unpublished paper available from http://home.fsw.vu.nl/hbg.ganzeboom/pdf/index.htm.
———. 2007. "Multiple Measurement of Occupational Status." Paper presented at the ESS Quality Enhancement Workshop, Mannheim, Sept. 27–28. Available as PowerPoint presentation, www.europeansocialsurvey.org/index.php?option =com_docman&task=doc_view&gid=310&Itemid=80.
Ganzeboom, Harry B.G., and Donald J. Treiman. 1996. "Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations." *Social Science Research* 25: 201–239.
Ganzeboom, Harry B.G., Paul de Graaf, and Donald J. Treiman. 1992. "An International Scale of Occupational Status." *Social Science Research* 21: 1–56.
Ganzeboom, Harry B.G., Donald J. Treiman, and Wout Ultee. 1991. "Comparative Intergenerational Stratification Research: Three Generations and Beyond." *Annual Review of Sociology* 17:277–302. [Reprinted, 2002, in Chinese, in the *Tsinghua Sociological Review* 3: 151–194.]
Gelman, Andrew. 2007. "Struggles with Survey Weighting and Regression Modeling." *Statistical Science* 22: 153–164.
Gibson, John, and David McKenzie. 2007. "Using Global Positioning Systems in Household Surveys for Better Economics and Better Policy." *The World Bank Research Observer* 22: 217–241.
Gile, Krista J., and Mark S. Handcock. 2010. "Respondent-Driven Sampling: An Assessment of Current Methodology." *Sociological Methodology* 40: 285–327.

Goyder, John. 1985. "Face-to-Face Interviews and Mail Questionnaires: The Net Difference in Response Rate." *Public Opinion Quarterly* 49: 234–252.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70 (special issue): 646–675.

Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in Household Interview Surveys.* New York: Wiley.

Gu, Dongfeng, Kristi Reynolds, Xigui Wu, Jing Chen, Xiufang Duan, Robert F. Reynolds, Paul K. Whelton, and Jiang He. 2005. "Prevalence of the Metabolic Syndrome and Overweight Among Adults in China." *Lancet* 365: 1398–1405.

Guo, Guang, Michael E. Roettger, and Tianji Cai. 2008. "The Integration of Genetic Propensities into Social-Control Models of Delinquency and Violence Among Male Youths." *American Sociological Review* 73: 543–568.

Hansen, Morris H., William N. Hurwitz, and William G. Madow. 1993. *Sample Survey Methods and Theory, Volume I: Methods and Applications*. New York: Wiley.

Hao, Dahai. 2007. "The Significance of Response Rates and Others: Another Explanation to 'High' Response Rates for Chinese Social Survey." *Sociological Research* 6: 180–194.

Harris, Donald R., Mavis MacLean, Hazel Glenn, Sally Lloyd-Bostock, Paul Fenn, Peter Corfield, and Yvonne Brittan. 1984. *Compensation and Support for Illness and Injury.* Oxford: Clarendon.

Hayes, Richard J., and Lawrence H. Moulton. 2009. *Cluster Randomised Trials.* Boca Raton, FL: CRC Press.

Heckathorn, Douglas D. 2007. "Extensions of Respondent-Driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment." *Sociological Methodology* 37: 151–208.

Jenkins, Paul, Charles Scheim, Jen-Ting Wang, Roberta Reed, and Allan Green. 2004. "Assessment of Coverage Rates and Bias Using Double Sampling Methodology." *Journal of Clinical Epidemiology* 57: 123–130.

Johnston, L.G., M. Malekinejad, C. Kendall, I.M. Iuppa, and G.W. Rutherford. 2008. "Implementation Challenges to Using Respondent-Driven Sampling Methodology for HIV Biological and Behavioral Surveillance: Field Experiences in International Settings." *AIDS and Behavior* 12: 131–141.

Kalton, Graham. 1983. "Models in the Practice of Survey Sampling." *International Statistics Review* 51: 175–188.

Kmenta, Jan. 1997. *Elements of Econometrics,* 2d edition. Ann Arbor: University of Michigan Press.

Kumar, Naresh. 2007. "Spatial Sampling Design for a Demographic and Health Survey." *Population Research and Policy Review* 26: 581–599.

Landry, Pierre F., and Mingming Shen. 2005. "Reaching Migrants in Survey Research: The Use of the Global Positioning System to Reduce Coverage Bias in China." *Political Analysis* 13: 1–22.

Laurie, Heather, Rachel Smith, and Lynne Scott. 1999. "Strategies for Reducing Nonresponse in a Longitudinal Panel Survey." *Journal of Official Statistics* 15: 269–282.

Levy, Paul S., and Stanley Lemeshow. 2008. *Sampling of Populations: Methods and Applications,* 4th edition. Hoboken, NJ: Wiley.

Li, Shi, Chuliang Luo, Zhong Wei, and Ximing Yue. 2008. "Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Description." Pp. 337–353 in *Income Inequality and Public Policy in China*, edited by Björn Gustafsson, Shi Li, and Terry Sicular. Cambridge: Cambridge University Press.

Little, Roderick J. 2004. "To Model or Not to Model? Competing Modes of Inference for Finite Population Sampling." *Journal of the American Statistical Association* 99: 546–556.

Liverman, Diana., Emilio F. Moran, Ronald R. Rindfuss, and Paul C. Stern (eds.). 1998. *People and Pixels: Linking Remote Sensing and Social Science*. Washington, DC: National Academy Press.

Lu, Yao. 2007. "Educational Status of Temporary Migrant Children in China: Determinants and Regional Variations." *Asian and Pacific Migration Journal* 16: 29–55.

Malekinejad, M., L. Johnston, C. Kendall, L. Kerr, M. Rifkin, and G. Rutherford. 2008. "Using Respondent-Driven Sampling Methodology for HIV Biological and Behavioral Surveillance in International Settings: A Systematic Review." *AIDS and Behavior* 12: 105–130.

Mare, Robert D., and Vida Maralani. 2006. "The Intergenerational Effects of Changes in Women's Educational Attainments." *American Sociological Review* 71: 542–564.

Mason, William M. 2001. "Multilevel Methods of Statistical Analysis." Pp. 14,988–14,994 in *International Encyclopedia of the Social and Behavioral Sciences*, edited by Neil J. Smelser and Paul B. Baltes. Amsterdam: Elsevier Science.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J. Edward Taylor. 1998. *Worlds in Motion: Understanding International Migration at the End of the Millennium*. Oxford: Clarendon Press.

Matthews, Stephen A. 2003. "GIS and Spatial Demography." GIS Resource Document 03–63 (GIS_RD_03–63), October, Population Research Institute, Pennsylvania State University. Available at www.pop.psu.edu/gia-core/pdfs/gis_rd_03–63.pdf.

McDade, Thomas W., Sharon Williams, and J. Josh Snodgrass. 2007. "What a Drop Can Do: Dried Blood Spots as a Minimally Invasive Method for Integrating Biomarkers into Population-Based Research." *Demography* 44: 899–925.

Merli, M. Giovanna, William Whipple Neely, Feng Tian, Gao Ersheng, Tu Xiaowen, Gu Weiming, and Yang Yang. 2010. "Sampling Female Sex Workers in Shanghai Using Respondent Driven Sampling." Pp. 293–308 in *Rational Judgement. Public Health and Social Development*, edited by Guomei Xia and Xiushi Yang. Shanghai: Shanghai Academy of Sciences Publishing House (in Chinese).

Moser, Claus A., and Graham Kalton. 1971. *Survey Methods in Social Investigation*. London: Heinemann Educational Books.

National Center for Health Statistics. 1987. *Linking the National Survey of Family Growth with the National Health Interview Survey* (Vital and Health Statistics, Series 2, No. 103). Washington, DC: U.S. Government Printing Office.

Organization for Economic Cooperation and Development (OECD). 2007. *OECD Principles and Guidelines for Access to Research Data from Public Funding*. Paris: OECD.

Owen, L., A. McNeil, and C. Callum. 1998. "Trends in Smoking During Pregnancy in England, 1992–7: Quota Sampling Surveys." *British Medical Journal* 317: 728–730.

Roberts, Kenneth D. 1997. "China's Tidal Wave of Migrant Labor: What Can We Learn from Mexican Undocumented Migration to the United States?" *International Migration Review* 31: 249–293.

Rosero-Bixby, Luis. 2004. "Spatial Access to Health Care in Costa Rica and Its Equity: A GIS-Based Study." *Social Science and Medicine* 58: 1271–1284.

Rosero-Bixby, Luis, and Alberto Palloni. 1998. "Population and Deforestation in Costa Rica." *Population and Environment* 20: 149–185.

Sampson, Robert J., Jeffrey D. Morenoff, and Felton Earls. 1999. "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children." *American Sociological Review* 64: 633–660.

Singer, Eleanor. 2002. "The Use of Incentives to Reduce Nonresponse in Household Surveys." Pp. 163–178 in *Survey Nonresponse,* edited by Robert M. Groves, Don Dillman, John L. Eltinge, and Roderick J.A. Little. New York: Wiley.

Singer, Eleanor, Nancy Gebler, Trivellore Raghunathan, John Van Hoewyk, and Katherine McGonagle. 1999. "The Effect of Incentives in Interviewer-Mediated Surveys." *Journal of Official Statistics* 15: 217–230.

Smith, James P., and Duncan Thomas. 2003. "Remembrances of Things Past: Test-Retest Reliability of Retrospective Migration Histories." *Journal of the Royal Statistical Society (A)* 166: 23–49.

Song, Shige, Wei Wang, and Peifeng Hu. 2009. "Famine, Death, and Madness: Schizophrenia in Early Adulthood after Prenatal Exposure to the Chinese Great Leap Forward Famine." *Social Science and Medicine* 68: 1315–1321.

Spencer, John, Brian Frizzelle, Philip Page, and John Vogler. 2003. *Global Positioning System: A Field Guide for the Social Sciences*. Oxford: Blackwell.

Stark, Oded. 1992. *The Migration of Labour*. Cambridge, MA: Blackwell, Oxford.

St Clair, David, Mingqing Xu, Peng Wang, Yaqin Yu, Yourong Fang, Feng Zhang, Xiaoying Zheng, Niufan Gu, Guoyin Feng, Pak Sham, and Lin He. 2005. "Rates of Adult Schizophrenia Following Prenatal Exposure to the Chinese Famine of 1959–1961." *Journal of the American Medical Association* 294: 557–562.

Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Jaswadi, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, Teresa Seeman, James P. Smith, Cecep Sumantri, Wayan Suriastini, and Siswanto Wilopo. 2006. "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence." Population Working Paper PWP-CCPR-2006–070. California Center for Population Research, University of California, Los Angeles.

Treiman, Donald J. 2009. *Quantitative Data Analysis: Doing Social Research to Test Ideas.* San Francisco: Jossey-Bass. [A Chinese translation will be published by the Social Sciences Academic Press, Beijing, in 2012.]

———. 2012. "The 'Difference Between Heaven and Earth': Urban-Rural Disparities in Well-Being in China." *Research in Social Stratification and Mobility* [An earlier, expanded version is available as *Population Working Paper PWP-CCPR 2011–006*. Los Angeles: UCLA, California Center for Population Research. Available at http://papers.ccpr.ucla.edu.]

Treiman, Donald J., William Cumberland, Xilai Shi, Zhongdong Ma, and Shaoling Zhu. 1996. "A Sample Design for the Chinese Life History Survey." Appendix D.1.a of the documentation for the survey, "Life Histories and Social Change in Contemporary China." Downloadable from the UCLA Social Science Data Archive: www.sscnet.ucla.edu/issr/da/.

Treiman, Donald J., William M. Mason, Yao Lu, Yi Pan, Yaqiang Qi, and Shige Song. 2006. "Observations on the Design and Implementation of Sample Surveys in China." *Social Transformations in Chinese Societies* 1: 81–111.

Volz, E., and D.D. Heckathorn. 2008. "Probability Based Estimation Theory for Respondent Driven Sampling." *Journal of Official Statistics* 24: 79–97.

Walsh, Stephen J., and William F. Welsh. 2003. "Approaches for Linking People, Place, and Environment for Human Dimensions Research." *Geocarto International* 18(3): 51–61.

Walsh, Stephen J., Tom P. Evans, and Billie L. Turner II. 2004. "Population-Environment Interactions with an Emphasis on Land-Use/Land-Cover Dynamics and the Role of Technology." Pp. 491–519 in *Geography and Technology*, edited by Stanley D. Brunn, Susan L. Cutter, and J.W. Harrington, Jr. Norwell, MA: Kluwer Academic.

Whyte, Martin King, and William L. Parish. 1984. *Urban Life in Contemporary China*. Chicago: University of Chicago Press.

Wooldridge, Jeffrey M. 2006. *Introductory Econometrics: A Modern Approach,* 3d edition. Mason, OH: Thomson South-Western.

Wu, Xiaogang, and Donald J. Treiman. 2007. "Inequality and Equality Under Chinese Socialism: The *Hukou* System and Intergenerational Occupational Mobility." *American Journal of Sociology* 113: 415–445.

Zhang, Zhuoni, and Donald J. Treiman. 2011. “Social Origins, Hukou Conversion, and the Wellbeing of Urban Residents in Contemporary China.” Population Working Paper PWP-CCPR-2011–009. Los Angeles: UCLA, California Center for Population Research. Available at http://papers.ccpr.ucla.edu.

Zhao, Yaohui, John Strauss, Albert Park, and Yan Sun. 2009. *China Health and Retirement Longitudinal Study, Pilot, User’s Guide.* Beijing, Peking University, National School of Development.

Zhuo, L., T. Ichinose, J. Zheng, J. Chen, P.J. Shi, and X. Li. 2009. “Modelling the Population Density of China at the Pixel Level Based on DMSP/OLS Non-Radiance-Calibrated Night-Time Light Images.” *International Journal of Remote Sensing* 30: 1003–1018.

Zou, Hong, Zhiyong Qu, and Qiuling Zhang. 2005. “Survey of Development of Migrant Children in Nine Cities in China.” *Youth Studies* 2: 1–7 (in Chinese).

About the Authors

Donald J. Treiman is Distinguished Professor of Sociology Emeritus at the University of California at Los Angeles (UCLA) and is a Faculty Associate at UCLA’s California Center for Population Research. He has a B.A. from Reed College (1962) and an M.A. and Ph.D. from the University of Chicago (1967). Prof. Treiman started his career as a student of social stratification and status attainment, particularly from a cross-national perspective, and this has remained a continuing interest. For the past 17 years his research has focused on China. With colleagues, he has carried out two national probability sample surveys in China, a 1996 survey of about 6,000 adults focused on inequality over the life course and a 2008 survey of 3,000 adults focused on the determinants, dynamics, and consequences of internal migration. Currently, with Yao Lu and others, he is developing a new survey focused on the consequences of migration for children. Direct all correspondence to Donald J. Treiman, California Center for Population Research, University of California at Los Angeles (UCLA), 4284 Public Affairs Bldg., Box 957236, Los Angeles, CA, 90095; e-mail: treimandj@gmail.com.

Yao Lu is an Assistant Professor of Sociology at Columbia University. Her research areas include social demography and social stratification. She is particularly interested in examining causes and consequences of migration and immigration, and how migration intersects with socioeconomic and health inequality across diverse contexts. In collaboration with Professor Treiman, Lu is developing a national survey on internal migration and children in China.

Yaqiang Qi is assistant professor at Department of Sociology, Renmin University of China. His research interests include social stratification and mobility, sociology of health, and social demography. Currently he is working on a series of studies of the impact of societal inequality on population health.

Appendix

Chinese Sample Surveys of Interest to Social Scientists (listed in the order of the year of initial data collection)

Date	1982
Title	China Census of Population
Sample size	10,039,191
Sample specifications	One percent public use sample drawn from census records
Design	Single cross-section
Principal investigator	National Bureau of Statistics
Available from	IPUMS International (https://international.ipums.org/international/)
Language of documentation	English
Additional comments	Chinese language documentation is available from DCS

Date	1985
Title	China In-Depth Fertility Survey—Phase I
Sample size	13,307
Sample specifications	Married women <50-years old who are permanent household members in Hebei, Shaanxi, and Shanghai provinces
Design	Single cross-section; individual, household, and community questionnaires
Principal investigator	National Bureau of Statistics
Available from	DCS
Language of documentation	English

Date	1987
Title	China In-Depth Fertility Survey—Phase II
Sample size	49,458 households; 39,210 women
Sample specifications	Married women <50-years old who are permanent household members in Gansu, Guangdong, Guizhou, Liaoning, and Shandong provinces
Design	Single cross-section; individual, household, and community questionnaires
Principal investigator	National Bureau of Statistics
Available from	DCS
Language of documentation	English

(*continues*)

Appendix (*continued*)

Date	1988, 1995, 2002, and 2007
Title	*Chinese Household Income Project (CHIP)*
Note on series	There are four surveys in this series, based on different samples and with different principal investigators. In some sense, this is a set of repeated cross-sections, although the design and the variables have changed from survey to survey. Also, the specificity of documentation varies from survey to survey. Where available, details are given separately for each year.
Sample size	1988: 10,258 rural households (51,352 people), 9,009 urban households (31,827 people); 1995: 7,998 rural households (34,739 people), 6,931 urban households (21,694 people); 2002: 7,000 rural households (20,632 people), 9,200 urban households (37,969 people); 2007: unknown
Sample specifications	While it usually is asserted that data were collected from national probability samples of China, it is evident that not all the provinces were included and that included provinces were chosen purposively. For details see Eichen and Zhang (1993) and Li et al. (2008).
Design	Cross sections; separate samples of urban and rural populations; varying questionnaires across surveys: individuals, households, school age children; migrants; community characteristics
Principal investigators	1988: Keith Griffin and Renwei Zhao; 1995: Carl Riskin, Renwei Zhou, and Shi Li; 2002: Shi Li; 2007: unknown
Available from	ICPSR (www.icpsr.umich.edu/icpsrweb/ICPSR/), 1988: study no. 9836; 1995: study no. 3012; 2002: study no. 21741. The 2007 study has not been deposited at ICPSR and does not appear to be publicly available as of January 2012.
Language of documentation	English
Additional comments	These surveys were carried out by the Chinese National Bureau of Statistics Urban and Rural Survey Units. The 1988 sample is described in Eichen and Zhang (1993), and the 1995 and 2002 samples are described in Li et al. (2008). There currently is no documentation of the 2007 survey.

Date	1989, 1991, 1993, 1997, 2000, 2004, 2006, and 2009
Title	*China Health and Nutrition Survey (CHNS)*
Sample size	Approximately 4,400 households, yielding approximately 19,000 individuals
Sample specifications	Multistage semiprobability sample of nine provinces

Design	Panel (rotating panel); also community questionnaires
Principal investigators	Barry Popkin and Fengying Zhai
Available from	Carolina Population Center (www.cpc.unc.edu/projects/china)
Language of documentation	English and Chinese
Date	1990
Title	China Census of Population
Sample size	11,835,947
Sample specifications	One percent public-use sample drawn from census records
Design	Single cross-section
Principal investigator	National Bureau of Statistics
Available from	IPUMS International (https://international.ipums.org/international/)
Language of documentation	English
Date	1993
Title	China Housing Survey
Sample size	2,096
Sample specifications	Probability sample of households in Shanghai and Tianjin
Design	Single cross-section
Principal investigators	John R. Logan and Yanjie Bian
Available from	ICPSR (www.icpsr.umich.edu/icpsrweb/ICPSR/)
Language of documentation	English
Date	1994
Title	The State and Life Chances in Urban China 1949–1994
Sample size	4,074
Sample specification	National probability sample of people ages 25–65 residing in urban areas in China
Design	Single cross-section
Principal investigators	Xueguang Zhou and Phyllis Moen
Available from	ICPSR (www.icpsr.umich.edu/icpsrweb/ICPSR/)
Language of documentation	English
Date	1996

(*continues*)

Appendix (*continued*)

Title	*Life Histories and Social Change in Contemporary China* (*LHSCCC*)
Sample size	6,090
Sample specifications	National probability sample of people ages 20–69
Design	Single cross-section
Principal investigators	Donald J. Treiman, Andrew Walder, and Qiang Li
Available from	UCLA SSDA (www.sscnet.ucla.edu/issr/da/)
Language of documentation	English and Chinese
Additional comments	Also available from CSSOD. The documentation at CSSOD is not complete but is in Chinese.

Date	1998, 2000, 2002, 2005, and 2008
Title	*Chinese Longitudinal Healthy Longevity Survey* (*CLHLS*)
Sample size	Initial wave: 9,073 people age 80+; see project Web site for other sample sizes
Sample specifications	People age 80 and older in 1998; also, a comparison group ages 65–79 and other special samples
Design	Panel, with replacements for deceased
Principal investigator	Yi Zeng
Available from	Center for the Study of Aging and Human Development, Duke University (http://centerforaging.duke.edu/index.php?option=com_content&view=article&id=115&Itemid=152)
Language of documentation	English

Date	1999–2000
Title	*Chinese Health and Family Life Survey* (*CHFLS*)
Sample size	3.426
Sample specifications	National probability sample of people ages 20–64, with an oversample of high sexually transmitted disease areas
Design	Single cross-section
Principal investigators	William Parish and Edward O. Laumann
Available from	www.src.uchicago.edu/prc/chfls.php
Language of documentation	English and Chinese

Date	2003, 2004, 2005, 2006, 2008, and 2010
Title	*Chinese General Social Survey* (*CGSS*)
Sample size	2003: 5,894 (urban only); remaining years, approximately 10,000 each

Sample specifications	National probability sample of people ages 18–69
Design	Repeated cross-sections
Principal investigators	Initially, Yanjie Bian; since 2008, Lulu Li
Available from	CSSOD
Language of documentation	Chinese and English
Additional comments	Currently only the 2003, 2005, 2006, and 2008 data are available for public distribution, with the 2010 data scheduled for release sometime in 2012.

Date	2008
Title	*Internal Migration and Health in China* (*IMHC*)
Sample size	3,000
	National probability sample of people ages 18–64, with an oversample of high in-migration areas
Design	Single cross-section
Principal investigators	Donald J. Treiman, William M. Mason, and Shige Song
Available from	UCLA SSDA (www.sscnet.ucla.edu/issr/da/)
Language of documentation	English and Chinese; currently, limited Chinese documentation

Date	2010–2020
Title	*Chinese Family Panel Studies* (*CFPS*)
Sample size	14,798 households
Sample specifications	Nearly national probability sample (excluding Tibet, Qinghai, Xinjiang, Ningxia, Inner Mongolia, and Hainan)
Design	Panel, with a new wave of household interviews every year and a new wave of adult interviews every other year
Principal investigator	Yu Xie
Available from	Institute of Social Science Survey, Peking University (www.isss.edu.cn/)
Language of documentation	Chinese and English
Additional comments	A pilot study (n = 2,375 households) was conducted in 2008 in Beijing, Guangdong, and Shanghai, with a second wave of interviews conducted in 2009. The main survey was first completed in 2010.

Date	2011
Title	*China Health and Retirement Longitudinal Study* (*CHARLS*)
Sample size	Approximately 10,000 households, 17,000 individuals

(*continues*)

Appendix (*continued*)

Sample specifications	National probability sample of adults age 45 and older; see Zhao et al. (2009) for details
Design	Panel, with a new wave of interviews every two years
Principal investigators	Yaohui Zhao, Justin Yifu Lin, John Strauss, and Albert Park
Available from	National School of Development (China Center for Economic Research) at Peking University (http://charls.ccer.edu.cn/charls/)
Language of documentation	English and Chinese
Additional comments	A pilot study (n = 1,570 households, 2,685 individuals) was conducted in 2008 in Zhejiang and Gansu provinces. A national sample survey was scheduled for 2011, but no documentation is currently available

Note: The surveys listed here are available, generally without charge, to members of the academic research community. We list the major surveys known to us. Additional surveys, often restricted to specific regions, cities, or subgroups of the population, may be obtained from the Chinese Social Survey Open Database (CSSOD) at Renmin University (www.cssod.org/index.php); the China Survey Data Network (CSDN) at the China Center for Economic Research, Peking University (www.chinasurveycenter.org/CSDN_EN/); and the Databank for China Studies (DCS) at the Chinese University of Hong Kong (www.usc.cuhk.edu.hk/Eng/AboutDCS.aspx). Note that the last of these archives charges user fees. The China Data Center at the University of Michigan (http://chinadatacenter.org) makes available a variety of aggregate data files constructed from various Chinese censuses. However, this center also charges fees, which often are beyond the capabilities of academic researchers. Finally, the NBS maintains a Web site, the NBS Statistical Data Link (www.stats.gov.cn/tjsj/), which shows various aggregate statistics drawn from the censuses and surveys conducted by NBS.