Studies in Sociology of Science **ISSN 1923-0176** [PRINT]
Vol. 2, No. 1, 2011, pp.14-21 **ISSN 1923-0184** [ONLINE]
www.cscanada.org **www.cscanada.net**

Error Analysis of Sampling Frame in Sample Survey*

LI Zhengdong[1]

Abstract: In our application practice of sample survey, we mostly neglect some non-sampling errors such as sampling frame errors. Actually, the influence of non-sampling errors to the total survey deviation can not be ignored. In view of this topic, this paper briefly discussed the sampling frame errors as non-sampling errors. First a brief review of the sampling frame, together with the type and structure of the sampling frame, is given. Next the distinction between sampling frame errors and sampling errors is made theoretically in general. Then through the analysis of a series of non-random impact factors and the application of corresponding improvements or solutions, the sampling frame errors are reduced or controlled within a certain range. Finally, this paper summed up and sorted out the influencing factors based on the sample units or elements for the sampling frame, and also discussed the problems and solutions.
Key words: Sampling Survey; Sampling Frame; Sampling Error; Sampling Frame Error

Sample survey is the process of observing, describing, and analyzing a sample of units drawn randomly from the survey population to draw an inference. In sample surveys, the sampling frame is a major component in the entire sampling design, and it plays an important role in sample surveys[2]. Ideal sampling frame means that every element in the population is recorded and only recorded once, excluding other elements which are not in the population. In practice, this ideal condition is very rare. As for probability sampling, essentially it is to select a subset of some units from a finite set of all units, with the probability of a subset being selected known. Sampling frame is a unit of finite population, however,

* This work is supported by Innovation Program of Shanghai Municipal Education Commission (Project Number: 09-YS399). Moreover, this work is also supported by Social Science Research Funds of Shanghai Institute of Technology (Project Number: SJ 2009-06). At the same time, this work was funded by the tenth key construction course of Shanghai Institute of Technology "Social Research Methods". Gratitude is expressed here!
[1] Department of Social Work, Humanity College, Shanghai Institute of Technology, Shanghai, China. Email: lizhengdong@sit.edu.cn
[2] Due to historical reasons, the domestic sociology society often calls "sample survey" as "social investigation." Because the three of sampling, questionnaires, statistical analysis surveys constitute the most critical aspects and the most essential characteristics for investigation and research, and thus investigation and research are called as the "sample survey", "questionnaire survey" or “statistical survey." (Feng Xiaotian. 2003)
* Received March 29, 2011; accepted April 19, 2011

the population of sampling frame units do not necessarily equivalent to the population of data collection[3] (Li Jinchang, 1996:30-32). As a way to identify the elements in the population, the sampling frame usually includes a large number of additional information useful for sampling design in its composition, which often has a strong influence on sampling design. The structure of sampling frame, together with its additional information and the quality of this information will determine the type of survey sampling design and the estimation procedure and its deviation.

1. SAMPLING FRAME AND ITS TYPE AND STRUCTURE

Simply put, sampling frame is the sampling range or the list of sampling units, and it is the list of all sampling units in the survey population. For the sampling frame, different researchers have different cognitive styles. In order to facilitate the implementation of probability sampling, sometimes the population will be divided into several parts that are non-overlapping and exhaustive, each part is called sampling unit, sampling unit is not necessarily the individual. In this sense, the sampling frame is actually the roll or list of all the sampling units. In preparing the sampling frame, if the sampling unit divided too finely, a lot of manpower resources and financial resources will be spent. For example, it is difficult to prepare the sampling frame of village committees and neighborhood committees in a province. At this time sampling units can be graded. For instance, to obtain the sampling frame of all village committees and neighborhood committees in a province, county can serve as the primary sampling units, so that it is easy to obtain the sampling frame of all county-level units in a province. Similarly the sampling frame of all townships in a county and the sampling frame of all villages in a township can be obtained easily.

American scholar Judith and William defined the sampling frame systematically based on the past concept. In their view, the sampling frame consists of the materials, procedures and methods to identify, differentiate, and approach the elements of a target population. Sampling frame is composed of a set of finite units, from which probability sampling can be conducted. The mechanisms or rules linking the sampling frame units and the population elements are an integral part of the sampling frame. The sampling frame also includes auxiliary information (such as size measurement, demographic information), which is used for special sampling techniques (such as stratification sampling and probability proportional to size sampling) and special estimation techniques (such as ratio estimate or regression estimate). the concepts involved in the sampling frame are: (1) The target population is a finite population with its elements identifiable; (2) The sampling is a set of selected units, but this set may not belong to the target population; (3) The linkage mechanism must exist between the target population and the set of selected units; (4) In order to obtain information from the elements, they must be able to distinguish and find; (5) There may be a variety of linkage forms between target elements and sample sets. The linkage procedure determines the type of sampling design and the estimation procedure in a survey; (6) For some sampling designs and estimation procedure, the auxiliary information about the population elements is required. The auxiliary information of every element in the target population must be known (Judith and William. 1997: 47-48). In fact, before we select the sample, the population must be divided into several parts, and these parts are sampling units. The list structure of the sampling units is the sampling frame in our practice and study.

The sampling frame includes the list sampling frame, mixed sampling frame and identification sampling frame.[4] In the sample survey, an ideal type of sampling frame is able to clearly distinguish and identify sampling units, and at the same time includes the clear linkage rules between each population element and each sampling unit. That is to say, according to the specific requirements of the sample survey, the sampling frame should not only describe the individual units, but also provide some additional information, such as the size of the components and the information of which part of the target

[3] Researchers usually want to take the survey population as the "target population", while inference population is the population deduced from statistics. There is only one target population, while there may be a few inference populations.

[4] It should be noted that the sampling frame types we mentioned here can manifest the simple one-to-one connection between the sampling units (units in sampling frame) and the target units (units in population).

population that they belong to.

A list of sampling frames is the list of a set of population elements numbered, and each element in the sampling frame is numbered in the serial order from 1 to N. For a list sampling frame, its target elements are all in the frame and the overall number of elements is known, and thus the random number table can be used to draw random samples easily. This kind of sampling frame is usually considered as the ideal type, which is relatively rare in our study and practice.

A mixed sampling frame is made up of a group of tangible sampling units, each of which is linked with a target element. These sampling units are mixed up to simulate the randomization state of location or sequence. It is a good example to select or draw lots using the pieces of paper with numbers written on them, and each piece of paper represents each element in the target population.

An identification sampling frame is composed of a group of tangible units in a certain order but unnumbered, and it designates a specific unit by a recognizable way. For example, specially numbered units are obtained by counting. In fact, as for the sampling frame, the correctness of the sampling process can not be verified correct from samples.

The type of sampling frame, mentioned above, is limited to one-to-one linkage between the sampling frame unit and the target element. It is also according to the linkage between the sampling frame unit and the target element that Judith and William summarize the following four structures of the sampling frame: (1) One-to-one structure. Each sampling frame unit corresponds to a unique population element, and each population element also corresponds to a unique sampling frame unit. (2) One-to-many structure. Each sampling frame unit may correspond to more than one population elements, while each population element only corresponds to a unique sampling frame unit. (3) Many-to-one structure. Each sampling frame unit corresponds to a unique general element, while each population element may correspond to multiple sampling frame units. (4) Many-to-many structure. Each sampling frame unit may be linked with multiple population elements, while each population element may be linked with multiple sampling frame units (Judith and William. 1997: 47-48).

In one-to-one and one-to-many links, the probability of the sample units being chosen can be calculated easily. In a general sense, the probability of a target element being chosen into samples is equal to the probability of its corresponding sampling frame being chosen inot samples. In many-to-one and many-to-many structures, in order to obtain unbiased estimators, it is necessary to know the quantity of the sampling frame units that are associated with the target elements (ZHU Tie and Zhu Zhenbao, 2001). It should be noted that when a survey is conducted, several sampling frame structures, discussed above, may exist at the same time. For example, the multi-stage sampling requirements a sampling frame at each stage of sampling. We take a nationwide household survey of urban poverty as an example. Roughly three-stage sampling will be adopted, namely, cities, city communities, and poor families in the community. At the first stage, the sampling frame is a list of cities, and this sampling frame belongs to a one-to-many structure, because each unit of the sampling frame is associated with multiple families in the target population (potential poor families). At the second stage, the sampling frame is a list of communities in the chosen city, and the sampling frame still belongs to a one-to-many structure. At the final stage of sampling, the survey or interview staff will visit the families in the communities and list all poor families, and then the units on the list should be associated with the population elements in a one-to-one structure.

In addition, regional sampling frames and schedule sampling frames are also commonly used. Regional sampling frame is to list all population units based on the division of the geographic regions, while the schedule sampling frame to list all population units in a chronological order.

2. DISTINCTION BETWEEN SAMPLING ERRORS AND SAMPLING FRAME ERRORS

Sample survey is to investigate a part of representatives selected from the population of study object, and then use the results to infer and indicate the population characteristics. The process of select a part of

representatives from the population is sampling. Sample survey follows the principle of random sampling, and thus can rule out the influence of subjective factors to ensure the representativeness of the sample. Of course, there are some differences between the results from sample survey and from census, and these differences are what we called sampling errors. In fact, we can control sampling errors in a certain range through scientific means. In other words, there are errors to use the sample statistics to estimate or infer the population parameters, which means that the results from estimation or inference may be inconsistent with the true value of the population parameters.

In general, there are two categories of the sampling error sources in sample survey: (1) Registration error. It is also known as survey error, it is caused owing to the errors from observation, measurement, registration, and calculation during the investigating, editing, encoding and aggregating process or because of the errors resulted from false information. The direct manifestation of this kind of errors is that the sign value or mark feature of the investigation unit is not collected or recorded truly and objectively, so that the statistics calculated deviate from its true value. Registration errors are not specific to sample surveys, but are likely to exist in all of the statistical surveys. Moreover, the greater the scope of investigation and the more the number of units, the greater possibility of registration errors resulted. This kind of errors is closely related to the accuracy of measurement tools, measurement techniques, the sense of responsibility of investigators, cooperative attitude of respondents; (2) Representation error. There are two types of such errors: systematic representation errors and accidental representation errors. Systematic representation error is called systematic errors for short. This kind of errors is caused by lack of sample representativeness owing to non-random factors, and its manifestation is the value of the sample statistics is high or low systemically. For example, better or worse units are chosen deliberately to be investigated; questionnaires are poorly designed (for example, some of the problems are easily misunderstood); the sampling frame and the target population are inconsistent; the sample is limited to the part easily to access (such as selecting households from telephone directory); fixed certain unit time is chosen to investigate the population with periodic changes (such as conducting a market price survey every Sunday); there are no answers or incomplete answers for some of the units in the sample. All of these would cause systemic representation errors. This kind of errors also belongs to issues of ideology, style and techniques, which should be avoided in a sample survey. Systematic errors, together with registration errors, are called deviation. Accidental representation errors are called accidental errors for short. This kind of errors is caused by the reason that the sample structure can not completely represent the population structure due to accidental factors following the principles of random sampling. It is inevitable in sample survey. Even if there are no registration errors or systematic representation errors, the errors between the sample and the population still exist. Although accidental errors are inevitable, but can be estimated and controlled, and it will increase or decrease with the number of investigation units. When the survey is comprehensive, sampling errors would not exist. Accidental errors can be positive or negative, but for all possible samples, the sum of accidental errors is equal to zero (L • Kish, 1997).

In practice, the sampling error has following forms (Li Jinchang.1996:32-45): (1) Actual sampling error. It refers to the absolute deviation between the sample statistics and the population parameters. (2) Standard sampling error. It is also known as the average sampling error. It is the indicator of the general level reflecting the size of the sampling error. The sampling error usually talked about, refers to the standard sampling error. We can take the standard deviation of the sample statistics as an indicator to measure the size of sampling errors, that is to say, the standard deviation of the sample statistics is used to describe the general level of sampling errors. The size of standard sampling errors is affected by many factors, including degree of differences within the population, sample capacity, sampling method, organization forms of sampling, estimation method of population parameters.[5] (3) Limit sampling error. It refers to the maximum allowed error range when the population parameter is estimated using sample statistics, and it is also the absolute value of the difference between the maximum or minimum allowed value of sample statistics and the population parameter in a sample inference.

During the sample design and sample selection, it is critical to determine the source of material sampled. This source of material is what we call the sampling frame, and it is some kind of list in general,

[5] The so-called sampling organization form refers to the specific sampling process, namely, how to arrange extraction of the samples. Currently the sampling organization form often used has several types such as simple random sampling, stratified sampling, systematic sampling, group sampling, and multi-stage sampling.

such as a list of all the households or a roster of all residents in a neighborhood committee. If there are no suitable lists about the population, the sampling frame may be composed of the geographical regions with clearly defined natural or man-made boundaries. Then we select a region (lot) as the sample, and let an interviewer to investigate the sampling "lot" in details, register appropriate units such as households or others to let some of these units elected to the final samples.

It is meaningless to discuss the sampling frame itself. In the discussion of the ideal sampling design, we need to consider the relative cost to design and construct different sampling frames and the influencing factors to the total survey error (Judith and William. 1997:53). The type of survey sampling frame and any of its defects may affect the total survey error. Similar as other factors in survey design, the choice of sampling frame must be based on the cost of evaluating, constructing and using the sampling frame and the total error of the survey estimation value using the specific sampling frame. To carry out the probability proportional to size sampling, more cost is needed to construct the sampling frame containing the corresponding information. However, the gain of sample benefit may balance the additional cost used to construct the sampling frame. It needs to be considered whether the cost and time spent are necessary in selecting the type of sampling frame. If the money is spent on a larger sample in the survey, better results may be obtained.

When sample survey is conducted in practice, the errors are often caused by "non-random factors ", and affect the accuracy of the sampling results greatly. This kind of errors is generally defined as "non-sampling errors." Only by carefully designed sampling process and seriously implement of the sampling process, can such errors be minimized. When the impact of the sampling frame to the survey is examined, we need to pay attention to not only the relative efficiency of the sampling design that the sampling frame allows, but also the defects and mistakes of the sampling frame itself, which would affect the variance of estimators. For example, the sampling frame may contain auxiliary information, which can be used for stratification, so that the optimal allocation of stratification can be adopted. If the information for stratification is not accurate, then the variance of the estimator will be larger than that from accurate stratification information. The relative efficiency of one type of the sampling frame contrasted with another type of the sampling frame is closely related to the efficiency of different sampling designs in each case. Although the available sampling design is limited by the existing sampling frame types, for the same type of the sampling frame, there still may be different sampling designs. Therefore, when considering the relative efficiency of the sampling frame, we have to consider the sampling designs. In addition, if the sampling frame is not of high quality, the bias of the estimator may be resulted. The quality of the sampling frame depends on whether it is outdated or complete and so on.

In a broad sense, the error brought about by the influence of a sampling frame to a sampling survey is the systematic representation error or systematic error. It is caused by lack of representiveness and so on due to non-random factors. And its manifestation is that the value of the sample statistics is systemically higher or lower. In our application in practice, it is often regarded as the deviation and neglected. In fact, we should treat sampling frame errors as non-sampling errors. Apparently this is a narrowly understood. However, it allows us to reduce or control the errors in a certain range through a series of analysis of the non-random influencing factors and the application of corresponding improvements or solving measures in the sampling design.

3. INFLUENCING FACTORS AND REDUCTION OF SAMPLING FRAME ERRORS

Ideal sampling frame means that every element in the population is recorded and only recorded once, excluding other elements which are not in the population. In practice, this ideal type is very rare. The survey and sampling personnel must be careful about the imperfection of the sampling frame. As mentioned above, the sampling frame is a way to identify population elements, and its composition usually includes a large number of auxiliary information useful for the sampling design and often has a strong influence on the sample design. The sampling frame structure, the auxiliary information and the quality of the information would determine the type of survey sampling design, the estimation procedure

and its deviation. Based on what we learn above, we summed up and sorted out the influencing factors based on the sample unit or element for the sampling frame, and also discussed the problems and solutions[6].

1) Omitting Sample Units or Missing Elements

Some of sample units are often omitted in the process of preparing the sampling frame, so that these samples can not be surveyed and their situation can not be reflected in the survey. Therefore, the sample can not represent the entire target population. Omission of sample units is the most serious problem of the sampling frame, because the omitted sample units can not be found in the sampling frame. Because the missing sample units can not be found and corrected, the estimation bias may be caused in the survey. In this case, the estimated total is always low. Omission of sample units is probably because the sampling frame is "inadequate", or because the sampling frame is "incomplete."[7] If for the list of a school, a part of part-time and full-time students is deliberately excluded from its target population, it is inadequate; and if the list is outdated, and therefore failed to include some new students, it is incomplete.

Unless we can find a way to make up, otherwise there is no chance that these sampling units can be selected as samples. Sometimes, this problem can be avoided through defining the survey population to rule out the omission of sample units. This imperfect solution is often adopted. A more desirable solution is to find supplement sampling frames so that the missing sample units can be included, such as the list of special students and newly enrolled students. For the missing sample units, we often can not find applicable supplementary sampling frames. Under these circumstances, we may find some kind of connection procedures to deal with. The purpose of the connection procedure is a clearly defined way to add omitting sample units to a specific record (Carlton, 2003:61).

2) Inclusion of the Non-target Population Sample Units or Vacancies, Alien Elements

Sometimes the sampling frame includes some sample units that does not belong to the target population, such as the elements that no longer exists but still in the sampling frame or the elements although in sampling frame but not within the survey scope. The main consequence that vacant elements bring about on the sampling frame is that the sample size will be smaller than the selected number due to some chosen vacancies being discarded. If this mistake is not corrected, there will be possibility that the bias of statistics will naturally be resulted. The errors caused by the sampling frame do not influence greatly, because those non-target elements can be recognized and removed in the survey, and unbiased estimation of the target population can be obtained by using the theory of domain estimation. Here the target population is a subset of the sampled population. If the non-target elements can be found in a survey, when compared to perfect sampling frames, the main influence of the errors for this type of sampling frames including the non-target elements is reduction of the estimation efficiency (Judith and William. 1997:55).

On some occasions, apparently the list or units in the sampling frame have no relationship with the members of the target population, for example, such as those who are dead or have moved out, those places that are empty or have been demolished, in this case, the sampling units are not related with any target elements, which are called "vacancy elements" or "empty list." On other occasions, some units are connected to the sampling frame, and they are similar to the target elements, but do not belong to the target population. For example, in a household survey, the purpose of the survey is to study the residents with Chinese nationality. However, at some places in the sampling frame, reside foreigners, and they are not members of the target population, which are referred to as "alien elements" that are connected with the sampling frame but do not belong to the target population. Judith and William argued the specific

[6] In the sample survey, the sampling units are individual units on the list off the sampling; the element is the smallest unit in the survey, usually refers to a person; sample refers to the sum of the sampling units taken from the sampling frame.

[7] It is very important to make a distinction between "inadequate" and "incomplete" for the sampling frame in practice, because “inadequate” is usually easier to identify. The sampling frame is inadequate means that it does not intend to include the entire target population, while the sampling frame is incomplete means that it fails to include some elements of the target population (Carlton, 2003:60).

reasons of including the sample units of non-target population are the "boundary problem." According to the rules, the population elements that are located in the region near the boundary do not belong to the sample region, but are included in the sample region, and thus the "boundary problem" arises.

The method dealing with vacancies are relatively simple. In fact, it is only option to remove or discard the selected non-population sample units or vacancies, alien elements. In practice, our best hope is to recognize vacancies and alien elements from the sampling frame, because we can remove non-population sample units before the survey is started. In fact, what we met most is that those units can not be recognized and can only be removed after contacting with them. In determining the sampling ratio required to achieve the desired sample size, Carlton pointed out that the common mistake of using systematic sampling is that when a vacancy is hit, then it will replaced by the next element on the list. This situation should be avoided because it increases the probability of the next element being selected. From beginning to end we should repeat the sampling interval in the whole population, discard the vacancies from samples (Carlton. 2003:65).

3) Multiple Lists and Multiple Sample units or Multiple Elements

Multiple units or multiple elements of the sample is the problem caused by representing two or more units or elements of the sample by one list or sample unit in the preparation process of the sampling frame. Once this situation happens, the probability of these sample units being selected is different from that of other sample units in the actual sample survey, which violate the random principle, and will inevitably affect the total sampling error of the sample survey. For example, in the survey of urban and rural residents' child-bearing will, it is feasible to take the doorplate numbers as sampling units in rural regions where one doorplate number corresponds to one family, while in the community committees of the city, one doorplate number often corresponds to more than 10 families. For the error reduction of this type of sampling frames, we must strictly review our sampling units, and change the sampling units when the case of multi-element sampling is encountered.

Multi-List means that some elements or units of the sample may appear on multiple lists when the sampling frame is composed by multiple lists. In other words, the target population elements may be associated with multiple sampling frame units. The problem caused by multiple-list is that the probability being selected of the elements changes with the number of their registration times. As for the reduction of these errors, one possible solution is to remove the repetitive lists from the entire sampling frame, which is often impossible anyway. Another possibility is to use the unique identification method, in which a clearly defined regulation is adopted for each element to be linked with one of its duplicated registrations (for example, the first registration or earliest registration), while the other elements of the registration are treated vacancies. For example, this approach is adopted to select households from the voter registration in the United Kingdoms. In urban regions, voters are registered in the constituency and numbered follow their street addresses. By using the voter identification number, it is not difficult to draw a voter sample through systematic sampling. If a voter drawn is the first person registered using the address, then this address is selected, while the second and subsequent voters registered using this address will be taken as vacancies. Then, by using the method of all selected to deal with the whole group, all households within the selected addresses will be included into the sample (Carlton. 2003:66). These two influencing factors above of sampling frame errors essentially involve the problem of complex connection and duplicate registrations. In order to correctly determine the probability of the sample unit or element being selected, we need to know the number of times for each sampling element to connect with the sampling frame units or lists, and eliminate this phenomenon before taking samples.

4) Group Element

The group element record is an important factor to bring the sampling frame problem, and is also a factor to affect the error of a sampling frame. For group elements mostly are involved in the group sampling. The sampling units in the group sampling frame are the group of target population elements. For example, the household includes more than one target population elements – person. If the size of groups is not taken into account, there will be errors for sample estimators when an element is taken from the group, because the probability of this specific sample element being selected is not known in the sample. The errors are often caused because the probability of each sample unit being selected is assumed equal

to that of the sampling frame being selected, while these sampling frame units are the group of elements. When the sample households are selected, the statistics is based on purposeful or randomly selected households of the respondents with person as measurements, while in the calculation of statistics the probability of households being selected is used. Therefore, this type of errors would occur (Judith and William. 1997:56).

In order to avoid the selection deviation, in practical application, we must use a strict probability mechanism to select elements from the selected group of elements. In this regard, Carlton proposed some solutions (Carlton. 2003:61-64). If the elements in the group are selected with equal probability as usual, the first method to solve the problem of the group of elements is to take the elements from the group of selected elements, which will produce an equal probability sample. However, because of the risk of mutual influence, this approach is often unacceptable. The second method is to select an element randomly from the group of selection elements, so as to avoid the risk of mutual influence, but at the cost of changing the selected probability. The third method is to use two-phase sampling, which is to select an element from each element and maintain an equal probability sample. For the "all selected" method, which gives equal opportunities to the various elements in the groups and their group elements, so that a sample composed of complete groups is obtained. However, because of the risk of mutual influence, this approach is often unacceptable.

5) Mistake Auxiliary Information and Inaccurate Sampling Frame

For special sample designs and estimation techniques, auxiliary information is needed to be included in the sampling frame, such as stratification sampling by using this auxiliary information, probability proportional to size sampling by the size measurement. In the cases that the characteristics of auxiliary information are highly relevant to the characteristics of variables in the survey, we also can use this auxiliary information to conduct ratio estimates and regression estimates. However, the mistakes of auxiliary information cannot be avoided. Moreover some sampling units do not have the auxiliary information. For this type of errors, their main influence is to reduce the accuracy of estimation. In addition, old sampling frames can result in that relevant information cannot be regained, or the sample units cannot be located. This type of sampling frames inevitably leads to the variance of the sampling and survey. To overcome these two problems, only through the rigorous process of careful design and proper sampling, and seriously implement of the sampling, can the error reduction be achieved to some extent.

REFERENCES

Li Peiliang.(2001). *Statistical Application of Social Research*. Social Sciences Academic Press.

Feng Xiaotian.(2001). *Research Approach of Sociology*. China Renmin University Press.

L • Kish.(1997). *Sample Survey*. China Statistical Publishing House.

Zhao Junkang. (2002). *Sampling Design Theory and Method of Statistical Survey*. China Statistical Publishing House.

Graham • Carlton. (2003). *Introduction to Sample Survey*. China Statistical Publishing House.

T • Judith, D • William. (1997). *Non-sampling Error of Survey.* China Statistical Publishing House.

Li Jinchang. (1996). *Sampling Investigation and Inference.* China Statistical Publishing House.

Li Qiang. (1995). *Applied Sociology*. China Renmin University Press.

Al • Bobby.(2009). *Social Research Metho*ds (11th edition), Huaxia Publishing House.

ZHU Tie, et al. (2001). *Confusion of Sample Survey*. Statistical Research, *12.*

Feng Xiaotian. (2003). *Result Representation and Method Application – Analysis and Research of 141 Surveys*. Sociological Research, *2*.