Chapter VIII
Non-observation error in household surveys in developing countries

James Lepkowski
University of Michigan
Ann Arbor, Michigan, United States of America

Abstract

Non-observation in a survey occurs when measurements are not or cannot be made on some of the target population or the sample. The non-observation may be complete, in which case no measurement is made at all on a unit (such as a household or person), or partial, in which case some, but not all, of the desired measurements are made on a unit. The present chapter discusses two sources of non-observation, non-coverage and non-response. Non-coverage occurs when units in the population of interest have no chance of being selected for the survey. Non-response occurs when a household or person selected for the survey does not participate in the survey or does participate but does not provide complete information. The chapter examines causes, consequences and steps to remedy non-observation errors. Non-coverage and non-response can result in biased survey estimates when the part of the population or sample left out is different than the part that is observed. Since these biases can be severe, a number of remedies and adjustments for non-coverage and non-response are discussed.

Key terms: non-response, non-coverage, bias, target population, sampling frame, response rates.

A. Introduction

1. Non-observation in survey research is the result of failing to make measurements on a part of the survey target population. The failure may be complete, in which case no measurement is made at all, or partial, in which case some, but not all, of the desired measurements are made.

2. One obvious source of non-observation is the sampling process. Only in a census, which is a type of survey designed to make measurements on every element in the population, is there no non-observation arising from drawing a sample. Non-observation from sampling gives rise to sampling errors that are discussed in chapters VI and VII of the present publication**.** This source of non-observation will therefore not be treated here.

3. The present chapter will discuss two other sources of non-observation, namely, non-coverage and non-response. As will be explained in more detail later, non-coverage occurs when there are units in the population of interest that have no chance of being sampled for the survey; and non-response occurs when a sampled unit fails to participate in the survey, either completely or partially. The chapter will address the causes of these sources of non-observation, their potential consequences, steps that can be taken to minimize them, and methods that attempt to alleviate the bias in the survey estimates that they can generate. The consequences of non-coverage and non-response include the possibility of bias in the results obtained from the survey. If the part of the population that is left out is different than the part that is observed, there will be differences between the survey results and what is actually true in the population. The differences are non-observation biases, and they can be severe.

4. Of course, non-observation bias may not occur at all, even when measurements are not made on a portion of the population. While recording instances of non-observation is somewhat straightforward, detection of non-observation bias is difficult. This difficulty is what makes consideration of non-observation bias an infrequently researched topic. It is possible to find examples where non-observation makes no difference at all in an entire survey, or as regards most survey questions. It is also possible to find examples where non-observation has led to substantial bias in the survey estimate from a single question, or substantial biases in the estimates from a set of questions, in which case all the results from the survey become suspect.

5. There has been a great deal of research on non-observation. This chapter can provide only an introduction to the nature of non-coverage and non-response errors in household surveys. The reader is referred to the references provided for more detailed treatments. The next section provides a framework for distinguishing between non-coverage and non-response and is followed by separate sections on each source of error.

B. Framework for understanding non-coverage and non-response error

6. Knowing the difference between non-coverage and non-response requires an understanding of the nature of populations and sampling frames. The target population is the

collection of elements for which the survey designer wants to produce survey estimates. For example, a survey designer may be called upon to develop a survey to study labour-force participation for persons aged 15 years or over living in a given country. The population clearly has geographical limits that are well defined (the borders of the country), and limits on the characteristics of the units, such as age restrictions.

7. There are other implicit aspects of the target population definition; for example, the meaning of a person living in the country. Many surveys use a definition of residence according to which a person must have lived in the country the majority of the past year or, having just moved into the country, must intend to stay there permanently. Some portions of the population may be out of scope for a certain survey topic. For example, persons living in prisons or jails, or other institutions such as the military, may be defined as out of scope for some surveys of economic conditions. Thus, institutions may be excluded because they contain persons who are not part of the conceptual basis for the measurement to be made. There is also an implied temporal dimension to the target population definition. The survey is probably interested in current labour-force participation and not historical patterns for the individual. If so, the survey is concerned to make estimates about the characteristics of the population as it exists at a particular point in time.

8. The target population is also the population of inference. The survey results will, in the end, be said to refer to a particular population. Surveys are often designed to measure the characteristics of persons in a given country. Regardless of whether some persons in the country are covered by the sampling process or not, the survey's final report may make unqualified statements about the entire population. For example, even though the survey excluded persons living in institutions, the final report may state that the results of the survey apply to the population of persons living in the country. The uninformed reader may then assume that the results represent persons living in institutions, even though they were not covered by the sampling process. It is thus important in describing the survey to include careful and complete statements about the target and survey populations in publications about the survey.

9. The target population will often differ from another important population, the set of elements from which the sample is actually drawn, called the sampling frame The sampling frame is the collection of materials used to draw the sample, and it may not match exactly with the target population. For example, in some countries, address registries prepared and maintained by a public security agency, such as the police, are used as a sampling frame. But some households in the population are not in those administrative systems. The frame then differs from the target population.

10. In other instances, the frame differs from the target population for structural, or deliberate, reasons. A portion of the population may be left out of the frame for administrative or cost reasons. For example, there may be a region, several districts, or a province in a country where there is current civil unrest. Public security agencies may place restrictions on travel into and out of the region. The survey designer may deliberately leave the region out of the frame, even though materials exist to draw the sample in the region.

11. Cost may also enter into a decision to exclude a portion of the population. In many countries, those living in remote and sparsely population areas are excluded from the sampling frame because of the high cost of surveying them if they are sampled. Furthermore, since in countries with many indigenous languages, separate translations and the hiring of interviewers who can speak all languages are expensive, survey designers may, in conjunction with survey sponsors, specifically exclude population members who do not speak one of the major languages in the country. In this case, it may not be possible to exclude a person until after a household has been identified and the language abilities of the persons in the household have been determined. The exclusion is made through a screening in the household.

12. On the other hand, survey designers may choose to classify this kind of a problem as non-response, that is to say, as non-coverage due to language exclusion or non-response due to inability to communicate. The decision about how to classify "language exclusions" depends in part on the size of the problem. For example, in one country the survey may be limited to populations who can speak one of several officially recognized languages. This decision may exclude substantial numbers of persons who do not speak those languages. In contrast, in another country, where nearly everyone speaks one of the official languages, small population groups speaking non-official languages for which questionnaire translations are not available may be contacted but not interviewed. In the former instance, it may be appropriate, with careful documentation, to classify the excluded language groups as non-coverage. In the latter, it is appropriate to classify the non-interviews as non-response.

13. Non-coverage arises when there are elements in the target population that do not correspond to listings in the sampling frame. In household surveys, typical non-coverage problems arise when housing units fail to be included in a listing prepared during field operations, when out-of-date or inaccurate administrative household listings are used, or when individuals within a household are omitted from a household listing of residents.

14. Non-coverage refers to a failure to give an element in the population a chance of being selected for the survey's sample, whereas non-response is due to an unsuccessful attempt to collect survey data from a sampled eligible unit, a unit in the target population. Non-coverage arises due to errors or problems in the frame being used for sample selection; non-response arises after frames have been constructed, and sample elements selected from the frame. For example, suppose that in a sampled household a male resident of the household is absent at the time of interview because he is spending the week away at a temporary job outside of the village where the household is located. If that resident is not listed on a household roster during initial interviewing because the household informant forgot about him, non-coverage has occurred. On the other hand, if a resident is listed on the roster, but he is away during the interviewing period in the village and the survey accepted only self-reported data from the resident himself, and hence no data were collected from him, that resident is a non-respondent.

15. Non-coverage typically involves entire units, such as households or persons. Non-response can involve entire units, or individual data items. For example, non-coverage might involve the failure to list a household in a village roster because it is located above a retail shop. The entire unit is absent from the frame. Non-response might occur because the household, when listed, refuses to participate in the survey, or because some members of the household

cooperate, and provide data, while others are not at home or refuse to respond to the survey entirely. These two forms of unit or total non-response, household or person, are in contrast to the case where a member of the household provides data in response to all survey questions except a subset. For example, a household respondent may refuse to provide data about his or her earnings in the informal economy, perhaps because of a concern about official administrative action on unreported income. This latter form of non-response is known as item non-response. Note that the type of non-response in this case also depends on whether the unit of analysis is the person or the household: person-level non-response is item non-response for analysis at the household level, but unit non-response for analysis at the person level.

16. It is also important to consider the trade-offs between non-coverage and non-response. While many sources of non-coverage or non-response might be identified for a given survey through careful study, and there may be a desire to reduce the size of either of these problems, reduction will require the expenditure of scarce, and limited, survey resources. There may then be a competition for these resources with respect to reducing these two sources of error.

17. For example, suppose that in a country with 40 major languages or dialects, the survey instrument is translated into 5 languages that are spoken in the households of 80 per cent of the population. The sixth most frequently spoken language group represents 3 per cent of the population. At the same time, suppose that survey operations specify two visits to a household over a two-day period in order to find someone at home, and that it is known that 10 per cent of the households visited twice will be non-responding because no one is at home during two days of the survey interviewing. The survey designer has a choice in terms of resources. More funds could be spent to translate the instrument into a sixth language to cover an additional 3 per cent of the population speaking the sixth language. Or more funds could be spent on having interviewers spend a third or fourth day in each village to conduct household visits to try to find a higher proportion of household members at home.

18. The decision about how to use any extra survey resources, for translation or for additional household visits, will depend on the size of the anticipated biases and the costs and resources involved. The biases depend on both the level of non-coverage or non-response and on the differences between covered and not-covered populations, or responding and non-responding sample persons.

19. These kinds of cost-error trade-offs occur frequently in survey design. It is beyond the scope of this chapter to consider in any detail the kind of data needed to make such trade-offs or how the trade-offs are made. In most surveys, such trade-offs are based on limited information and made informally.

C. Non-coverage error

1. Sources of non-coverage

20. The sources of non-coverage in household surveys depend on the frame materials used to select the sample. Since many household surveys in developing countries, and some transition

countries, involve area sampling methods, the present discussion will limit the frame and non-coverage problems to household surveys based on area samples.

21. Area sampling is also usually coupled with multistage selection. Primary and sometimes secondary stages of selection involve geographical areas that can be considered clusters of households. In some subsequent stage of selection, a list of households must be obtained, or created, for a set of relatively small geographical areas. At the last stage of selection, a list of persons or residents in the household is created in each sampled area. There are thus three types of units that need to be considered when examining non-coverage in such surveys: geographical units, households, and persons. As discussed later, these units also may be separate sources of non-response in household surveys.

22. Non-coverage of geographical units as a result of deficiencies in the sampling frame is rare, because most area frames will be based on census materials that cover the entire geographical extent of a population. Non-coverage of a geographical area does arise, but in a more subtle form, as mentioned above. A survey may be designed to provide inferences to the entire population of a country or region within a country, and references to the population in the final report may indeed include the population living in the entire area, but the sample may not be selected from the entire country.

23. For example, during the survey design, the survey designers may identify some geographical areas with limited shares of the population that are extremely costly to cover. They may make a deliberate decision to exclude those geographical areas from the frame. Yet, in reporting results for the survey, the deletion of these areas is not mentioned, or only mentioned briefly. Report readers may have, or be given implicitly, the impression that survey results apply to the entire country or region, when in fact a portion of the population is not covered. In practice, the size of the non-coverage error arising in such situations is generally small, and typically ignored.

24. It is important to keep in mind that the distinction remains between a desired target population (that is to say, the population living in the entire geographical area of the country) and a restricted "survey population" living in the included geographical area. There is a danger, though, that through incomplete documentation, the user of the data may be under the impression that the survey sample covers the entire population, when in fact it does not.

25. A more important source of non-coverage occurs at the household level. Most surveys consider households to be the collection of persons who usually reside in a housing unit. Two components are thus important: the definition of a usual resident and the definition of a housing unit.

26. Housing unit definitions are complex, inasmuch as they take into account whether a physical structure is intended as living quarters, and whether the persons living in the structure live and eat separately from others in the same structure (as in multi-unit structures such as apartment buildings). Living separately implies that the residents have direct access to the living quarters from the outside of the structure, or from a shared lobby or hallway. The ability to "eat

separately" usually involves the presence of a place to provide and prepare food, or the complete freedom of the residents to choose the food they eat.

27. Applying this kind of broad definition to the many diverse living situations across countries, or across regions of a country, is difficult. Most housing units are readily identified, such as single family or detached housing units, duplexes where separate housing units share a wall but have separate entrances, and apartments in multi-structure buildings. However, there are many housing units that are difficult to classify or find. For example, in urban slum areas, separate housing units may be difficult to identify when people are living in structures built from recycled or scrap materials. Housing units may be located in places that cannot be identified by casual inspection of entrances from a street, lane or pathway.

28. In rural areas, a structure intended for dwelling may be easily identified, but complex social arrangements within the structure may make separate housing unit identification difficult. For instance, in a tribal group, long-houses with a single entrance are used for housing; they contain separate compartments for family unit sleeping arrangements, but there is a common food preparation area for group or individual family meals, that is to say, the individual compartments are not themselves housing units, because they do not have a separate entrance or their own cooking and eating area. In such an arrangement, the notion of a household as the group of persons who usually reside in a specific housing unit is more difficult to apply. It is not clear whether the entire structure, or each compartment, should be treated as a housing unit. In practice, the entire longhouse is treated as a housing unit or dwelling and, if sampled, all households identified during the field listing of households are included in the survey.

29. There are also living quarters that are not considered housing units. Institutional quarters occupied by individuals under the care or custody of others, such as orphanages, prisons or jails, or hospitals, are not considered to be housing units. Student dormitories, monasteries and convents, and shelters for homeless persons are special types of living quarters that do not necessarily provide the care or custody associated with an institution. Living quarters for transitional or seasonal living are also a problem. For example, there may be separate housing units present in an agricultural area for housing seasonal labour, which are occupied for only one season, or a few seasons each year. Presumably, the seasonal residents usually live elsewhere, and should not be counted as part of a household in the seasonal unit.

30. Multistage area sampling in developing countries requires that at some point in the survey process lists of dwellings be created for small geographical areas, such as a block in a city or an enumeration area in a rural location. Non-coverage often arises when part-time survey staff are sent to the field to list housing units, and encounter the kinds of complex living quarters described above. Identification of most housing units is straightforward; but the missing of housing units may still be common to the extent that the part-time staff has limited experience applying to complex living quarter arrangements a definition that has several components.

31. The non-coverage problem in housing unit listing is made more difficult by the temporal dimension. A housing unit may be unoccupied at the time of listing, or under construction. If the survey is to be conducted at some point in the future, these types of units may need to be included in the listing. In surveys where housing unit listings are used across multiple waves of

a single panel survey, or across several different surveys, it is common to try to include construction units that are unoccupied or under construction.

32. In surveys in transition countries, it may be possible to use a list already prepared by an administrative authority. However, the quality of those lists for household surveys needs to be carefully assessed. The same kinds of problems outlined here that could arise in survey listing are likely to occur in respect of administrative lists.

33. Thus, the housing unit listing process can generate non-coverage of certain types of households. This non-coverage may be difficult to identify without substantial investment of additional survey resources.

34. Finally, within a sampled housing unit, listing of persons who are usual residents is a part of the household listing process as well. Operational rules are required to instruct interviewers regarding whom to include in the housing unit as a usual resident. As in the case of housing units, most determinations are straightforward. Most persons encountered are staying at the housing unit at the time of contact, and it is their only place of residence. There are others who are absent at the time of contact, but for whom the residence is an only residence.

35. However, there are persons for whom the housing unit is one of several in which they live. A decision must be made in the field by part-time staff about whether the sampled housing unit is the usual place where this person resides. It is also difficult for household informants to report accurately on the living arrangements of some residents. This reported proxy information about another resident may not be completely accurate.

36. Informants may also have personal reasons for deliberately excluding persons whom they know to be usual residents. For example, a person may be living in a housing unit who would make the household ineligible for receiving the government benefits that it is already receiving. Also, an informant may deliberately exclude a resident who does not want to be identified by public or private agencies because of financial problems (such as debt) or legal problems (such as criminal activity).

37. Informants may also not include someone in the household for cultural or cognitive reasons. An informant may not report an infant less than one year of age because the culture does not consider these persons old enough to be regarded as persons. They may also exclude infants, because they believe that the survey organization is not interested in collecting data about young children; or they may simply forget to include someone, whether it is an infant or someone older.

38. Non-coverage in household surveys may thus arise from a variety of definition and operation circumstances. The concern must be the extent to which non-coverage leads to error in survey results.

2. Non-coverage error

39. Suppose that the survey is to estimate the mean for some characteristic Y for a population of N persons, N_{nc} of whom are not covered by the survey's sampling frame. Let the mean in the

population of size N be $\overline{Y}$, let $\overline{Y}_c$, be the mean of those covered by the sampling frame, and let $\overline{Y}_{nc}$ be the mean of those not covered by the frame. . The error associated with the non-coverage is referred to as the non-coverage bias of the sample mean, $\overline{y}_c$, which is based only on those covered in the sample, and which in fact estimates $\overline{Y}_c$ rather than $\overline{Y}$.

40. The bias of the sample mean, $\overline{y}_c$, depends on two components, the proportion of the population that is not covered, N_{nc}/N, and the difference in the means of the characteristic Y between covered and not-covered persons. Hence,

$$B\left(\overline{y}_c\right) = \left(N_{nc}/N\right)\left(\overline{Y}_c - \overline{Y}_{nc}\right)$$

41. This formulation of the non-coverage bias is helpful in understanding how survey designers deal with non-coverage. In order to keep the error associated with non-coverage small, or to reduce its effect, the survey designer either must have small differences between covered and non-covered persons, or must have a small proportion of the persons who are not covered by the survey.

42. An important difficulty with this formulation is that, in most surveys, neither the difference $\left(\overline{Y}_c - \overline{Y}_{nc}\right)$ nor the proportion $\left(N_{nc}/N\right)$ not covered is known. Further, the non-coverage rate $\left(N_{nc}/N\right)$ may also vary across subclasses. The difference may vary across different variables and across subclasses of persons (such as a region, or a subgroup, defined by some demographic characteristic such as age). Thus, non-coverage error is a property not of the survey but of the individual characteristic, and of the statistic estimated.

43. In many government survey organizations, estimates of a total are frequently required. The non-coverage bias associated with a total depends on not only the differences between covered and non-covered units on the characteristic of interest but also on the number (and not the rate) of non-covered, that is to say, for an estimated total for respondents $\hat{Y}_r = N\overline{y}_r$, the bias is $B\left(\hat{Y}_r\right) = N_{nc}\left(\overline{Y}_r - \overline{Y}_m\right)$.

Reduction, measurement and reporting of non-coverage error

44. There are four possible means of handling non-coverage error in household surveys:

- Reducing the level of non-coverage through improved field procedures.
- Creating procedures to measure the size of the non-coverage error and reporting the level in the survey.
- Attempting to compensate for the non-coverage error through statistical adjustments.

- Reporting non-coverage properties of the survey as fully as is possible in the survey report.

45. The reduction of non-coverage error in household surveys is usually attempted either through the use of multiple frames or through methods to improve the listing processes involved in the survey. Multiple frames are more likely to be used for housing units rather than persons. They require the availability of separate lists of housing units that pose particular problems for field listing.

46. For example, suppose that seasonal housing units for agricultural workers are known to be difficult to list properly in the field in a given country. Suppose also that an agency responsible for agricultural production, education, or social welfare has a list of the number and type of seasonal housing units on farms or enterprises where seasonal labour is employed and housed. The list of seasonal housing units from the alternative source may be used as a separate frame. Field interviewers preparing housing unit lists would be given a list of farms or enterprises where agency lists were already available in the area they are to list, and told not to list seasonal housing units there. Samples of housing units for the survey would then be selected from the housing unit list prepared by the interviewer and from the list maintained by the government agency. There will no doubt remain some non-coverage across both lists, and possibly some "over-coverage" may occur as well; but the use of both frames may reduce the level of non-coverage, and the error associated with it.

47. It is also important to consider methods to improve the listing processes. When housing unit lists are available from an administrative source, they may be checked by a field update before the sample is drawn. Interviewers may be sent to geographical areas with a list of housing units from the administrative source, and given instructions on how to check and add, or delete, housing units from the list as they examine the area.

48. Interviewers may also be trained to use a "half-open interval" procedure in the field to capture missed housing units from administrative lists or field lists that have missing units. The half-open interval procedure involves the selection of a housing unit from an address list, a visit by an interviewer to the sampled unit, and an implied or explicit list order. At the unit, the interviewer is instructed to enquire about any additional housing units that might be present between the selected housing unit and the next one on the list.

49. The next unit on the list is defined by some kind of pre-defined route through a geographical area. For example, on a city block, interviewers preparing a listing are instructed to start on a particular corner, and then proceed in a clockwise direction around the block. The housing unit list is to be assembled in that clockwise order.

50. If an interviewer finds a housing unit that is not on the list, and between the selected housing unit and the next on the list, he or she is instructed to add the missed housing unit to the sample and attempt an interview. If there are several such missed units, the interviewer may need to contact the survey central office for further instructions so as to avoid disruptions to field operations.

51. Within households, improved listing procedures may involve question sequences administered by the interviewer to the housing unit informant to identify missed persons. For example, the survey interviewer may be instructed to ask about any infants who may have been left off the list of usual residents. The household listing may also be improved if interviewers are given guidelines about the choice of suitable informants or instructions to repeat the names on the list of persons to the informant to be sure no one was overlooked.

52. Measurement of non-coverage bias is also an important consideration, although a difficult problem to address. How does a survey organization identify units that are not included in any of its lists? As measurement of non-coverage can be an expensive survey task, it is one that is undertaken only occasionally.

53. A common way to assess non-coverage error is to compare survey results, for those variables for which comparisons can be made, with findings from external or independent sources. To assess the size of non-coverage, a survey may compare the age and gender distribution of its sample persons with the distribution obtained from a recent census, or from administrative records. Differences in the distributions will indicate non-coverage problems. To assess the non-coverage error associated with a variable, a comparison of values of the statistic of interest to an independent source may be made. For example, total wage and salary income reported in a survey, for the total sample and for key subgroups, may be compared to administrative reports on wage and salary income. In a classic study, Kish and Hess (1950) compared the distribution of housing units in a survey with recent census data on the distribution of housing units at the block level. The comparison provided insight into the nature of the non-coverage problem in the survey data collection.

54. A more expensive non-coverage error assessment can be made through dual system measurement, or related case matching procedures. Censuses employ dual system methods to assess coverage of a census operation [see, for example, Marks (1978)]. In a census, a separate survey is compared with census results to identify non-coverage problems. The assessment of the size of the non-coverage depends on a case-by-case matching of survey sample to census elements to determine which sample elements did not appear in the census. These procedures are closely related to the methods of "capture-recapture sampling" used in environmental studies of animal populations.

55. Since household surveys are universally affected by non-coverage error, many surveys will employ post-stratification or population control adjustments as statistical procedures to adjust survey results so as to compensate for non-coverage error. These adjustments are very similar to the method outlined above for assessing the size of the non-coverage error. The sample distribution by age and gender, for example, may be compared with the age and gender distribution from an outside source, such as a recent census or population projections. When the sample distribution is low (or high) for an age-gender group, a weight may be applied to all sample person data from that age-gender group to increase (decrease) their contribution to survey results. Weighted estimators will be required to properly handle the weights in analysis.

56. As a final consideration for non-coverage, good reporting is important for any statistical organization. Analytical reports ought to give clear definitions of the target population,

including any exclusions. The frame should be described in enough detail for the reader to see how non-coverage might arise, and even make an informal assessment of the size of potential error. It would be helpful to include as references or appendices, any quality assessments of the frame, such as checks of the quality of housing unit lists or administrative lists, or comparison of original lists of persons within housing units with those lists obtained from reinterviews carried out for the purpose of quality control assessment.

57. A more difficult problem is the reporting of any coverage rates or non-coverage bias for the population and subclasses of the population. These kinds of assessments may be possible only for ongoing surveys where at some time there has been an attempt to assess the size of the non-coverage problem. It is very difficult if not impossible to make such assessments for one-time cross-sectional surveys.

58. Finally, if post-stratification or population control adjustments are made, the survey documentation must contain a description of the adjustment procedures and the magnitudes of the adjustments for important subgroups of the population.

D. Non-response error

59. Non-response error suggests a number of parallels with non-coverage error in terms of definitions, measurement, reduction, compensation and reporting. The organization of the present section is thus very similar to that of section C. It is important to make clear, however, that non-response and non-coverage are quite separate problems, having different sources and, in a few instances, different solutions. While in non-coverage survey designers almost never know anything other than the location and general characteristics of the non-covered portion of the population, in non-response they know at least frame information for non-respondents. Non-response is also believed to be more extensive in household surveys, and thus its contribution to the bias of survey estimates may be larger.

60. As noted above, two types of non-response are often identified in household surveys, namely, unit non-response and item non-response. These two types have quite different implications for survey results, and the methods used to measure, reduce and report them, and to compensate for them, are in some ways distinct as well. While a separate section could be devoted to each type, both will be addressed together in this section.

1. Sources of non-response in household surveys

61. In household surveys, unit non-response can occur for several different kinds of units. As is the case for non-coverage, non-response may occur for primary or secondary sampling units. For example, a primary sampling unit might consist of a district or sub-district in a country. Weather conditions or natural disasters may prevent survey operations from being conducted in a district or sub-district that has been selected at a primary, or secondary, stage of sampling. The unit is covered by the survey, but during the survey period, it is not possible to collect data from any of the households in the unit.

62. Non-response is more frequent at the household level. A listed housing unit chosen for the sample may be found occupied, and an interview attempted. However, as the interviewer visits the housing unit, several adverse events may prevent data collection. A household member may refuse participation as an individual or as a representative of the entire unit.

63. Although a housing unit is occupied, its residents may be away from home during the entire survey period. In some developing countries, a considerable problem is encountered with housing units clearly lived in but locked during the entire data-collection period.

64. In many countries, although occupied housing units have individuals home at the time of data collection, language may pose a barrier. A version of the survey's questionnaire may not have been translated into the language of the household, or the interviewer may not speak the local language. To avoid non-response, surveys may hire translators locally to accompany interviewers to the doorstep and translate interactively. Other surveys reject this practice because of concerns about whether the translation is correct, and whether the translation is consistent across households. Households that cannot provide responses, though, because of language difficulties, can be classified as non-responding units. As an alternative approach, it is the practice of some survey organizations to exclude from the survey households that do not speak a translated language. These households then become non-covered, rather than non-responding. The particular approach chosen by the survey organization, whether to handle such units as not covered or to handle them as non-responding, must be clearly described in the survey documentation.

65. Person-level unit non-response also may occur. For surveys that allow proxy reporting on survey questions, data can be collected from other household members for persons in the household who are not at home at the time of interview. For surveys, though, that require self-report for some or all questions, a person who is not at home during the survey, refuses to participate, or has another barrier (such as language) that precludes interviewing is a non-respondent. Health conditions, whether permanent, such as hearing impairment or blindness, or temporary, such as an episode of a severe acute illness, may preclude an individual from responding as well.

66. As for households with language problems, some survey organizations choose to classify persons with language barriers or permanent health conditions as not covered, and those with temporary conditions as non-responding (Seligson and Jutkowitz, 1994). There are no widely accepted rules for deciding how to make such a classification. For a survey of income or expenditures, persons with temporary health conditions are few enough in number for the organization to be able to treat them as not covered. For a survey of health conditions, though, the responses of these individuals may differ enough for there to be concern about excluding them. They may then be classified as non-response. In view of the lack of widely agreed practice, it is important that survey organizations report clearly in survey reports exactly how such cases have been handled in a given survey.

2. Non-response bias

67. A great deal more research has been devoted to the problem of non-response in household surveys than to non-coverage [see for example, reviews by Groves and Couper (1998), and Lessler and Kalsbeek (1992)]. This increased emphasis in research is related to several factors.

68. Non-coverage is, in a certain sense, less visible than non-response. The non-covered households or persons are simply not available for study, while non-responding units can be observed and counted, and possibly persuaded to participate.

69. There is a presumption in developed countries that non-coverage is less important than non-response because the non-coverage rate is lower than the non-response rate. The opposite may be true for developing countries where non-response rates are lower and non-coverage rates much higher than in developed countries. Recall that non-coverage bias for a sample mean is attributable to two sources, the size of the non-coverage rate and the size of the difference between the means for the covered and not covered population groups. Similarly, for non-response, the size of the non-response bias for a sample mean can be attributed to the proportion of the population that does not respond and the size of the difference in population means between respondent and non-respondent groups.

70. Following the development for non-coverage, suppose that the survey is to estimate the mean for some characteristic *Y,* and that the mean in the population $\overline{Y}$ is composed of a mean for persons who respond, say $\overline{Y}_r$, and a mean for those not responding, $\overline{Y}_{nr}$. Let N_{nr} denote the number of persons who would not respond if they were sampled. The bias of the sample mean for respondents $\overline{y}_r$ is then $B(\overline{y}_r) = (N_{nr}/N)(\overline{Y}_r - \overline{Y}_{nr})$. As for non-coverage, the survey designer must either keep the non-response rate small, or anticipate small differences between responding and non-responding households and persons. This general framework can be used to understand further non-response at the item level. The problem of item non-response bias is more complicated, though, because often items are considered in combinations, and item non-response is the union of non-responses across several items.

71. While in non-coverage neither the difference nor the rate is known, for non-response, carefully designed surveys will provide good estimates of the non-response rate. Carefully designed surveys maintain detailed records of the disposition of every sample unit, whether household, person, or individual data item, that is selected for study. They can then estimate the non-response rate directly from survey data. They may also have data to observe if response rates differ across important subclasses, particularly geographical subclasses for households.

72. Evaluating differences between respondents and non-respondents requires more extensive data collection and measurement. It is often impossible during survey data collection to attempt measurement of characteristics of interest for survey non-respondents. Special studies designed to elicit responses from non-responding units can, however, be conducted during the course of a survey.

73. Non-response in later waves of panel surveys provides more data for studying and adjusting for the effects of potential non-response bias than non-response in one-time or cross-sectional surveys. Panel surveys are ones in which the same units are followed and data are collected from the panel units repeatedly over time. A portion of the units can be lost to follow-up, leading to panel or attrition non-response over the course of the survey. Investigations of panel non-response can, however, use the data collected on previous panel waves to learn more about differences between respondents and non-respondents, and to serve as the basis for the kind of adjustments described below. Techniques for compensating for panel non-response are described in Lepkowski (1988).

74. The availability of slightly more information about non-respondents than about non-covered persons, and the potential use of behavioural models to study and compensate for non-response have also led to more research on non-response than on non-coverage. When careful records are kept on all sample units, and not just responding ones, comparisons between respondents and non-respondents can be made directly from sample data. Further, non-response is partly generated by household or person behavior: it is a self-selection phenomenon. The survey designer can turn to an extensive literature in sociology, psychology and social psychology to study how individuals and groups make decisions about participation in various activities. Behavioural models can be examined, provided some data are available for non-respondents, to understand the determinants of non-response in a survey.

3. Measuring non-response bias

75. Measurement of non-response bias requires measurement of non-response rates and measurement of differences between respondents and non-respondents on survey variables. Non-response rate calculation for households or persons from sample data in turn requires definition of possible outcomes for all sampled cases, and then specification of how those outcomes should be used to compute a rate. For example, completed and partial interviews (those that have sufficient data to provide information on key study concepts) are often grouped together.

76. Eligible non-interview cases are those that are in the population and identified through the survey operation, but from whom no data were collected. For example, if a survey is restricted to persons aged 15 years or over, then eligible non-interviews are those person aged 15 years or over for whom no data were collected. There are usually at least three sources of non-interviews: refusals (Ref) or persons or households that have been contacted, but will not participate in the study; non-contacts (NC) or eligible persons or households where contact cannot be established during the course of the data collection; and other (Oth) or those non-interviews occurring for some other reason, such as language difficulty or a health condition. Finally, there are also cases that are not eligible (Inelig) for the survey (for example, those under age 15), and those with unknown eligibility (Unk).

77. The response rate in this simplified set of outcomes can be computed in several different ways. A commonly accepted method of response rate calculation (where “Int” denotes the number of completed and partial interviews in a survey) is

$$\bar{R} = \frac{\text{Int}}{\text{Int+Ref+NC+Oth+}\varepsilon \times \text{Unk}}$$

Here, some proportion, ε, of the unknown eligibility cases are estimated to be eligible. Often, this estimated eligibility is computed from the existing data by using the rate of known eligibility (those cases with outcomes Int, Ref, NC and Oth) among all cases for which eligibility has been determined. Hence

$$\hat{\varepsilon} = \frac{\text{Int+Ref+NC+Oth}}{\text{Int+Ref+NC+Oth+Inelig}}$$

78. Household surveys that repeatedly interview the same households, or a panel of persons selected from a household sample, have additional non-response considerations that affect the calculation of response rates. Such longitudinal panel surveys have unit non-response at the initial wave of interviewing as in a cross-sectional survey, and in addition may be unable to obtain data at later waves from some panel members. Response rate calculations must take into account the losses due to non-response for the initial as well as the subsequent waves of data collection. It is beyond the scope of the present publication to address the calculation of response rates in panel surveys. More on this subject can be found on the American Association for Public Opinion Research web site (http://www.aapor.org. Path: Survey Methods).

79. Measures of differences between respondent and non-respondent means, or other statistics, are more difficult to obtain. One can compare survey results with those of outside sources for some variables in order to assess whether there is a large difference between the survey and the external source in terms of the value of an estimate; this approach, however, may be difficult to apply because there may be differences in definitions and methodology between the survey and the external source that complicate interpretation of any observed difference. In other words, the difference between the survey estimates and the external source estimates may be attributed to causes other than non-response.

80. The measurement of differences between respondents and non-respondents is expensive. In principle, with sufficient resources, it is sometimes assumed that responses can be obtained from non-responding cases. However, the resources are seldom available for the attempt to obtain data from every non-responding case. As an alternative, a second phase or double sample can be drawn from among the non-respondents, and all remaining survey resources devoted to collecting data from this subsample.

81. Statistically, there is a modest literature about two-phase sampling for non-response concerning a number of design features (see, for example, Cochran, 1977, sect. 13.6). In the case when complete response is obtained from the two-phase non-response sample, it is possible to determine an optimal sampling fraction in the second phase, given cost constraints, that minimizes the sampling variance of a two-phase estimate of the mean.

4. Reducing and compensating for unit non-response in household surveys

82. Reducing unit non-response is, in many circumstances, achieved through ad hoc methods that appear to be sensible ways to reduce non-response rates. More recently, comprehensive

theories based on sociological and psychological principles have been posited [see Groves and Couper (1998)], from which may flow non-response reduction methods based on a more complete understanding of how non-response operates in household surveys. It is beyond the scope of this chapter to describe these more comprehensive theoretical frameworks. Instead, several techniques that have been shown to be effective in reducing non-response in experimental studies are described.

83. Repeated visits, or "callbacks", are a standard procedure in most sample surveys. Survey interviewers do not make just one attempt to contact a household, or an eligible person, but "callback" on the household or eligible person to try to obtain a completed interview. The number of callbacks to be made, callback scheduling, and interviewer techniques for persuading reluctant or difficult-to-contact respondents to participate are all subjects of research in the field. However, there is no single recommended standard for these survey features. Differences between countries in response rates, public acceptance of surveys, and population mobility make it impossible to establish a unified theory on callbacks. Public receptiveness to surveys on different topics makes it difficult to establish callback standards even in a single country across different kinds of surveys. However, it is always advisable to use the best interviewers for the difficult task of refusal conversion.

84. There is no empirical evidence that a single technique, including callbacks, yields high response rates in household surveys. Often a combination of techniques is employed. Interviewer-administered household surveys that use advance notification in the form of a telephone call or advance letter, personalization of correspondence, information about sponsorship of the surveys, and providing potential respondents with illustrations of how the data are being used have all been shown to increase response rates. Incentives are controversial in surveys in developing and transition countries, and they are discouraged in many countries. They are becoming widespread in surveys in developed countries [see Kulka (1995) for a review of research literature on the technique].

85. Response rates can also be improved through attention to interviewer technique. Interviewer training to prepare interviewers to tailor their approach to the different reactions they receive from respondents can appreciably improve response rates. Incentives paid to interviewers based on monitored production and quality of work exceeding survey goals have also had a beneficial impact on survey response rates.

86. It is inevitable in every household survey that there will be unit non-response. Survey designs often adjust for sample size for unit non-response, as well as compute compensatory weights to provide an adjustment in estimation and analysis.

87. The sample size adjustment for non-response requires estimation prior to data collection of an anticipated unit non-response rate. The estimation is often ad hoc or particular to a survey, based on data from past survey experience with the population of interest, the topic of the survey, and other factors. In a one-time cross-sectional survey, the estimation often requires assumptions that the experience from other surveys will be reproduced in the forthcoming survey. In repeated cross-section surveys where the same population is sampled at regular, or irregular, time intervals, the data for estimating anticipated response rates are readily available.

In panel surveys, where the sample units are followed over time, the estimation requires anticipation not only of initial first-wave unit non-response but also of subsequent attrition non-response in which subjects who cooperated in earlier waves cannot be interviewed at later waves (owing to refusal, or the inability to locate them, or other factors).

88. The sample size adjustment increases sample size required for cost or precision reasons in order to have sufficient units in the sample to yield the desired outcome. Say, for example, that a final sample size of 1,000 completed interviews with households is required, and that there is an anticipated non-response of 20 per cent. In order to obtain the final 1,000 completed household interviews, the survey operation draws a sample of 1,000/(1-0.2) = 1,250. The final sample size will, to the extent that the anticipated response rate is correct, yield approximately the final required number of completed interviews. The interviewers are given an assignment of units to interview, and instructed to obtain responses from as many as possible. No substitution is allowed.

89. Another approach to handling unit non-response is substitution. This approach leaves the decision about whether to approach a unit to the interviewer, that is to say, it is subjective interviewer judgement, and not an objective probability selection, that determines which sample units are to be approached. Substitution methods for handling non-response can lead to exact sample sizes. However, there is substantial evidence [see, for example, Stephan and McCarthy (1958), who deal with a closely related non-probability procedure, quota sampling] that substitution methods lead to samples that do not match known population distributions well.

90. Statistical adjustments can be applied to the final survey data so as to compensate in part for the potential of non-response bias. The most common kind of compensation entails developing non-response adjustment weights.

91. Non-response adjustment weights require that the same information be available for all respondents and all non-respondents. Since little is known about non-respondents, the type of variables that are available for this kind of an adjustment is limited in most household surveys. In most cases, the primary information known about non-respondents is geographical location, that is to say, where the household was located.

92. For example, suppose that a household survey uses an area sampling method in which census enumeration areas are selected at the first stage of selection. During data collection, not all households chosen for the survey in a given enumeration area provide data. A simple non-response weighting adjustment scheme would assign increased weights to all responding households in an enumeration area in order to compensate for non-responding households in that area. If 90 per cent of the households in an enumeration area responded, then the weights of responding households in the area would be increased by a factor of 1/0.9 = 1.11. If in another area, 80 per cent responded, the factor would be 1/0.8 = 1.25. The weights of all responding households in the enumeration area are increased by the same factor. All non-responding households are dropped from the final sample, effectively weighting each of them by zero.

93. In some cases, weighting adjustments can be developed from a comparison of administrative data with survey respondent data. For example, administrative data may have

been used to select the sample. The sample respondents can then be assigned weights that make the distributions of weighted respondents on some key variables correspond to the distributions reported in the administrative data.

94. Non-response adjustments can also be made on the basis of a model. When response status of sampled households in a survey as simply responded or not responded, and there are data available for responding and non-responding households, response status can be regressed on the available variables. Logistic regression coefficients may be then used to predict the probability of each household responding. The inverse of the predicted probabilities can be used, much as above, to compute a weight, sometimes referred to as a response propensity weight. Since the weights computed directly from predicted probabilities tend to be quite variable, the predicted probabilities are often grouped in classes, and a single weight is assigned to each class using the inverse of the midpoint, the median, or the mean-predicted probability, or the weighted response rate in the class, as the weight.

5. Item non-response and imputation

95. An area of more recent active research has been item non-response [see, for example, the recent review by Groves and others (2002)]. With item non-response, there is a great deal of data available for each non-responding case. These data afford the opportunity for more complete understanding of item non-response, and the potential for measurement, reduction and compensation based on more complex statistical models.

96. For example, suppose that 90 per cent of the respondents to a household survey on health and health-care service availability provide answers to all questions, but 10 per cent answer all questions except one about wage and salary earnings in the previous month. The information available from the 90 per cent providing complete data can be used to develop statistical models to understand the relationship between health and health care and wage and salary income. Those models can in turn be used to posit methods for reducing the level of non-response to wage and salary income to compensate, or to predict missing values of wage and salary income.

97. The replacement of item missing values is referred to as imputation, which has been used in surveys for decades now. See Kalton and Kasprzyk (1986) and Brick and Kalton (1996) for reviews of imputation procedures used in household and other surveys. Imputation is a procedure that has been used in surveys to compensate for missing item values for decades. The basic idea is to replace missing item values with a value that is predicted using other information available for the subject (household or person, for instance) or from other subjects in the survey.

98. Imputation can be implemented, for example, through a regression model. For a variable Y in a survey, a model may be proposed for Y that "predicts" Y using a set of p other variables $X_1,\ldots,X_p$ from the survey. Such a model can be written as:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + \varepsilon_i$$

This model is fitted to the set of subjects for whom the survey variable Y and the "predictor" variables $X_1,\ldots,X_p$ are not missing. Then, the value of Y is predicted for the missing cases

using the estimated parameters obtained from fitting the above model. The predicted value of the variable *Y* for the i^{th} unit is given by:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \cdots + \hat{\beta}_p X_{pi}$$

99. This regression model for imputation is implemented in several forms. The regression prediction can include a predicted "residual" to be added to the predicted value. A technique called sequential hot deck imputation implements a form of the regression imputation that effectively adds a residual "borrowed" from another case in the data file with similar values on the $X_1, \ldots, X_p$ as the case to be imputed.

100. Recent advances in the area of imputation have also considered the problem arising from the fact that imputation introduces additional variability into estimates that use the imputed values. This variability can be accounted for through variance estimation procedures such as the "jackknife" variance estimate, or through models for the imputation process, or through a multiple imputation procedure in which the imputation is repeated multiple times and variability among imputed values is included in variance estimation.

101. There are a few techniques that can be used to reduce the level of item non-response in a survey. Survey interviewers can be trained to probe any non-codable or incomplete answer provided to any question in the survey questionnaire. Survey designers do add scripted follow-up questions to selected items that probe further when an answer such as "I don't know" or "I won't answer that question" is obtained. For example, questions about income have higher item non-response rates than other items. Surveys concerning income sometimes add a sequence of questions for some income items that "unfold" a series of ranges within which income may be reported. If the respondent refuses to answer or does not know the income amount, the unfolding questions may be: Is the income more than XXX units?, between YYY units and XXX units?, etc. These questions allow the construction of ranges within which an income is reported to occur.

102. Organizations conducting household surveys should routinely examine the frequency of item non-response across survey items to gauge the importance of the problem in the survey. Item non-response rates are seldom published, except for a few key items. The user is often left to determine the extent to which item non-response would be a problem for their analysis. Survey documentation should include item non-response rates for key items and for items with high non-response rates.

Acknowledgements

The author thanks Kenneth Coleman, Master of Science candidate in the University of Michigan Program in Survey Methodology, for his valuable assistance examining survey methods in Latin and South America.

References

Brick, J.M., and G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, vol.5, pp. 215-238.

Cochran, W.G. (1977). *Sampling Techniques.* 3rd ed. New York: John Wiley and Sons.

Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley and Sons

_____ , and M.P. Couper (1998). *Non-response in Household Interview Surveys*. New York: John Wiley and Sons.

Groves, R.M., and others (2002). *Survey Non-response*. New York: John Wiley and Sons.

Kalton, G., and D. Kasprzyk (1986). The treatment of missing survey data. *Survey Methodology*, vol. 12, pp. 1-16.

Kish, L., and I. Hess (1950). On non-coverage of sample dwellings. *Journal of the American Statistical Association*, vol. 53, pp. 509-524.

Kulka, R. (1995). The use of incentives to survey "hard-to-reach" respondents: a brief review of empirical research and current research practices. *Seminar on New Directions in Statistical Methodology*. Statistical Policy Working Paper, no. 23. Washington, D.C.: U.S. Office of Management and Budget, pp. 256-299.

Lessler, J., and W. Kalsbeek (1992). *Non-sampling Error in Surveys*. New York: John Wiley and Sons.

Lepkowski, James M. (1988). The treatment of wave non-response in panel surveys. In *Panel Survey Design and Analysis*, D. Kasprzyk, G. Duncan and M.P. Singh, eds. New York: Wiley and Sons

Marks, E.S. (1978). The role of dual system estimation in census evaluation. In *Developments in Dual System Estimation of Population Size and Growth*, K.J. Krotki, ed. Edmonton, Alberta, University of Alberta Press.

Seligson, M.A., and J. Jutkowitz (1994). *Guatemalan Values and the Prospects for Democratic Development*. Arlington, Virginia: Development Associates, Inc.