

## DEM 2110

### ADVANTAGES AND LIMITATIONS OF SURVEYS

#### Advantages of survey

- 1) **Less costly:** the overall cost of a survey is generally lower compared to a census, as the latter requires large amounts of manpower, financial, logistical and material resources.
  - From a probability sample, properly selected and implemented, accurate and reliable results can be a basis for making inferences on the total population.
- 2) **Produce statistical information of better quality** because, it is more feasible to engage better and well-trained interviewers.
  - It is also easier to provide better supervision because supervisors are usually experienced and well trained and supervisor/interviewer ratio is low.
- 3) **More detail is provided on subject matter** as surveys allow for longer time for interviews and small workloads for interviewers. *For example,*
  - The weighing of food and other measurements in a nutrition study are feasible in a survey than in a census
  - Medical examination to collect health information such as the incidence of HIV/AIDS infection by taking blood samples are feasible in a survey but not in a census
- 4) Surveys are **more flexible with respect to the subject matter covered**, making it a good source of data for meeting users' needs
- 5) The **shorter time interval** between surveys makes them more suitable for studying those population characteristics that change frequently in some countries, such as household formation, fertility, and employment status

- 6) Surveys such as the DHS simultaneously **obtains information related to enumerated events and population at risk**

#### **Limitations of surveys**

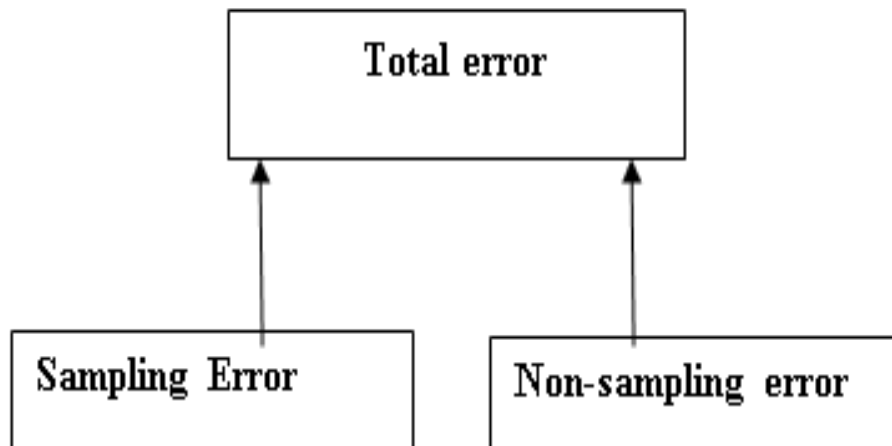
- 1) Detailed cross-classifications are subject to sampling errors
- 2) Inability to produce estimates for small areas
- 3) Coverage is only extended to a limited geographical area
- 4) Poor response rate is inevitable
- 5) Requires close supervision of field work

#### **Link between census and surveys**

- 1) Census is often used as a sampling frame for selecting the population to be surveyed.
- 2) Surveys have been used to determine the amount of error in a census, e.g. through the post enumeration survey.
- 3) Census data/statistics serve as benchmark for analyzing and evaluating survey data.

## ERRORS IN SURVEYS

- Among the disadvantages of surveys is that they are **prone to sampling and non-sampling errors**
- All survey data are subject to error from various sources
- **Sampling error** is as a result of selecting a sample instead of canvassing the whole population yet inference is made to the whole population
- **Non-sampling error** is mainly due to adopting wrong procedures in the system of data collection and processing
- The quality of a sample estimator of a population parameter is a function of total survey error, comprising both sampling and non-sampling errors
  - Total survey error is equal to sampling error plus non-sampling error
  - Graphical illustration of Total Survey Error



## NON-SAMPLING ERRORS

In general, **non-sampling errors may arise** from one or more of the **following factors**:

1. **Data specification being inadequate** and/or **inconsistent** with **respect to objectives** of the survey
2. **Duplication** or **omission of units** due to inaccurate definition of the boundaries of area units

3. Incomplete or wrong identification particulars of units or faulty methods of enumeration
4. **Inappropriate methods of interview, observation or measurement** using ambiguous questionnaires, definitions or instructions
5. **Lack of trained and experienced field enumerators** including **lack of good quality field supervision**
6. **Inadequate scrutiny of the basic data** to correct obvious mistakes
7. **Errors in data processing operations** such as **coding, keying, verification, tabulation** etc.
8. **Errors during presentation and publication of tabulated results**

## COMPONENTS OF NON-SAMPLING ERRORS

- Brieumer and Lyberg (2003) identify five components of **non-sampling error**, namely **specification, frame, nonresponse, measurement** and **processing error**. We may add that **estimation error** is another error, which should be considered.

### 1) **Specification Error**

- This occurs when the **concept implied by the question** is **different** from the underlying **construct** that **should be measured**.
- A simple question such as **“how many children does a person have”** can be **subject to different interpretations** in some cultures.
- In households with **extended family** member's **biological children** may not be distinguished from **children of brothers or sisters** living in the same household.
- In a **disability survey**, a general question asking people whether or not they have a disability can be subject to **different interpretations** depending on the **severity of the impairment** or the **respondent's perception of disability**.
- **People with minor disabilities may perceive themselves to have no disability**.

- Unless the right screening and **filter questions** are included in the questionnaire, the answers may not fully bring out the total number of people with disabilities.

## 2) Coverage or Frame Error

- **Frame imperfections** can bias the estimates in the following ways:
- If **units are not represented** in the **frame** but should have been part of the frame
- This **result in zero probability** of selection for those units omitted from the frame.
- On the other hand, if some **units are duplicated**, this results in **overcoverage** with such units having larger probabilities of selection.
- Errors associated with the frame may, therefore, result in both **overcoverage** and **undercoverage**.
- The **latter is the most common in large-scale surveys in most African countries**.
- **Non-coverage** denotes **failure to include some sample units** of a defined survey population in the **sampling frame**. Because such units have zero probability of selection, they are effectively excluded from the survey results
- **Non-coverage is often associated with problems of incomplete frames**.
- Thus, **non-coverage** refers to the **negative errors** resulting from failure to include elements that would, under normal circumstances, belong in the sample.
- **Positive errors of overcoverage** also occur due to inclusion in the sample of elements that do not belong there.
- It should be noted that non-coverage errors differ from nonresponse.

## 3) Nonresponse error

- **Nonresponse** refers to the **failure to measure some of the sample units**. Thus, **failure to obtain observations** on some units selected for the sample.

- Results from **failure to obtain** observations on some sample units, **due to refusals, failure to locate addresses or find respondents at home** and **losses of questionnaires**.
- In summary, **the types of non-respondents** include:
  - 1) **Not-at-homes**: prospective respondents who may not be at home when enumerators visit their households.
  - 2) **Refusals**: respondents who refuse to give information for whatever reasons.
  - 3) **Not identifiable respondents**

#### 4) Measurement Error

- These errors arise from the fact that **what is observed or measured** *departs* from the **actual values** of sample units
- These errors centre on the **content** of the survey such as **definition of survey objectives**, their **transformation into usable questions**, and the **obtaining** of responses.
- These errors concern the accuracy of measurement at the level of individual units.
  - 1) For example at the initial stage wrong or **misleading definitions** and **concepts** on **frame construction** and **questionnaire design** lead to **incomplete coverage** and **varied interpretations** by different interviewers **leading to inaccuracies** in the collected data
  - 2) **Inadequate instructions to field staff** are another source of error. For some surveys instructions are **vague** and **unclear** leaving interviewers to use their **own judgment** in carrying out fieldwork.
  - 3) The interviewers themselves can be a source of error. At times, the **information collected** on a given item for all units **may be wrong**; this is mainly due to **inadequate training of field workers**.
  - 4) **Age reporting** in **Africa** is a **common measurement problem** through **age heaping** and **digital preference**. These and other examples of

measurement error may be **attributable to respondents** or **interviewers** or both. At times, there may be **interaction between the two**, which may contribute to **inflating such errors**.

- **Respondents** can introduce errors because of the following reasons:
  - 1) **Failure to understand the survey question(s)**
  - 2) **Careless** and **incorrect answers** due to, for example, **lack of adequate understanding** of the objective(s) of the survey; in particular, the respondent may not give sufficient time to think over the questions.
  - 3) **Desire to “cooperate”** by answering questions even when they do not know the correct answer
  - 4) **Deliberate inclination to give wrong answers**, for example, in surveys dealing with sensitive issues, such as income and stigmatized diseases.
  - 5) **Memory lapses** if there is a long reference period, a case in point being the collection of information of non-durable commodities in expenditure surveys.

## 5) Processing Errors

- Processing errors arise during the **data processing stage** and comprise:
  - **Editing** errors
  - **Coding** errors
  - **Data entry** errors
  - **Programming** errors, etc.

## Summary

- The various sources of non-sampling error are present right from the when the survey is being planned and designed to when data are processed and analyzed. These errors should be given due attention because they can cause huge biases in the survey results if not controlled.
- How can non-sampling errors be controlled?
  - The best is to follow the right procedures of all survey activities from planning, sample selection up to the analysis of results.

- Examples include careful and intensive training of field personnel
- Age data - collect not only age at last birthday, but also date, month and year of birth. The discrepancy will reveal any recall error that may be present in the reported age.

## **SAMPLING ERRORS**

- Sampling errors arise due to the fact that a **sample rather than the entire population is surveyed**
- These errors centre on the process of sample design
- In order to draw inferences from the sample survey about the population, probability sampling is used
- Even when the same sample design is repeated, it would result in different estimates depending on the actual units, which happened to have been selected
- Therefore, good survey practice includes calculation and estimation of errors. That is, margin of error and amount of variability

## **The Margin of Error**

- This is a **measure of the difference** between the **sample estimate** and the **actual population value**
- The **better the design**, the **less is the margin of error** and the **more precise the estimates**
- **Design** = (stratification, clustering, and proportionate sampling. etc.)

## **Example (opinion poll)**

- Votes that the candidate for a particular office are expected to get, the results could be stated that candidate A's votes are *estimated at 52%* with the error unlikely to be *more than 4%*, so that this candidate's votes are expected to fall in the range of 48%-56%.

## **SUMMARY**

- All survey data are subject to error from various sources
- The broad fundamental distinction of errors is between errors in the measurement process and errors in the estimation of population values from measurement of a sample of it.
- Non-sampling are unpredictable and not easily controlled. If not properly, controlled non-sampling errors can be more damaging than sampling error for large-scale household surveys.
- Both sampling and non-sampling errors need to be controlled and reduced to a level at which their presence does not defeat or eliminate the usefulness of the final survey results.

## **References**

Shryock, H.S and J.S Siegel (2004). *The Methods and Materials of Demography*. Academic Press: New York.

United Nations (2005). *Designing Household Survey Samples: Practical Guidelines*. New York

Zambia Statistics Agency, Ministry of Health (MOH) Zambia, and ICF. 2019. *Zambia Demographic and Health Survey 2018*. Lusaka, Zambia, and Rockville, Maryland, USA: Zambia Statistics Agency, Ministry of Health, and ICF.