Handling Nonsampling Errors—Case Salo

Maria Valaste
University of Helsinki, Finland
The Social Insurance Institution of Finland (KELA), Finland

20–24 August 2018

Introduction

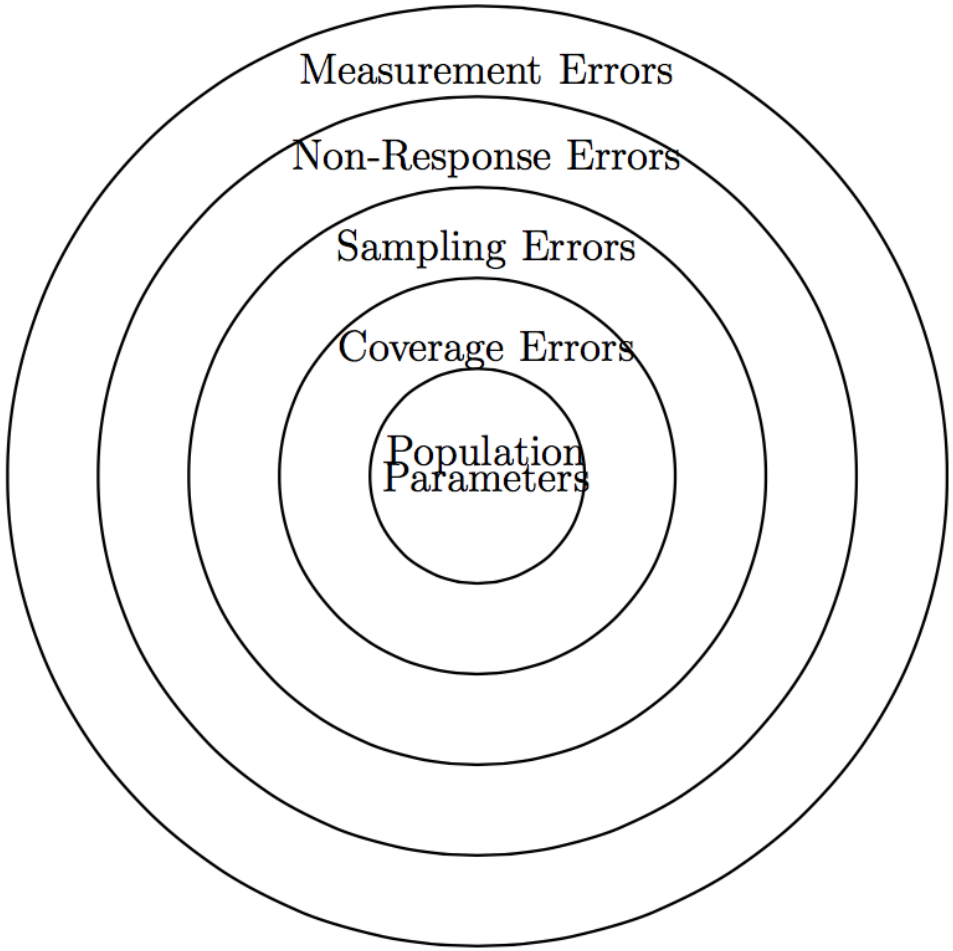
- In real life nonsampling errors are almost inevitable.
- In a perfect case the variable of interest is measured on every unit in the sample without error, so that errors in the estimates occur only because just part of the population is included in the sample. Such errors are referred to as *sampling errors*. (Thompson, 2012). In real life *nonsampling errors* may also arise.
- Groves (1989); Alwin (1991, 2007); de Leeuw et al. (2008); Groves et al. (2009) specify four sources of error in surveys:
  - coverage error,
  - sampling error,
  - nonresponse error and
  - measurement error.

Introduction (cont.)

- Most important types of nonsampling errors are nonresponse, coverage errors and measurement errors (de Leeuw et al., 2008). Lehtonen & Pahkinen (2004) also adds to this list processing errors.

Introduction (cont.)

Figure: A sequence of sources of survey errors (Alwin, 1991, 2007).



Introduction (cont.)

- Two types of coverage errors exist: undercoverage and overcoverage errors. An undercoverage error arises when some population elements are not included in the sampling frame. An overcoverage error is present when a unit from the target population appears more then once in the sampling frame.
- Sampling errors exist in the surveys because only a subset of the population elements is used to represent the population.
- A nonresponse error occurs when the survey fails to get a response to one, or possibly all, of the questions.
- A measurement error is a lack of measurement precision due to weakness in the measurement instrument.

Missing data

- Two types of missing data: unit nonresponse and item nonrespose
  - Unit nonresponse is the failure to obtain any information from an sample unit.
  - Item-nonresponse refers to the failure to obtain information for one or more questions in a survey, given that the other questions are completed. (de Leeuw et al., 2008).
- The methodologies for handling unit non-response and item non-response can differ but in both cases the reasons for missing values has to be investigated.
  - Usually indicator variable is created for unit response or item response and missingness rates and descriptive statistics are computed.
- There are basically three techniques to deal with nonresponse/missing data:
  - Weighting and reweighting
  - Analysis so that missingness has been taken into account by modelling
  - Imputation

Example: Salo

This study is a large research project "Sudden structural change – case study of Nokia-city Salo" (2013 – 2023)*

* Ylikännö & Kehusmaa, 2015

Example: Salo (cont.)

- Salo is a middle size town in South-Western Finland approximately 50 kilometers from Turku and 100 kilometers from Helsinki.
- Until 2012 the assembly factor of Nokia mobile phone company was situated in Salo. It employed more than 4 000 persons and was the largest private employer is the area.
- In the summer 2012 the factory was closed down causing Salo to become an area of sudden structural change. It has already since 2009, when some of the major subcontractors of Nokia were transferred to Asia, received millions of euros in order to minimize the negative effects of sudden structural change.
- One aim in the research project is to follow the inhabitants of Salo and their well-being for several (appr. 10) years in order find out how they cope with the sudden structural change and its (negative) effects.
- The baseline survey data of this case study was gathered from the mailed questionnaire which was distributed in spring 2013.
- The questionnaires were distributed to everyone living in Salo and representing the following birth cohorts: 1961–1963, 1971–1973, 1981–1983, and 1991–1993.

Example: Salo (cont.)

- Of the study population, 2133 subjects completed and returned the questionnaire. The response rate was 29%.
- Subjects were asked to answer to the questions about their background, educational level and main type of activity, residency, willingness to relocate, use of services, health, social well-being and income.
- The *second follow-up survey* in spring 2015 utilized the mailed questionnaire but also web survey.
- Target population was those who participated the first survey and also those who have moved to Salo after the first survey.
    - 2287 subject completed the questionnaire. 1285 subjects participated in the baseline and follow up survey and 1002 subjects were new subjects. The response rate for the follow up survey was 29%.

Example: Salo (cont.)

- Currently the third survey is planned.
  - As earlier survey rounds, this also will have challenges especially nonresponse issues.
  - How to improve the response rate? And in general, what is a lesson learned from the previous surveys?

References

D. F. Alwin (1991). Reasearch of survey quality. *Sociological Methods and Research 20*, 3–29.

D. F. Alwin (2007). *Margins of Error: a study of reliability in survey measurement.* New Jersey: Wiley.

de Leeuw, E. D., Hox, J. J. & Dillman, D. A. (2008). The cornerstones of survey research. In de Leeuw E. D., H. J. J. & D. D. A., eds., *The international handbook of survey methodology*. Erlbaum/Taylor & Francis, New York/London.

Groves, R. M. (1989). *Survey errors and survey costs*. John Wiley & Sons, New York.

Groves, R. M., Jr., F. J. F., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology*. New Jersey: John Wiley & Sons.

Risto Lehtonen and Erkki Pahkinen (2004). *Practical Methods for Design and Analysis of Complex Surveys.* Second Edition. Chichester: Wiley.

Thompson, S. K. (2012). *Sampling*. New Jersey: John Wiley & Sons.

Valaste, M. (2015). Aineisto ja menetelmät. In M. Ylikännö & S. Kehusmaa, eds., *Muuttuva Salo. Kyselytutkimus äkillisen rakennemuutoksen alueen asukkaiden hyvinvoinnista*, chap. 3. Kela, Helsinki, pp. 13–16.

Ylikännö, M. & Kehusmaa, S. (2015). *Muuttuva Salo. Kyselytutkimus äkillisen rakennemuutoksen alueen asukkaiden hyvinvoinnista*. Kela, Helsinki.