

gesis

Leibniz Institute  
for the Social Sciences



# Sampling in Practice

## GESIS Survey Guidelines

*Sabine Häder*

These slides are based on the GESIS Survey Guideline paper about open-ended questions:

Häder, Sabine (2016). Sampling in Practice. Mannheim, GESIS – Leibniz Institute for the Social Sciences (*GESIS Survey Guidelines*). doi: 10.15465/sdm-sg\_014

Please cite the slides as:

Häder, S. (2017). Slide Set: Sampling in Practice. GESIS Survey Guidelines. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences.

A complete list of all references used on these slides can be found in the above mentioned Survey Guideline paper.

# Sampling for Surveys

- What has to be done in the preparatory phase?
  - ▶ Sample surveys, be they general population surveys or surveys of specific populations, can take many different forms
  - ▶ May be telephone surveys, face-to-face surveys, postal surveys, or online surveys; Decision: time- and cost constraints or specific characteristics of the survey
  - ▶ Always first question: the availability of a suitable sampling frame
  - ▶ Furthermore: sampling frame contains supplementary characteristics?
  - ▶ Finally: specific sampling procedure has to be determined

# Sampling for Telephone Surveys 1

- For landline phone surveys:
  - ▶ Telephone book not appropriate as sampling frame
  - ▶ GESIS sampling frame or ADM sampling frame (with supplementary information for stratification)
  - ▶ How does it work? For every local network area (area code) it is possible to determine the exact 100-blocks in which at least one listed telephone number occurs. All possible sequences of digits are generated for every 100-block with at least one listed number (i.e., in the case of  $K$  blocks with at least one listed number, a total of  $K \cdot 100$  numerical sequences). A predetermined number of numerical sequences is then drawn from this set (e.g., by means of simple random sampling or stratified random sampling).
  - ▶ However: growing proportion of households can be reached only by mobile phone

## Sampling for Telephone Surveys 2

- For mobile phone surveys:
  - ▶ In 2010, 26% of the Eurobarometer respondents were “mobile only” (Germany: 12.4% in 2011)
  - ▶ Serious undercoverage! GESIS sampling frame: is constructed in such a way that all meaningful numerical sequences are generated for all valid provider dial-in numbers (e.g., 0171, 0165, 0177). In addition, these sequences are cross-checked with current information from the Internet. In this way, a sampling frame is achieved that comprises around 212.19 million numerical sequences (i.e., actual and potential phone numbers). A predetermined number of numerical sequences is then drawn from this set (e.g., by means of simple random sampling).

# Sampling for Telephone Surveys 3

- For dual frame surveys
  - ▶ Parameters required for the dual frame model

	Landline		Mobile
$M^F$	Size of the landline sampling frame	$M^C$	Size of the mobile phone sampling frame
$m^F$	Size of the landline sample	$m^C$	Size of the mobile phone sample
$k_i^F$	Number of landline phone numbers at which person $i$ can be reached	$k_i^C$	Number of mobile phone numbers at which person $i$ can be reached
$z_i$	Number of persons in the household of person $i$ who belong to the target population		

$$\pi_i \approx k_i^F \frac{m^F}{M^F} \cdot \frac{1}{z_i} + k_i^C \frac{m^C}{M^C}$$

# Sampling for Telephone Surveys 4

- Additional requirements for telephone surveys:
  - ▶ For each numerical sequence, it is advisable to make at least ten contact attempts on different days of the week and at different times of day.
  - ▶ If the target persons of a landline survey are not pre-determined, for example because they are on a list, they must be randomly selected within the contact household (e.g. last birthday method)
  - ▶ In telephone surveys, the target persons do not have equal probabilities of selection. For this reason, design weighting must be carried out when the data are being analysed. This applies to landline samples, mobile phone samples, and dual frame samples.
  - ▶ In every survey, careful documentation of the response rate is of importance for the assessment of the quality of the sample. The response rate should be computed in accordance with AAPOR Standards.

# Sampling for Postal Surveys

- Requirements for postal surveys:
  - ▶ The prerequisite for drawing a sample for a postal survey is a list of names and addresses of the persons who belong to the frame population. From this list, a predetermined number  $n$  of persons can be selected.
  - ▶ The situation is more difficult when a list of the elements of the frame population is not available. Alternative?
  - ▶ One possibility: conduct telephone screening interviews to select target persons, ask for their address and send the questionnaire by mail -> very low response rates
  - ▶ Another possibility is to distribute the questionnaires during a so-called “random walk” -> inclusion probabilities can not be determined exactly
  - ▶ Use of population register -> time and cost intensive



# Sampling for Face-to-face Surveys 1

- ADM-Design:
  - ▶ Stage 1 - Selection of areas: The entire inhabited area of the Federal Republic of Germany can be divided into around 50,000 to 60,000 small areas on the basis of official statistics. This is the finest division supported by official statistics. ADM draws a random sample from these areas; the thus determined elements are referred to as sample points. 210 sample points in Western Germany and 48 sample points in Eastern Germany have to be selected for every 1,000 interviews to be realised. These sample points are structured in such a way that they comprise, on average, 600 to 700 households.
  - ▶ Stage 2 - Selection of Households: Beginning at a certain starting address, a random walk in the sample point is conducted. The survey institutes draw up a set of random route instructions that the interviewers must follow when doing the random walk
  - ▶ Stage 3 - Selection of a target person: Kish grid or Last birthday method

## Sampling for Face-to-face Surveys 2

- Population register samples: Best practice
- Stage 1 - Sample of around 150 municipalities (stratified, proportional to population size)

As a general rule, the higher the number of municipalities, the greater is the precision of the estimators. However, the size of the sample, in particular, is a key cost- and time factor

- Stage 2 – Sample of individuals (fixed number) from the population register

Advantage: This design is self-weighting on the individuals' level.

Disadvantage: The design is clustered (to be considered in the calculation of the design effect)

# Sampling for Online Surveys

- ▶ Email Surveys
- ▶ Web Surveys (following Couper 2000):

Nonprobability Methods	Probability-Based Methods
Polls as entertainment	Intercept Surveys
Unrestricted self-selected surveys	List-based samples
Volunteer opt-in panels	Web option in mixed-mode surveys
	Pre-recruited panels of Internet users
	Pre-recruited panels of full population

gesis

Leibniz Institute  
for the Social Sciences

Member of  
*Leibniz*  
Leibniz  
Association