SSOAR
Open Access Repository
www.ssoar.info

Sampling in Practice (Version 2.0)

Häder, Sabine

Erstveröffentlichung / Primary Publication
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Häder, S. (2016). *Sampling in Practice (Version 2.0).* (GESIS Survey Guidelines). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.15465/gesis-sg_en_014

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

Terms of use:
This document is made available under a CC BY-NC-ND Licence (Attribution-Non Comercial-NoDerivatives). For more Information see:
https://creativecommons.org/licenses/by-nc-nd/4.0

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

Leibniz-Institut
für Sozialwissenschaften

GESIS Survey Guidelines

Sampling in Practice

Sabine Häder

December 2016, Version 2.0

Abstract

This contribution is devoted to the practical side of survey sampling. It demonstrates how sampling frames and sampling strategies are determined for telephone surveys, in particular, and for postal, face-to-face, and online surveys. And finally, taking the European Social Survey as an example, it deals with sampling in cross-cultural surveys.

Citation

Häder, S. (2016). Sampling in Practice. *GESIS Survey Guidelines*. Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_014

1. What is it all about?

Sample surveys, be they general population surveys or surveys of specific populations, can take many different forms. For example, they may be telephone surveys, face-to-face surveys, postal surveys, or online surveys. The main factors influencing the choice of survey mode are often time- and cost constraints or specific characteristics of the survey in question, for example the proportion of sensitive questions or the complexity of the survey instrument.

Irrespective of the survey mode chosen, the first question that must always be addressed when sampling in practice relates to the availability of a suitable sampling frame. Moreover, it must be clarified whether the sampling frame contains supplementary characteristics that could be used as stratification variables. And finally, the specific sampling procedure is determined.

The following chapters provides concrete examples of sample surveys conducted in the various survey modes.

2. How do telephone samples work?

Between 1998 and 2008, telephone surveys were the most frequently used survey mode in Germany. In the mid-1990s, GESIS developed an adequate sampling frame for landline numbers (the "Gabler-Häder design"). However, a trend has since emerged that makes the sole use of this frame appear inadequate to cover the entire population of private households – namely, the fact that a growing proportion of households can be reached only by mobile phone. These households do not have a positive chance of inclusion in landline surveys. This can give rise to systematic bias in the samples because landline households and mobile-only households differ in terms of characteristics that are of relevance to social research. To avoid further excluding the young, mobile population group with a low level of formal education, who are frequently underrepresented in telephone surveys anyway, it became necessary to consider integrating mobile connections into telephone samples. Results of research on sampling strategy, response quality, and people's willingness to participate in mobile surveys will be presented in what follows.

2.1 Is the telephone book an adequate sampling frame for landline telephone surveys of the general population?

Up to 1992, the telephone book was an acceptable sampling frame for landline telephone surveys in Western Germany because, with only very few exceptions, telephone subscribers were obliged to allow their telephone numbers to be published in the telephone directory. This obligation was abolished in 1992, and the proportion of so-called „non-pubs," that is, unlisted telephone subscribers, has increased sharply in recent years. By now, some 50 percent of landline numbers assigned to private households are unlisted.

Table 1: Ratio of assigned landline numbers to listed landline numbers in Germany, 2009

Total private households in 2009 in millions	Listed residential telephone numbers in millions	Residential telephone numbers assigned in millions	Proportion of private households listed	Proportion of assigned telephone numbers listed
39.331	22.658	536.221	57.6 %	4.2 %

Source: Heckel, 2011, p. 105

It should be noted that the proportion of non-pubs may be significantly higher in urban areas, especially in cities. By contrast, in rural areas, especially in Western Germany, the proportion of listed landline numbers is higher.

Drawing a sample from the telephone book would lead to biased samples because listed subscribers differ from unlisted subscribers in terms of key characteristics (e.g., compared to listed telephone subscribers, a greater proportion of unlisted subscribers live in cities rather than in rural areas, are younger, have a higher level of formal education, and are divorced). For this reason, the telephone book is no longer suitable for use as a sampling frame for surveys of the general population or larger municipalities. However, this does not mean that it can no longer be used in sample planning at all. It is still useful for constructing migrant samples (see Humpert/Schneiderheinze, 2002) or samples of enterprises or institutions, and for drawing samples for individual rural regions in Western Germany.

2.2 How does the "Gabler-Häder design" for the selection of landline samples work?

This sampling frame is used when a complete list of telephone numbers is not available.

In order to include both listed and unlisted telephone numbers in the sample, it would be conceivable to use random digit dialling (RDD), a method that originated in the USA. To apply this method, the range between the lowest and the highest listed telephone number would have to be determined for all 5,200 local network areas (area codes) in Germany (or only for the area codes of interest). This range would comprise the set of numerical sequences that could conceivably be telephone numbers. From this set, the sample would be drawn. For example: In a CD-ROM telephone directory published in January 2014, the lowest residential telephone number in the local network area "06321" was 2001 and the highest was 95600100. The total number of residential listings in this local network area was 12868. Let us assume that there were just as many unlisted numbers in this local network area as there were listed numbers. If a sample were drawn from the full range between the lowest and the highest listed number, only 0.027% of the numerical sequences selected would actually be activated telephone numbers. it is clear, therefore, that because of the complicated numbering system in Germany, this approach is not feasible as it is far too time- and cost-intensive.

Another solution to the problem of also including unlisted telephone subscribers in the sample is the so-called "randomise last digits" (RLD) method, which also originated in the USA. Here, numbers are selected from the telephone book and then their last two digits are replaced with randomly generated digits. However, under this procedure, the inclusion probabilities, that is, the chances that the telephone numbers will be selected into the sample, are not equal. Rather, they depend on how many other numbers in the same "100-block" are listed. A "100-block" is defined as the stem of a telephone number that remains after the last two digits have been removed. It comprises the set of all the different numerical sequences that can be generated by randomly replacing the last two digits – that is, exactly 100. For example, the telephone number 5129815 is in block 51298xx, which comprises all sequences of digits from 5129800 to 5129899. To correct for the unequal inclusion probabilities under the randomise last digits method, the data sets obtained in the interviews would have to be retrospectively weighted by the inverse of the number of listed numbers in that block, which would be an extremely time-intensive operation. For example, a data set from an interview with a participant whose block contains a total of 40 listed numbers would be multiplied by the factor 1/40. Another participant, in whose block only two listed numbers occur, would be assigned the factor 1/2.

A method developed at GESIS in the mid-1990s gives listed and unlisted telephone numbers an equal chance of inclusion in the sample (see Häder & Gabler, 1998). This method will be briefly described in what follows.

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.

There are often large gaps between these 100-blocks. These gaps are the main reason for the low efficiency of the application of pure random digit dialling. In our model, we assume that no unlisted numbers occur in these gaps. In the past, this model was adequately realistic. Of late, however, it transpires that, especially because of the presence of additional providers in the market, new blocks are being activated in which none of the numbers are listed in the telephone book. These numbers are now missing from our sampling frame. Hence, enhanced efforts are needed to incorporate information from the Federal Network Agency (*Bundesnetzagentur*; Sand 2014a, S.13, Sand 2015).

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*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).

This is by now the standard sampling procedure for telephone surveys in Germany. It was further developed by the Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM; the association that represents the interests of private-sector market and social research agencies in Germany) insofar as regional stratifications are now also possible (Heckel 2002, von der Heyde 2002).

2.3 How are mobile numbers handled in telephone samples?

Surveys conducted by landline telephone can be regarded as methodologically well researched with regard to sampling and data quality. However, landline samples no longer provide adequate coverage of the general population in Germany.

For some years now, mobile phone usage has been widespread throughout Europe. In 2010, 87% of the respondents of the Eurobarometer had a mobile phone (Eurobarometer, 2011: 6). What is striking is the trend whereby mobile phones are becoming the only means of telephony in increasingly broad sections of the population – in other words, these households are doing without a landline telephone. In 2010, 26% of the Eurobarometer respondents were "mobile only". However, because "mobile only" and "mixed use" respondents differ with regard to socio-demographic characteristics, the non-inclusion of the mobile only group constitutes undercoverage, which may lead to systematic bias of the estimators.

In Germany, mobile phone penetration increased from 96% in 2005 to 137% in the third quarter of 2011. In other words, a large number of people had several mobile phones or SIM cards. In 2011, 83% of the German population had (at least) one mobile phone; 12.4% belonged to the "mobile only" group (ADM 2012). Mobile phone penetration in Germany is thus well below the European average, which is probably due to the fact that landline telephone connections are still the preferred means of obtaining comfortable Internet access. However, in Germany, too, the mobile only group differs from the rest of the population in terms of socio-demographic characteristics: They tend more to be young and male; a disproportionately high percentage hail from Eastern Germany; and they tend to have a lower level of educational attainment. Hence, only 16.8% of the mobile only respondents of the TNS-Infratest F2F-Bus 2010 (n=30,000) held a general higher education entrance qualification (*Abitur*), whereas the national average is 24.6%. Some 55,4% of the mobile only respondents in this survey programme were male, whereas the national figure is only 48.4%.

As international experience shows, it is to be expected that the number of mobile only users will continue to grow. It could therefore be concluded that telephone surveys should be conducted only by mobile phone in future. However, it must be taken into account that 18.6% of the population in Germany can be reached only by landline telephone (ADM 2012). These tend to be older people living in rural areas in Western Germany. Thus, the only way of avoiding considerable undercoverage is to use both telephone modes in parallel – in other words, to implement a dual-frame approach.

Only around 2% of mobile phone numbers are listed in the telephone book – and most of them are business listings. Here, the 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).

2.4 How do dual frame samples for telephone surveys work?

As we argued above, using only one of the two sampling frames (landline or mobile) would mean that all the elements of the population would not be covered. Therefore, if a second sampling frame is available that contains a large proportion of the missing elements (e.g. mobile only users), so that both frames together (almost) cover the total population, the samples should be independently drawn from both sampling frames. The only problem is that both frames overlap. This raises the question of how this overlap should be dealt with in order to achieve almost unbiased estimators. A relatively elegant and simple way of doing this is to determine the inclusion probabilities of the elements of the population and to apply the Horvitz-Thompson (HT) estimator for the estimation of the total of a variable of interest.

For a general formula for inclusion probabilities (see Gabler /Ayhan 2007) the relevant parameters must first be defined (see Overview 1):

Overview 1. 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		

To simplify the formulae that are then to be derived, the following basic assumption should be made:

The probability that two (not necessarily different) members of the same household will be selected from different frames is negligible.

This assumption may be problematic only in small regional samples.

Because of the basic assumption, we neglect samples that contain a household with several different landline numbers. Moreover, the probability of selecting a person from both the landline frame and the mobile frame is negligible, so that it can also be neglected.

$$\pi_i^{F \cap C} = \pi_i^F \pi_i^C \approx 0$$

For the inclusion probability of the person *i*, we therefore obtain in a good approximation

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

This approach has become standard practice in Germany. Experience in this regard was gathered in the studies CELLA1 and CELLA2 (GESIS and TU Dresden), which were funded by the German Research Foundation (DFG).

The size of the respective sampling frames – the landline frame, M^F, and the mobile frame, M^C – currently amounts to:

M^F = 158.88 million

M^C = 212.19 million.

For the parameters m^C and m^F, the actual number of numbers selected per mode should be entered, irrespective of whether they were actually called or not. Moreover, care should be taken to ensure that similar samples (with or without unlisted numbers, respectively) are used. In the case of landline samples, unlisted numbers can be eliminated via prediallers. HLR Lookup (Sand 2014a) can be used to determine the status of mobile phone numbers (Struminskaya 2009). The other parameters must be collected during the interview (Gabler et al. 2013).

German market and social researchers are currently discussing which ratio of landline to mobile sample size is to be recommended. Simulations carried out by the ADM suggest that, in nationwide studies with sufficiently large samples, a ratio of 70:30 yields stable estimators for mobile only users and should thus prove adequate (see ADM 2012).

2.5 How does telephone sampling work when a sampling frame is available?

Sometimes, lists, or frames, are available that represent the target population "acceptably" in the sense that the difference between the target population and the frame population is negligible. This may be the case, for example, for members of certain associations, organisations, and institutions, or for the customers of a firm.

The following example – a customer satisfaction survey commissioned by a furniture store – illustrates the usefulness of such lists as a sampling frame for telephone surveys. It is conceivable that this furniture store will have a list of its customers' names and addresses, because every time a large item of furniture is purchased, these details are registered in order to be able to inform the customer by telephone of the date and time of delivery. If the furniture store wishes to have a survey of customer satisfaction conducted, it can make the aforementioned particulars available to a market research institute. This is in line with data protection regulations, provided the information is used exclusively for the purpose of measuring customer satisfaction. By contrast, it is not legally permissible to link the measurement of customer satisfaction with advertising campaigns, or the like. It must now be decided how far back the registration (or the last update) may be. This decision relates not only to the up-to-dateness of the particulars but also to substantive considerations, for example the customer's ability to recall the purchase of the furniture. Hence, it is conceivable that all particulars registered in the last two years would be used. The frame population would then comprise those customers who purchased a large item of furniture from this furniture store in the last two years and, when doing so, gave their telephone number. If the list also contained details of the price of the purchased furniture and/or how often the person has purchased furniture from the store in the past, these details could be used as stratification characteristics. Telephone numbers could then be randomly drawn from the cells formed by stratification, and the customers in question could be contacted by phone.

If, in addition to the names and telephone numbers, the list also contained the addresses of the customers, it would be possible to prepare the selected customers for the planned study by sending them an advance letter. However, the extent to which such contacts prior to the actual telephone interview contribute to increased willingness to participate on the part of potential respondents has not yet been clearly established (see, e.g., Hüfken 2000). It is also equally unclear to date whether it is beneficial to send the questionnaire to the interviewees in advance if their postal addresses are known. Especially in the case of complex questions, this makes sense in theory. However, there are no clear findings that support an increase in the response rate as a result (see Friedrichs 2000).

2.6 What should be taken into consideration when using telephone samples?

Irrespective of the way in which the actual sample for a telephone survey is constructed, a number of generally applicable tips can be given for its implementation.

- Telephone samples should be "shuffled" before the survey begins if they are initially sorted according to the length of the area code, the provider, and/or subscriber number, regions, or business units. It is then advisable to process the numbers "in packs", for example in sets of 100 numbers. In this way, one can avoid a situation where all the numbers are rung up at once and a large number of interviews take place on the first contact. Target persons who are harder to reach would then have less chance of making it into the net sample, and the sample would be biased with regard to the characteristic of "reachability" as a result. For this reason, one set of numerical sequences should first be processed completely, so that if a person cannot be contacted, the maximum number of contact attempts can be carried out before a new set is begun. "Shuffling" the sample before the survey begins is important especially because, in this way, it is not always necessary to process all the "packs" in the gross sample in order to achieve the targeted number of interviews.

- For each numerical sequence, it is advisable to make at least eight contact attempts on different days of the week and at different times of day. In the guidelines issued by the Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM; see http://www.adm-ev.de/homepage.html) it is assumed that contact attempts made between the hours of 9 a.m. and 9 p.m. do not constitute a violation of privacy.
- 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. There are several standard procedures for this (see Lavrakas 1993). The "last birthday method" has proved particularly viable. Naturally, it is not necessary to select a target person if the aim of the study is to collect information about the household that each household member is equally able to provide.
- 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 with the help of the AAPOR Response Rate Calculator (http://www.aapor.org/Response_Rates_An_Overview1.htm#.U0PvGhBI85Q).

3. How are samples for postal surveys drawn?

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. If the list is computer-based, the selection can be carried out digitally by assigning every data set (every person) a random number, sorting the data sets by the size of these random numbers, and selecting the first *n* data sets. If, on the other hand, the list is available only in printed form, systematic random sampling is recommended. This entails randomly determining the first sample element to be drawn as a number between 1 and the sampling interval. The sampling interval is computed by dividing the size of the sampling frame by the desired size of the sample. Beginning at the randomly determined starting point, the next element is determined by adding the sampling interval to the number of the first element drawn.

For example, the sample for a project aimed at exploring the intellectual history of the tropical rainforest was drawn according to this principle. Two hundred geography teachers in the school district of Stuttgart were randomly selected from the address book of the Philological Association (*Philologenverband*) and surveyed by post.

In another study, which was devoted to determining the noise disturbance experienced by the local residents in the vicinity of the EXPO 2000 exhibition in Hanover, it was first determined how many residents in the targeted age group lived in the individual streets adjacent to the EXPO site. Then, a sample of potential respondents proportionate to the number of street residents in the targeted age group was selected from the population register of the municipality of Hanover, and a postal survey of the selected persons was conducted.

The situation is more difficult when a list of the elements of the frame population is not available. In this case, an alternative solution must be found. The disadvantage of such alternative solutions is that the inclusion probabilities cannot be determined exactly and the response rates are very low.

One possibility is to conduct telephone screening interviews to select target persons. To this end, telephone numbers are generated (e.g., following the Gabler-Häder design) and the corresponding households are phoned. A target person must then be selected within the household (normally using the so-called "last birthday method") and asked for their address. The questionnaire is then sent to the target person. Unfortunately, the response rates achieved with this method are very low. Moreover, the procedure is relatively cost- and time intensive. Another possibility is to distribute the questionnaires during a so-called "random walk". Following a set of random route instructions, the interviewer puts the questionnaires into letterboxes or delivers them personally to the household. It is problematic when the interviewer does not manage to select the respondent within the household. In this case, there must be a corresponding instruction on the questionnaire and it must be assumed that the target household will independently select the target person. The random route procedure is recommended for use only in surveys in smaller regions because it is otherwise too time intensive. A further problematic aspect of this design should also be noted, namely, that the inclusion probabilities of the potential respondents cannot be determined exactly.

4. How are samples for face-to-face surveys drawn?

The two main sampling designs used in face-to-face population surveys based on random samples in Germany are presented in what follows:

4.1 The ADM Sampling System

For face-to-face population surveys in market, media, and social research, the Arbeitsgemeinschaft Deutscher Markt- und Sozialforschungsinstitute (ADM), the association that represents the interests of private-sector market and social research agencies in Germany, employs, as a rule, a specially developed design, the so-called "ADM design," which is a three-stage stratified random sampling 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. Moreover, in the case of municipalities with over 10,000 inhabitants, these areas can be delimited with the help of digitised street maps. ADM draws a random sample from these areas; the thus determined elements are referred to as sample points. From 1997 to 2003, the approximately 80,000 constituencies in the Federal Republic of Germany formed the basis for the determination of the sample points. Since 2003, however, a refined division of the sample points into street segments has been undertaken.

The ADM uses the smallest administrative unit areas available – down to city-block level – to generate the sample points. Official inner-city divisions and attribute data can be systematically used by now. These data include urban district data, entries in telephone directories, etc.

Since 2003, following tried-and-tested convention, 210 sample points in Western Germany and (since the territorial reform in 1996) 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.

In order to be able to conduct interviews in the sample points, a so-called network is needed – that is, a certain number of interviewers who are distributed across these areas.

The model developed above represents a stratified sampling design. The stratification criteria in the ADM design are the federal states *(Bundeslaender)*, the administrative regions (*Regierungsbezirke*), and the districts (*Kreise*) in combination with a classification of the municipalities. "The cells in the sampling system are formed from the rural districts [*Landkreise*]/towns or cities not attached to a district [*kreisefreie Städte*] and the municipality size classes [*Gemeindegrößenklassen*]. The municipality size classes are used in an extended form, divided into groups of 10 ... The municipalities that do not belong to a metropolitan region are assigned within a stratum (rural districts) according to their actual (political) size" (ADM 1999:86f.; our translation)

Stage 2: Selection of the target household

The aim of the second stage of sample selection is to select the target households in the sample points in which the interviews are to be conducted at a later date. 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 (Häder 2014).

Stage 3: Selection of the target person

The third stage entails finding the actual target person in the target households that were selected in the second stage. In survey practice, two different methods are used to do so. In the first variant, a list is drawn up of all the household members who belong to the target population and who would therefore be eligible in principle to participate in the survey. These persons should be entered into the list in order of age. The random numbers on the address list sheet should then be used to determine which person in the household should be interviewed. In the second variant, the person who had the most recent birthday or who will have the next birthday in the household is requested to participate in the survey.

Under the ADM design, each household theoretically has an equal chance of selection. By contrast, the chance of selection of the individual household members depends on the number of persons in the household who belong to the target population. Therefore, design weighting must be applied when the data are analysed.

In practice, a number of problems occur when the ADM design is applied. Although they cannot be discussed in detail here, they will at least be mentioned:

- The degree of personal discretion that the interviewers have when selecting the households is relatively large, especially in the case of the random route procedure. It is somewhat more limited when the random walk and the interviews are conducted separately and the target households are randomly selected by the surveying institute from the list of households compiled by the interviewer during the random walk (*Adressvorlaufverfahren*; see Häder/Häder 1997).
- Compared, to a population register sample, for example, the institute has very limited possibilities of controlling the work of the interviewers when they are selecting households/persons.
- A list of households is not available initially but must first be generated by means of a random walk in the sample point. This in turn implies a relatively high degree of spatial clustering, which does not occur in population register samples, for example.

- It is sometimes difficult to comply with the definition of the target population. For example, it cannot be readily decided how persons with foreign-sounding names should be dealt with in a survey whose target population is "German citizens".
- The design weighting in the case of sampling plans that cannot be fully implemented because of nonresponse may possibly lead to a bias of the estimator (Rothe 1994, Rothe/Wiedenbeck 1994).
- Because of the starting-address provision in the random route procedure, households located in the lower sections of long roads have no chance of selection.

Despite these problems, the ADM design is frequently employed in market, media, and social research because it is considerably less cost- and time intensive than the higher quality population register sample (see Häder & Häder 1997). Examples of prominent surveys that have used the ADM Design are the German General Social Survey (ALLBUS) up to 1994 and once again in 1998, and the recruitment of the German Internet Panel.

4.2 Population register samples

At present, the best practice design for population surveys in Germany is the population register sample (aka register sample), which is a two-stage stratified random sampling procedure.

As there is no central population register in Germany but rather municipality-level registers, a sample of municipalities must be drawn in the first stage. Simulations have revealed that relatively stable distributions for many variables of interest can be achieved with a sample of around 150 municipalities. This means that the number of municipalities (i.e., primary sampling units = PSUs) drawn should not be significantly lower. However, this also implies that there is no fixed master plan for drawing population register samples. As a general rule, the higher the number of PSUs, the greater is the precision of the estimators. However, the size of the sample, in particular, is a key cost- and time factor (vgl. Albers 1997).

By way of example, the construction of the population register sample for the German sub-study of the European Social Survey (ESS) 2003 will be presented in what follows.

The target population of this survey was all persons aged 15 years and older resident within private households in Germany. A disproportionate sampling approach was chosen for Eastern and Western Germany insofar as 1,000 interviews were to be conducted in Eastern Germany and 2,000 in Western Germany. The samples for the two regions were independently generated.

First, the municipalities were stratified following the BIK classification according to district (*Kreis*) and municipality size class (*Gemeindegrößenklasse*). This yielded 1,085 non-empty stratification cells in Western Germany and 435 non-empty stratification cells in Eastern Germany.

In the first sampling stage, 100 municipalities in Western Germany and 50 municipalities in Eastern Germany were selected from these strata proportionally to the population aged 15 years and older. The number of municipalities that were ultimately drawn from each stratum was specified by controlled rounding following Cox (1987). The number of sample points amounted to 108 in Western Germany and 55 in Eastern Germany (a number of large municipalities were represented in the sample by more than one sample point).

In the second sampling stage, a systematic random sample comprising a fixed number of persons was drawn for each sample point from the registers of the respective sample points. The specification of the size of the gross sample was determined by the expected rate of non-eligibles, and the expected response rate.

5. How are samples recruited for online surveys?

It should be noted that online surveys can naturally address only those who have access to the respective media. However, because Internet users cannot be regarded as a random subsample, representative population surveys are not (yet) possible.

E-mail surveys

If a list of e-mail addresses is available, it can be used to draw a random sample. This can be done by means of simple random sampling, systematic random sampling, or stratified random sampling. Unless the questionnaire is exceptionally short, it should be sent as an attachment. The problem when returning the questionnaire is that the respondent may have concerns about the anonymity of his data because the e-mail sender address is given. It is therefore important to emphasise that the questionnaire data will be handled separately from the respondent's personal data, for example his e-mail address. Moreover, the rules of netiquette should be observed. Thus, it is advisable to send an advance e-mail. The questionnaire should not be dispatched until the potential respondent has confirmed his willingness to participate in the survey.

Web surveys

In the case of sampling in web surveys, a distinction must first be made between designs in which researchers do not select any of the interviewees themselves and designs in which researchers at least attempt to influence the selection of the interviewees. The first category includes surveys in which potential respondents are targeted by means of

- banners (hyperlinks referring to the website of the survey)
- newsletters, or
- news groups and mailing lists (for target-group-oriented surveys).

The disadvantage of these designs is complete self-selection on the part of the potential target persons. At the same time, low response rates are to be expected.

With the help of the following designs, researchers at least attempt to select the target persons:

- List-based selection. The potential respondents receive an e-mail requesting them to participate and providing a web link to the survey.
- Pop-up or intercept. Every n^{th} visitor to a website is requested to participate in the survey. The problem with this design is that it is possible to determine only the inclusion probabilities for the visit to the site but not those for the visitor.
- Recruitment via another survey medium (e.g., telephone screening). Here, too, the participants in the actual survey are initially Internet users only. In the German Internet Panel, persons who are willing to participate but who do not have a computer are equipped with an Internet access. By contrast, the GESIS Panel takes a mixed-methods approach and offers these persons the option of taking part in a postal survey. Because screening is an extremely time-intensive procedure, the effort is worthwhile only in the case of repeated surveys, that is, for the recruitment of access panels, as the response rate is, as a rule, very low.

The following overview shows the division of sampling methods for online surveys into probability-based and nonprobability methods following Couper (2000):

Overview 2: Probability-based and nonprobability sampling methods for online surveys

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

Overall, a satisfactory solution with regard to random sampling procedures for online surveys has yet to be found. The main problem is that none of the aforementioned methods enables the inclusion probabilities to be determined exactly. Moreover, it is not usually possible to increase the response rate by sending people a reminder letter. Hence, most online surveys are characterised by a high degree of self-selectivity. This frequently leads to systematic deviations between the participants in the study and the target population of Internet users. Furthermore, should adjustment weighting be necessary, the absence of distributions of meaningful adjustment variables for the target population is problematic.

6. How are samples drawn for cross-cultural surveys?

Sampling designs for cross-cultural or multi-population surveys should meet the requirement of generating equivalent samples for the various participating countries. At the same time, the circumstances in the individual countries or cultures with regard to experiences, traditions, the availability of sampling frames, and costs must be taken into consideration.

Thus, the first question that arises is what form an optimal sampling strategy should take. There are quite conflicting views on this in the profession. For example, the guiding principle for the Eurobarometer is to use the same sampling design for all participating countries. Other cross-nationally comparative studies (e.g., previous ISSP surveys) have used both random sampling and quota designs.

The following general advice on drawing cross-cultural samples can be found in Kish (1994: 173): "Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of choice is particularly advisable for multinational comparisons, because the sampling resources differ greatly between countries. All this flexibility assumes probability selection methods: known probabilities of selection for all population elements." Häder and Gabler (2003) added that if equivalent estimators are used, this results in high comparability of results (see also Häder und Lynn 2007). This strategy is employed for the European Social Survey (ESS) and it has also been used for other European studies, for example the European Values Study 2008.

6.1 Sampling frame

The most important point when seeking the best possible sampling design in the ESS countries is the determination of the respective sampling frames. Here, it is a question of finding lists that cover the defined target population as fully as possible. In other words, they should not contain any elements that do not belong to the target population (overcoverage), nor should any elements of the target population be missing (undercoverage). The target population of the ESS is defined as follows: "The survey will be representative of all persons aged 15 and over – no upper age limit – resident within private households in each country, regardless of their nationality, citizenship or language."

The following types of lists are used to draw samples for the ESS:

a) Countries with reliable lists of *residents* that are available for social research, for example Sweden, Finland, and Norway. These countries have central population registers from which samples for the ESS can be drawn. Moreover, additional information (e.g., age and nationality) about the individuals is usually available. This is helpful both for monitoring the interviewers and for nonresponse analyses.

b) Countries with reliable lists of *households* that are available for social research, for example Cyprus. In ESS Round 3, Cyprus used a list of households from the census and a supplementary list of consumers provided by the Electricity Authority of Cyprus (EAC).

c) Countries with reliable lists of *addresses* that are available for social research, e.g. the list of postal addresses "PTT-afgiftenpuntenbestand" in the Netherlands or the "Postcode Address File" in the United Kingdom.

And finally, there are participating countries who have not succeeded in finding suitable lists from which to draw samples, or who have excellent lists that are not, however, available for social research. Portugal, for example, has a central population register, but legal provisions preclude the country's statistical office from granting access to it. A similar situation exists in Austria. These countries are therefore obliged to have recourse to area sampling in combination with a random route procedure. However, the Sampling Expert Panel of the ESS makes sure that these countries observe the principle that the actual interviews should be conducted separately from the recording of the sampled addresses during the random walk. The Sampling Expert Panel's declared aim is to have these countries replace the random route procedure with sampling from lists as soon as possible because there are doubts in the statistical profession as to the extent to which random route techniques can be judged to be "strictly random" (Lyberg, evaluation of IALS DATE).

6.2 Sampling design and Sample Design Data File

Depending on the sampling frame available, the sampling designs are then determined. These designs vary considerably across the countries participating in the ESS. They range from simple random sampling (Finland, ESS1–6) to four-stage stratified and clustered designs (Russia, ESS5). The designs are documented on a standardised *sampling sign-off form*. The following categories are captured for each country:

- Target population, population coverage
- Sampling frame
- Sampling design
- Inclusion probabilities (formula)
- Design effects
- Target response rate
- Sample size
- Special features of the design (e.g., oversampling)

The basic principle when choosing a sampling plan is that the sample must be random – that is, all inclusion probabilities must be known and documented. A *Sample Design Data File* is used to record the inclusion probabilities of each sampling stage and other information on the sampling design. The information from this data set enables users to understand all the details of the sampling procedure, compute design weights and design effects, and estimate variance. The Sample Design Data Files of those countries whose data protection regulations do not prohibit the publication of these data are accessible to the public (http://ess.nsd.uib.no/).

The following screen shot is an excerpt from the French Sample Design Data File for ESS Round 3 (see Figure 1):

	CNTRY	ROUND	IDNO	PROB1	PROB2	PROB3	PROB4	PSU	SAMPPOINT	STRATEX1	STRATIM1	OUTCOME
1	FR	3	68000426	0,000041372282988	0,019586507072905	0,250000000000000	.	45	.	43	.	1
2	FR	3	68000429	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
3	FR	3	68001671	0,000041372282988	0,019586507072905	1,000000000000000	.	45	.	43	.	1
4	FR	3	68001860	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
5	FR	3	68001878	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
6	FR	3	68002037	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
7	FR	3	68002733	0,000041372282988	0,019586507072905	0,333333333333333	.	45	.	43	.	1
8	FR	3	68002997	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
9	FR	3	68003906	0,000041372282988	0,019586507072905	0,333333333333333	.	45	.	43	.	1
10	FR	3	68004503	0,000041372282988	0,019586507072905	0,500000000000000	.	45	.	43	.	1
11	FR	3	68007299	0,000041372282988	0,019586507072905	0,500000000000000	.	45	.	43	.	1
12	FR	3	68008502	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
13	FR	3	68009867	0,000041372282988	0,019586507072905	0,500000000000000	.	45	.	43	.	1
14	FR	3	68010377	0,000041372282988	0,019586507072905	1,000000000000000	.	45	.	43	.	1
15	FR	3	68010779	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
16	FR	3	68011601	0,000041372282988	0,019586507072905	0,250000000000000	.	45	.	43	.	1
17	FR	3	68012390	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
18	FR	3	68012891	0,000041372282988	0,019586507072905	.	.	45	.	43	.	2
19	FR	3	68001389	0,000130644217580	0,005303476723630	.	.	119	.	44	.	2
20	FR	3	68002439	0,000130644217580	0,005303476723630	.	.	119	.	44	.	2
21	FR	3	68002856	0,000130644217580	0,005303476723630	0,500000000000000	.	119	.	44	.	1
22	FR	3	68003168	0,000130644217580	0,005303476723630	.	.	119	.	44	.	2
23	FR	3	68003630	0,000130644217580	0,005303476723630	.	.	119	.	44	.	2
24	FR	3	68003633	0,000130644217580	0,005303476723630	0,333333333333333	.	119	.	44	.	2
25	FR	3	68007302	0,000130644217580	0,005303476723630	.	.	119	.	44	.	2

Figure 1. Excerpt from the French Sample Design Data File for ESS Round 3

6.3 Design weights

Unequal probabilities of selection that occur during sampling must be corrected for in the estimation. The Sampling Expert Panel of the ESS computes the necessary design weights for each country on the basis of the information from the respective Sample Design Data Files about the selection probabilities of the individual sampling stages. In the French design (see Figure 1), for example, the product of PROB1 and PROB2 is constant; only PROB3 varies. PROB3 is the individual's probability of selection in the third stage. Because this probability cannot be determined in advance but rather only when the household is contacted, this information is to be provided only for the net sample. In this case, the probability of selection is inversely proportional to the number of target persons in the household – that is, 1 in the case of one person, 0.5 in the case of 2 persons, 0.33 in the case of three persons, etc.

The total design weight is $w_i = 1/(PROB1_i \cdot PROB2_i \cdot PROB3_i)$. It is then standardised in such a way that the sum of the weights corresponds to the size of the sample.

Table 2 shows the effect of design weighting for data from ESS Round 3. In the case of Austria, Portugal, and Russia, the proportion of single-person households would be considerably underestimated without the necessary design weighting. In the case of Finland and Sweden, by contrast, the design weights amount to 1 because all elements had an equal probability of selection. In the case of the variable V2 (persons who spend three hours and more per week watching television) the difference between the weighted and the unweighted data is not that large because this variable has less to do with the household size, which has, after all, been taken into account in the weighting.

Table 2: Weighted and unweighted estimators for the variables “share of single-person households” (%) and “television viewing, three hours and more per week” (%) for five countries

	Austria	Portugal	Russia	Finland	Sweden
V1, unweighted	8,8	7,6	10,8	23,4	21,1
V1, weighted	21,2	17,2	23,7	23,4	21,1
V2, unweighted	13,7	22,2	23,2	14,5	9,8
V2,weighted	12,5	21,5	22,8	14,5	9,8

V1 = Single-person households (%)

V2 = 3 hours+ of TV per week (%)

7. References

Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V / A.G.M.A. e.V.. (1999): *Stichprobenverfahren in der Umfrageforschung*. Opladen: Leske + Budrich.

Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V./ (2012). ADM-Forschungsprojekt „Dual-Frame-Ansätze“. Research report. Accessed 15.4.2014.

https://www.adm-ev.de/forschungsprojekte/

Albers, I. (1997): Einwohnermelderegister-Stichproben in der Praxis. Ein Erfahrungsbericht. In S. Gabler & J. H. P. Hoffmeyer-Zlotnik (Eds.), *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*(4), 464–494.

Cox, L. H. (1987): A constructive procedure for unbiased controlled rounding. *JASA*, Vol. 82, Issue 398: 520-524.

Eurobarometer (2011). Special Eurobarometer 356: Roaming im 2010. Accessed 4.6.2016. http://ec.europa.eu/public_opinion/archives/ebs/ebs_356_en.pdf

Friedrichs, J. (2000): Effekte des Versandes des Fragebogens auf die Antwortqualität bei einer telefonischen Befragung. In: V. Hüfken (Ed.), *Methoden in Telefonumfragen*. Opladen: Westdeutscher Verlag.

Gabler, S., & Ö. Ayhan (2007). Gewichtung bei Erhebungen im Festnetz und über Mobilfunk: Ein Dual Frame Ansatz. In: S. Gabler & S. Häder (Eds.), *Mobilfunktelefonie – Eine Herausforderung für die Umfrageforschung*. Mannheim: GESIS, pp. 39–45.

Gabler, S., S. Häder, I. Lehnhoff, & E. Mardian (2012). Weighting for unequal inclusion probabilities and nonresponse in dual frame telephone surveys. pp. 147-167. In S. Häder, M. Häder & M. Kühne (Eds.), *Telephone surveys in Europe*. Research and Practice. Heidelberg: Springer.

Häder, M. (2014). *Empirische Sozialforschung. Eine Einführung*. 3rd ed. Wiesbaden: VS Verlag für Sozialwissenschaften.

Häder, M., & Häder, S. (1997). Adreßvorlaufverfahren: Möglichkeiten und Grenzen. In S.Gabler & J. H. P. Hoffmeyer-Zlotnik (Eds.), *Stichproben in der Umfragepraxis*. Opladen: Westdeutscher Verlag.

Häder, S. & S. Gabler (1998). Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In S. Gabler, S. Häder & J.H.P. Hoffmeyer-Zlotnik (Eds.), *Telefonstichproben in Deutschland*. Opladen: Westdeutscher Verlag.

Häder, S., & S. Gabler (2003). Sampling and estimation. In J. Harkness, F. van de Vijver & P. Mohler (Eds.), *Cross-cultural survey methods*. New York: Wiley.

Häder, S., & P. Lynn (2007). How representative can a multi-nation survey be?. In: R. Jowell et al (Eds), *Measuring attitudes cross-nationally*. London: Sage Publications.

Heckel, C. (2002). Erstellung der ADM-Telefonauswahlgrundlage. In S. Gabler & S. Häder (Eds.), *Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland*. Münster: Waxmann Verlag.

Heckel, C. (2011). Auswahlrahmen Dual Frame. Paper delivered at the meeting of the working group MOBILSAMPLE on 21 October 2011 at the TU Dresden.

Hüfken, V. (2000): Kontaktierung bei Telefonumfragen. Auswirkungen auf das Kooperations- und Antwortverhalten. In: V. Hüfken (Eds.), *Methoden in Telefonumfragen.* Opladen: Westdeutscher Verlag.

Humpert, A., & K. Schneiderheinze (2002). Stichprobenziehung für telefonische Zuwandererumfragen. In: S. Gabler & S. Häder (Eds.), *Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland.* Münster: Waxmann Verlag.

Kish, L. (1994). Multipopulation survey designs: Five types with seven shared aspects. *International Statistical Review 62*, 167–186.

Lavrakas, P.J. (1993): *Telephone Survey Methods.* Newburry Park: Sage.

Rothe, G. (1994): Wie (un)wichtig sind Gewichtungen? In: S. Gabler, J.H.P. Hoffmeyer-Zlotnik & D. Krebs (Eds.), *Gewichtung in der Umfragepraxis.* Opladen: Westdeutscher Verlag.

Rothe, G. & M. Wiedenbeck (1994). Stichprobengewichtung: Ist Repräsentativität machbar? In: S. Gabler, J.H.P. Hoffmeyer-Zlotnik & D. Krebs (Eds.), *Gewichtung in der Umfragepraxis.* Opladen: Westdeutscher Verlag

Sand, M. (2014a). Dual-Frame-Telefonstichproben. Entwicklung. Handhabung und Gewichtung. GESIS Technical Reports 2014/02.

Sand, M. (2015). "Überarbeitung des GESIS Auswahlrahmens für Telefonstichproben - Führt die Anreicherung durch die Angaben der Bundesnetzagentur zu einer Verbesserung der Auswahlgrundlage?" In: M. Häder & R. Neumann (Eds.), *Dresdner Beiträge zur Soziologie*. Vol. 5.

Struminskaya, B., L. Kaczmirek, I. Schaurer, W. Bandilla, S. Gabler & S. Häder (2011) Identifying non-working numbers in cell phone RDD samples via HLR-Lookup technology: Reduced survey costs and higher precision in response rate calculation. In: *Survey Practice*. August. http://surveypractice.wordpress.com/2011/08/30/identifying-nonworking-numbers/.

von der Heyde, C. (2002). Das ADM-Telefonstichprobenmodell. In: S. Gabler & S. Häder (Eds.), Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann Verlag.