Frost Hubbard, Yu-chieh (Jay) Lin, Dan Zahs, and Mengyao Hu, 2016

Introduction

Although one could employ a census to measure the entire population, it is more common to take a sample of the population. A properly designed probability sample (see probability sampling) can be used to make estimates for no the sample itself, but also for the underlying population from which it was selected. A probability sample is one in v each element of the (underlying) population has a known and non-zero chance of being selected. That is, every pers chance to be included in the study and have their characteristics, opinions, etc., become part of the data. It should be that everyone does not have to have an *equal* chance of being selected—just a *known non-zero* chance of being sele

Probability samples have several desirable characteristics. They enable us to put a margin of error or confidence int our estimates—essentially, a measure of how accurate the estimate is compared to the same estimate calculated on t population. Probability samples make it possible to not only compare the sample to the population, but also to comp sample from one population to a sample from another population, as occurs in multinational, multicultural, or multi surveys, which we refer to as '3MC' surveys.

An optimal sample design is one that maximizes the amount of information obtained per monetary unit spent within allotted time and meets the specified level of precision . One important prerequisite for 3MC surveys is that all sam full probability samples from comparable target populations .

Different nations have different sampling resources and conditions. For a multinational survey, this means that the o sample design for one country may not be the optimal design for another. (Please note this chapter uses the term 'participating country' to encompass any participating country, culture, region, or organization in a 3MC study.) The allowing each participating country some flexibility in its choice of sample design is highly recommended, so long a sample designs use probability methods at each stage of selection .

This chapter outlines the decisions that need to be made when designing a 3MC probability survey sample. It encou survey organizers to allow sample designs to differ among participating countries while, at the same time, ensuring standardization on the principles of probability sampling.

Although flexibility will usually be necessary in sample design and implementation, the guidelines and recommend discussed in this chapter are only directly applicable if a probability framework is followed. Probability designs can expensive. To reduce costs, some survey organizations select nonprobability samples, including convenience sampl quota samples. While not all nonprobability samples are biased, the risk of bias is extremely high and, most importa cannot be measured—a survey that uses nonprobability sampling cannot estimate the error in the sample estimates .

Please note that this chapter assumes that the reader has a basic understanding of statistics and terms such as 'varian 'standard deviation.' Please refer to Further Reading or an introductory statistics textbook if a refresher in statistics i

↑ Ba

Guidelines

Goal: To select an optimal, cost-efficient probability sample in each participating country that is representative of th population and allows researchers to make inferences to the target population, and to standardize sample designs wi hampering optimal designs in each participating country.

↑ Ba

1. Define the target population for the study across all countries and the survey population within each partic country.

Rationale

The survey planners of any 3MC survey need to develop a detailed, concise definition of the target population in or ensure that each participating country collects data from the same population. Without a precise definition, one cour collect data that include a certain subgroup, such as noncitizens, while another country excludes this subgroup. This difference in sample composition may influence the estimates of key statistics across countries. In addition, a precis definition will let future users of the survey data know to which exact population the survey data refer. The data use then make a more informed decision about whether to include the survey data in their analyses.

Procedural steps

1.1 Define the target population across all participating countries as clearly as possible, including the elements population and the time frame of the group . For example, a target population might be defined as "all persons the age of eighteen, who usually slept most nights in housing units in South Africa, Zimbabwe, Lesotho, and Swaziland during April, 2007." Note that this definition would, in turn, require definitions of the terms "usual "most," and "housing unit."

1.2 To ensure a clear description of the target population, think about all potential inclusion/exclusion criteria. F example, the target population might exclude:

1.2.1 Persons outside a defined age range.

1.2.2 Persons in institutions, such as hospitals, nursing homes, prisons, group quarters, colleges, monasterie military bases.

1.2.3 Persons living in certain sparsely populated or remote geographic regions.

1.2.4 Non-citizens, ethnic minorities, homeless or nomadic populations, language groups.

1.3 Define the survey population within each participating country by refining the target population based on co security, or access restrictions to all target population elements .

1.3.1 In some cases the target and final survey populations are the same. However, there can be differences. target population can be thought of as 'who we *want* to include in our survey' while the survey populatio we *can* include in our survey.' Differences can occur due to frame availability, screening ability, cost, etc example, the survey population may exclude those residing in war-torn areas, or the data collection peric be narrowed in areas with civil disturbances that are threatening to escalate.

1.3.2 Ideally, the resulting survey populations are comparable across all countries. To help analysts determi comparability, the decisions involving the definition and implementation of the target and survey popula should be well documented.

Lessons learned

1.1 Large established 3MC surveys have defined their target and survey populations differently, depending upo goals and topics of the study.

1.1.1 The is an independent, nonpartisan research project that measures the social, political, and economic atmosphere in Africa. Afrobarometer surveys are conducted in more than a dozen African countries and repeated on a regular cycle. Participants in Round 6 of the Afrobarometer Survey had to be citizens of th country and of voting age on the day of the survey. They had to complete the interview in their country's language or in an official local language for which a translation was available. People living in areas of a conflict or natural disasters, national parks and game reserves, and in institutionalized settings were excl Special cases, like areas of political unrest, were reviewed on a case-by-case basis.

1.1.2 The Asian Barometer is an applied research program studying public opinion on political values, dem and governance in thirteen East Asian political systems (Japan, Mongolia, South Koreas, Taiwan, Hong China, the Philippines, Thailand, Vietnam, Cambodia, Singapore, Indonesia, and Malaysia) and five Sou countries (India, Pakistan, Bangladesh, Sri Lanka, and Nepal). The target population of the Asian Barom defined as citizens who were at least 20 years of age and were eligible to vote (i.e., were not disenfranch to mental illness or incarceration).

1.1.3 The is an academically-driven social survey designed to chart and explain the interaction between Eur changing institutions and the attitudes, beliefs, and behavior patterns of its diverse populations. Round 7 ESS covers more than 22 nations and includes persons 15 years or older who are resident within private households, regardless of nationality, citizenship, or language; homeless and institutional populations are excluded from the sample.

1.1.4 The Living Standards Measurement Study (LSMS) was established by the World Bank in 1980 to exp ways of improving the type and quality of household data collected by statistical offices in developing c Its goal is to foster increased use of household data as a basis for policy decision making. Respondent requirements and exclusions vary across participating countries .

1.1.6 The World Values Survey is conducted by a nonprofit association seated in Stockholm, Sweden to hel scientists and policy makers better understand worldviews and changes that are taking place in the belie values, and motivations of people throughout the world. Respondents are adults, 18 years and older; som countries also place upper limits on age.

1.1.7 The World Mental Health (WMH) Survey studies mental illness in selected countries in Europe, Asia, and North and South America. One of the major goals of the WMH Study was to compare the age of ons disease across countries. Best practice might suggest strictly defining the age of majority (e.g., 18 years However, the WMH study organizers recognized that strictly defining this inclusion criterion would be d given that age of majority varies by country (and even within a country). Also, a strict definition would a study protocols such as ethics reviews and informed consent (seeking permission to interview minors). Therefore, the WMH Study had to make a difficult decision to allow the age eligibility criterion to vary a countries. In the end, the WMH Study allowed the age range to vary, with 16 years of age being the you

lower age limit; some countries also set upper age limits. This was taken into consideration in the analys
Participating countries were also allowed to vary in whether or not respondents must be citizens or be flu
specific languages .

1.2 An increasingly common form of housing seen in international studies is workers' quarters. Survey designe want to explicitly state in the definition of the target population whether workers' quarters should be included excluded.

⇡ Ba

2. Identify and evaluate potential sampling frames. Select or create the sampling frame that best covers the ta population given the country's survey budget and timeliness or feasibility of access to the sample.

Rationale

An ideal sampling frame contains all of the elements of the target population. However, very few sampling frames e allow access to every element in the target population. The goal, then, is to choose a sampling frame or a set of sam frames that allows access to the largest number of elements in the target population and contains the fewest number ineligible elements, given the constraints of the survey budget.

Procedural steps

Note: Although this chapter focuses heavily on the method and practice of in-person interviews, telephone interview random-digit dialing (RDD) frames or other lists are also widely used. In cross-national surveys, one country may c interviews over the telephone while another country conducts face-to-face interviews (see, for example, the [Gallup Poll]). This difference in the mode of data collection, driven by the available sampling frames, might lead to differen the results. (See [Data Collection: General Considerations] for more information about mixed-mode design and mode

2.1 Have each participating country identify a preexisting list (or lists) of desired elements or clusters of elemen target population to create a sampling frame. Examples include:

2.1.1 Official population registries.

2.1.2 Health registries.

2.1.3 Lists of schools.

2.1.4 Postal registries.

2.1.5 Electoral rolls.

2.1.6 Utility customer lists.

2.1.7 Pre-existing sampling frames used by other surveys.

2.1.8 Telephone directories.

2.1.9 Random-digit dialing (RDD) telephone frames.

- The phone system in many nations makes it possible to generate telephone numbers for sampling purpo without first generating all possible telephone number combinations.
- In some places, there might be separate lists or frames for landline and cellular telephone numbers.

2.1.10 Lists of email addresses.

- These are only appropriate in special circumstances.
- In some populations, many people may not have an email address; at the same time, some people may h multiple email addresses.

2.1.11 Other list(s) of addresses, phone numbers, or names.

2.2 Create a sampling frame via area probability sampling methods if there are no appropriate pre-existing lists elements (or if the lists are not made available for use by the survey project) of the target population. Even if s do exist, it is wise to assess the cost and coverage errors associated with creating an area probability sampling and how this might compare to the use of a pre-existing list. Many texts and documents provide detailed guida regarding the development of area probability samples . Additionally, cross-national surveys face a host of ch specific to the within-household phase of selection, particularly in non-Western countries: defining what cons household, large household sizes, and the availability of information on household composition (e.g. knowled birthdays of all household members) . Below, we outline a simple two-stage area probability sample of housel including the following steps used in many 3MC surveys. (Additional information can be found in Appendice A and C).

2.2.1 Create a list of primary sampling units (PSUs) based on geographic clusters. In the United States, for e these clusters are typically census enumeration areas.

2.2.2 Using a probability sampling method, select a sample of PSUs.

2.2.3 Determine the appropriate method for listing the housing units (secondary sampling units (SSUs)) wit selected PSUs.

2.2.4 Send staff to list the housing units in selected PSUs, maintaining a uniform definition of what constitu 'housing unit.'

2.2.5 Once the housing units in a PSU have been enumerated, select a random sample of housing units from

2.2.6 During data collection, ask the selected housing units within the PSUs to participate. Once the housing agreed to participate, complete a list of all eligible members within the housing unit. (See Appendix B fo detailed instructions on enumerating eligible members of the housing unit.)

2.2.7 Using a probability method, select one or more eligible members within the housing unit .

- Train the interviewer or, where possible, program the computer, to select an eligible respondent based o selection method specified.
- While some 'quasi-probability' and 'non-probability' or 'quota' within-household selection methods can be aware that such procedures produce a non-probability sample.
- Some studies may want to survey the most knowledgeable adult, such as the one with primary child car responsibilities.

2.3 Evaluate how well each potential sampling frame covers the target population .

2.3.1 Examine the sampling frame(s) for a one-to-one mapping between the elements on the sampling frame target population. There are four potential problems:

- Undercoverage (missing elements): elements in the target population do not appear on sampling frame.
- Ineligible elements: elements on the sampling frame do not exist in the target population.
- Duplication: several sampling frame elements match one target population element.
- Other mismatches: for example, one sampling frame element matches many target population elements only a street address is listed but there might be several apartment units at one address).

2.3.2 Area frames generally have better coverage properties than preexisting lists of addresses, names, or tel numbers because area frames have fewer missing eligible elements, fewer duplications, and fewer inelig elements. For more information on the creation of area probability frames, see Appendices A and C.

2.4 Consider combining multiple sampling frames which cover the same population if the union of the differen would cover the target population better than any one of the frames on its own . When combining multiple lis create a sampling frame, the following steps should be considered :

2.4.1 First, determine for each element on the combined frame whether it is a member of Frame A only, Fra only, or both Frames A and B, and calculate probabilities .

- If the membership of each element can be determined before sampling, duplicates can be removed from sampling frame.
- A variation on this is to use a rule that can be applied to just the sample, rather than to the entire frame. might be designated as the controlling frame, in the sense that a unit that is in both frames is allowed to sampled only from A. After the sample is selected, determine whether each unit from B is on the A fran retain the unit only if it is not on frame A. This method extends to more than two frames by assigning a order to the frames.
- If the membership cannot be determined prior to sampling, then elements belonging to both frames can weighted for unequal probabilities of selection after data collection (see Data Processing and Statistical Adjustment for best practices for weighting and nonresponse).

2.5 Assess the cost of obtaining or creating each potential sampling frame.

2.5.1 In most circumstances, it is less expensive to purchase preexisting lists than to create area probability

2.5.2 While three-stage area probability samples are more costly to develop than preexisting lists, they facil cost-effective clustering for interviews.

2.5.3 If pre-existing lists are not up-to-date, potential respondents may no longer live at the address on the li may have changed phone numbers; tracking these individuals can be very expensive.

2.5.4 Pre-existing lists for household surveys often contain more ineligible elements than area probability fr increasing survey costs.

2.6 Update an already existing frame, if necessary. For example, World Health Survey (WHS) administrators ha suggested that frames that are two years old or more require updating . However, that is only a rough rule of t mobile societies, a frame that is one year old may need to be updated, while in other societies, even older fran might still be accurate.

2.6.1 If the frame is a preexisting list, contact the provider of the list for the newest version and its quality documentation.

2.6.2 If the frame is an area probability sample and the target population has undergone extensive movemen substantial housing growth since the creation of the frame, then updating the PSUs and SSUs will be req However, what is most important is the quality of the enumerative listing.

2.6.3 Select the sampling frame based on the undercoverage error vs. cost tradeoff.

Lessons learned

2.1 Most countries do not have complete registers of the resident population and, therefore, construct area fram sample selection.

2.1.1 Some surveys in countries with limited resources have found that it can be difficult to enumerate the r impoverished areas and, consequently, surveys in these countries may under-represent poorer or more ru residents. However, not all survey methodologists agree with the opinions expressed by these authors reg enumeration in rural, poor areas. Those who disagree argue that the poor enumerations are mainly due tc expectations and insufficient training and supervision. Regardless, if the statistic of interest is correlated income and/or urbanicity, the sample estimate will be biased. For example, the Tibet Eye Care Assessme study on blindness and eye diseases in the Tibet Autonomous Region of China, used an area sampling fr One of the PSUs was the township of Nakchu, an area of high elevation that is primarily populated by nc herders. Because of the elevation and rough terrain, Nakchu proved difficult to enumerate accurately. As the survey sample underrepresented the residents of the roughest terrain of Nakchu. This was potentially important, as ophthalmologists believe that Tibetans who live in the most inaccessible regions and the hi elevation have the highest prevalence of eye disease and visual impairment.

2.1.2 Even when available, the quality and recency of census data and/or administrative lists can vary across countries. Data from other sources might be available. For example, IPUMS-International is an "effort tc inventory, preserve, harmonize, and disseminate census microdata from around the world. The project ha collected the world's largest archive of publicly available census samples. The data are coded and docum consistently across countries and over time to facilitate comparative research. IPUMS-International mak data available to qualified researchers free of charge through a Web dissemination system" . Information available from the Demographic and Health Survey (DHS) Program on survey boundaries, geospatial da maps for the most recent DHS and past surveys for over 90 countries.

2.2 Local residents can help produce maps for an area probability sample. When measuring the size of the rural population in Malawi, researchers used statistical methods to determine the sample size and selection of villag they asked members of the selected communities to help draw maps, including exact village boundaries, key landmarks, and each individual household .

2.3 Detailed information on area probability sampling is available through the DHS, including the Sampling an Household Listing Manual, sampling and weighting tutorials, and periodic online courses.

2.4 Technology can also permit the associated use of GPS to facilitate sample selection. used personal digital as (PDAs) equipped with GPS units in household surveys to rapidly map all households in selected areas, choose random sample, and navigate back to the sampled households to conduct an interview in Togo and Niger. Incorporating GPS into the sampling method allows researchers to select a random sample of households for interviewing from a complete and up-to-date listing of the households and, as a result, the probability of selec known. The GPS data collected also provides geospatial information for reports and analyses. were able to ge preliminary report of survey findings from the aggregated data (including maps of the districts sampled showi the households and their sample inclusion status) within a few days of completion of data collection. This can particularly useful in situations where data are needed quickly to guide public health action, such as in routine monitoring and evaluation and rapid needs assessments.

2.5 discuss several sampling approaches using geographic information system (GIS) tools for face-to-face hous surveys. Examples include the use of Google Maps, Google Earth, satellite photos, handheld global positionir system (GPS) devices, and location-enabled applications on mobile phones. These technologies can be used to and select clusters in early stages of selection through the use of a variety of methods (e.g., grid method, segm of nighttime lights, LandScan population estimate data, random geographic cluster sampling (RGCS)) and to households in a later stage of selection (satellite maps, Qibla method, reverse geocoding, data from unmannec vehicles).

2.6 provides an overview on within-household selection methods for face-to-face household surveys. Various m exist, and the two commonly used ones are Kish (see Appendix B) and birthday methods. The Kish method is believed to be the gold standard for within-household selection. The technique requires all eligible persons in household to be listed. A person is sampled with equal probability from all eligible persons. Birthday methods household members' birthday information to select the respondent. Using data from the European Social Surv (ESS), provide evidence that the Kish method performs better than the birthday method with regard to gender representation.

⇡ Ba

3. Choose a selection procedure that will randomly select elements from the sampling frame and ensure that important subgroups in the population will be represented.

Rationale

Sample selection is a crucial part of the survey lifecycle. Since we cannot survey every possible element from the ta population, we must rely on probability theory to make inferences from the sample back to the target population.

Procedural steps

3.1 Consider only selection methods that will provide a probability sample.

3.1.1 Statisticians have developed procedures for estimating sampling errors in probability samples which a any type of population.

3.1.2 Random sample selection protects the researcher against accusations that their bias, whether conscious unconscious, affected the selection.

3.2 Identify the optimal sampling method available in each country. Below are summaries of each selection me See Appendix C for additional information about each selection method.

3.3 Consider simple random sampling (SRS) without replacement. In SRS, each element on the frame has an ec probability of selection, and each combination of n elements has the same probability of being selected. Du benefits of stratification, this technique is seldom used in practice.

3.3.1 Advantages of SRS:

- The procedure is easy to understand and implement.

3.3.2 Disadvantages of SRS:

- The costs in attempting to interview a simple random sample of persons can be quite high.
- SRS provides no assurance that important subpopulations will be included in the sample.

3.4 Consider stratified sampling (see Appendix C for a detailed description). Stratified sampling uses auxiliary information on the sampling frame to ensure that specified subgroups are represented in the sample and to im survey precision. Virtually all sample implementations use some form of stratification.

3.4.1 Some examples of commonly used stratification variables are:

- Region of the country.
- State/province.
- City/town, community, municipality.
- Postal code.
- Metropolitan status/urbanicity.
- Size of sampling unit (e.g., population of city).
- Race/ethnicity.

- National origin.

3.4.2 Advantages of stratified sampling:

- Stratified samples can be ether proportionate (which is general used to improve the precision of estimat total population) or disproportionate (to help improve sub-group estimates).
- Depending on the allocation of elements to the strata, the method can produce gains in precision (i.e., d in sampling variance) for the same efforts by making certain that essential subpopulations are included sample.

3.4.3 Disadvantages of stratified sampling:

- For any given frame stratification variables may be limited.
- No gains in precision will be seen if the stratification variables are not correlated with the statistic(s) of In some cases, the precision may even decrease.

3.5 Consider systematic sampling to reduce the operational effort needed to select the sample. In systematic sar every k^{th} element on the sampling frame is selected after a random start.

3.5.1 Advantages of systematic sampling:

- The operational time necessary to select the sample can be reduced substantially.
- If the sampling frame is sorted into groups or ordered in some other way prior to selection, the systemat sampling method will select a proportionately allocated sample (see description below of stratified sam This is often referred to as 'implicitly stratified sampling.'

3.5.2 Disadvantages of systematic sampling:

- If the key selection variables on the sampling frame are sorted in a periodic pattern (e.g., 2, 4, 6, 2, 4, 6 the selection interval coincides with periodic pattern, the systematic sampling method will not perform periodicity is a problem, several systematic samples can be selected and concatenated to form the total sample.
- If the list is sorted in a specific order before selection, the repeated sampling variance.

3.6 Consider cluster sampling (see Appendix C for a detailed description). With cluster sampling, clusters of fr elements are selected, rather than selecting individual elements one at a time. Within the selected clusters, we interview all or a sample of households. We can even do multi-stage cluster sampling where we select an addi sample of clusters inside the larger selected clusters.

3.6.1 Advantages of cluster sampling:

- When survey populations are spread over a wide geographic area and interviews are to be done face-to- can be very costly to create an element frame and visit n elements randomly selected over the entire
- A full frame of all elements in the entire population is not required only the elements within selected cl are needed. This will reduce listing costs.

3.6.2 Disadvantages of cluster sampling:

- Estimates are not as precise as with SRS, necessitating a larger sample size in order to get the same leve precision. See Appendix D for more information about effective sample size.

3.7 Consider two-phase (or double) sampling (see Appendix C for a further description). The concept of two-pl sampling is to sample elements, measure one or more variables on these 1st-phase elements, and use that info to select a 2nd-phase subsample.

3.7.1 A common application is to collect 1st-phase data that is used to stratify elements for the 2nd-phase subsample.

3.7.2 Survey samplers use two-phase sampling to help reduce nonresponse, with the stratifying variable fro one being whether the person responded to the initial survey request. For example, samplers might selec subsample of nonrespondents and try to entice the nonrespondents to participate by offering incentives.

3.8 Consider replicated (or interpenetrated) sampling. Replicated sampling is a method in which "the total sam made up of a set of replicate subsamples, each of the identical sample design" .

3.8.1 Advantages of replicated sampling:

- It allows the study of variable nonsampling errors, such as interviewer variance.
- It allows for simple and general sampling variance estimation (see Data Processing and Statistical Adju for further explanation, especially regarding the methods Balanced Repeated Replication and Jackknife Repeated Replication).

3.8.2 Disadvantages of replicated sampling:

- There is a loss in the precision of sampling variance estimators; a small number of replicates leads to a in the number of degrees of freedom when calculating confidence intervals.

3.9 Consider using a combination of techniques such as a stratified multistage cluster design.

3.9.1 Most surveys in countries with limited resources are based on stratified multistage cluster designs . Th combination of these techniques reduces data collection costs by clustering while striving to increase or precision through stratification.

Lessons learned

3.1 Probability sampling at every stage generally requires more labor and funding than other methods. Therefo 3MC studies have used probability sampling in the first stage of selection and then allowed quota sampling o substitution to occur at later stages . However, a survey that uses a nonprobability sampling method at any sta selection cannot estimate the sampling error of the estimates . Therefore, the coordinating center should make effort to promote the use of a full probability sample and remove any obstacles that would prevent participati countries from using probability methods at each stage of selection. For the first few waves of data collection, International Social Survey Programme (ISSP) allowed countries to use nonprobability methods at the househ level. After recognizing the problem this caused in variance estimation, the ISSP has required countries to use probability samples since 2000 .

3.2 Existing 3MC surveys have employed various strategies for selecting a probability sample.

3.2.1 Round 6 of the Afrobarometer Survey uses a clustered stratified multi-stage area probability sample. T sampling design has four stages in urban areas: (1) stratify and randomly select primary sampling units, randomly select sampling start-points, (3) randomly choose households, and (4) randomly select individu respondents within households; and five stages in rural areas: (1) randomly select secondary sampling u (SSU), (2) randomly select two primary sampling units (PSU) from each SSU, (3) randomly select samp start-points from each PSU, (4) randomly choose households, and (5) randomly select individual respond within households .

3.2.2 Sample designs for Round 5 of the must use random probability sampling at every stage. Samples are by a sampling expert or panel and may include clustering and stratification. Quota sampling and substitu not allowed, although subgroups may be over-sampled. Sample designs and frames must be documented and be pre-approved by a sampling expert or panel. The target minimum response rate is 70%.

3.2.3 Sampling frames and designs for the vary across participating countries, but generally consist of two s the first stage, the sample frame is developed from census files and Primary Sampling Units are randoml selected with probability proportionate to size; in the second stage, households (usually 16) are randomly selected from each of the designated Primary Sampling Units. Clustering and stratifying are permitted, b sampling procedures must be documented and made available to data analysts.

3.2.4 Survey of Health, Aging and Retirement in Europe (SHARE) sampling designs vary by country but al required to be probability samples. Three sampling designs may be used: (1) stratified simple random sa from national population registers, (2) multi-stage sampling using regional or local population registers, single or multi-stage sampling using telephone directories followed by screening in the field .

3.2.5 Sampling frames for the World Mental Health Survey vary across participating countries, but generall of three types of sampling frames: (1) individual contact information databases such as national populati registries, voter registration lists, or household telephone directories; (2) multistage area probability sam frames; or (3) hybrid multistage frames that combine area probability methods and a individual contact d in the final stages. Sampling designs vary across participating countries, including stratification and clus but probability sampling is required at all stages. The target minimum response rate is 65% .

3.2.6 Probability sampling is strongly recommended, but not required, in the World Values Survey; any dev from probability sampling are to be reported in the Methodology Questionnaire report.

3.3 When there is very little information on the population, surveys (including Europe-wide surveys, surveys in developing countries, and worldwide international surveys) sometimes use random route (or 'random walk') a the multistage face-to-face interview sampling method.

3.3.1 For each randomly-chosen sampling point (e.g., urban units, small cities, or voting districts), interview assigned a starting location and provided with instructions on the random walking rules—e.g., which dir start, on which side of the streets to walk, and which crossroads to take. Households are selected by inte following the instructions. The routes end when the predefined number of respondents (or households) is achieved . Since the probability of the selected household is unknown, this method is categorized as non probability sampling methods . See Table 1 in for a list of surveys which used this method in data collec

3.3.2 The advantage of random route sampling is that it is usually cheaper and easy to conduct. Most import this sampling method makes it possible when there is no list of respondents available .

3.3.3 This method assumes that each household in this sampling point has equal probability of being selecte However, a recent study evaluated the bias in random route sampling using registration office data (to ve impact of selection errors on survey results), and found that "all tested routes strongly violate the equal probability assumption and lead to biased expected values in multiple variables" . Additional error sourc includes incorrect interviewer behaviors, which can lead to coverage and sampling error (also see Parada Other Auxiliary Data).

3.3.4 In a review of a survey of blindness in war-torn southern Sudan, note that the estimate of trachoma pr is higher than previously reported for Africa, including other studies in conflict-affected places. Kuper a

Gilbert suggest that methodological problems, including an element of subjectivity resulting from the ra route method employed for household selection, may have resulted in an overestimate of the disease.

⇡ Ba

4. Determine the sample size necessary to meet the desired level of precision for the statistics of interest at po or subgroup levels for the different potential sample selection procedures.

Rationale

After choosing a sample design, and before selecting the sample from the sampling frame, the sample size must be determined. The sample size takes into account the desired level of precision for the statistic(s) of interest, estimates statistic of interest from previous surveys, the design effect, and estimated outcome rates of the survey. (See for a d treatment of the approach used in the European Social Survey. For a more extensive example of sample size calcula Appendix D).

Procedural steps

4.1 Specify the desired level of precision, both overall and within key subgroups. Practical experience has dete that often it is easiest for sponsors to conceptualize desired levels of precision in terms of 95% confidence int

4.2 Convert these 95% confidence intervals into a sampling variance of the mean or proportion.

4.3 Obtain an estimate of S^{2} (population element variance).

4.3.1 If the statistic of interest is not a proportion, find an estimate of S^{2} from a previous survey on t target population or from a small pilot study.

4.3.2 If the statistic of interest is a proportion, the sampler can use the expected value of the proportion (\(p\ if it is a guess, to estimate S^{2} by using the formula $S^{2}=p \left (1-p \right)$.

4.3.3 If no information about p is available, a researcher can assume that it is 50%, which will yield a conse estimate of the sample size.

4.4 Estimate the required number of completed interviews for an SRS by dividing the estimate of S^{2} by desired sampling variance of the mean. See for more on sample size computation for SRS.

4.5 Multiply the number of completed interviews by the design effect to account for a non-SRS design.

4.6 Calculate the necessary sample size by dividing the number of completed interviews by the expected respo eligibility rate, and coverage rate.

4.6.1 The sampler can estimate these three rates by looking at the rates obtained in previous surveys with th or similar survey population and survey design.

Lessons learned

4.1 Prior to the first implementation of the European Social Survey (ESS), many of the participating survey organizations had never encountered the concepts of sample size determination and calculating design effects Therefore, the ESS expert sampling panel spent considerable time explaining these. In return, the organizatio were new to these methods were very enthusiastic to learn about them, and eager to meet the standards of the coordinating center. In fact, after completing Round 1 of the study, many nations commented that designing t sample was one of the most educational aspects of the entire survey process, and had significantly improved t survey methods within their country.

4.2 Sample size frequently varies among countries participating in cross-cultural surveys. In Round 6 of the Afrobarometer Survey, sample size ranges from a minimum of 1,200 respondents to 2,400 or more in extreme heterogeneous areas ; sample size ranges from 800 to 3,200 respondents in the study; Round 5 of the Europea Survey (ESS) requires a minimum of 800 respondents for participating countries that have a population of les two million, 1,500 from larger countries ; the requires a minimum of 1,000 respondents, with a goal of 1,400 respondents; sample size ranges from 1,600 to 5000 households in the ; the Survey of Health, Ageing and Ret in Europe (SHARE) requires 1,500 respondents from each participating country ; samples in the World Menta Survey range from 2,357 (Romania) to 12,992 (New Zealand) ; and in most countries in the World Values Sur minimum sample size is 1,200 respondents.

⇡ Ba

5. Institute and follow appropriate quality control procedures at each step of the sample design process.

Rationale

Development and implementation of quality control procedures for the sample design are necessary to ensure the hi level of coverage possible and to maintain a probability sample that meets the desired level of precision for key surv statistics. If a failure to meet those standards is detected, protocols should be in place to remedy the failure. In additi monitoring of procedures related to the sample design of the study should inform efforts to improve the quality and effectiveness of the study over time.

Procedural steps

5.1 Define the target population for the study across all participating countries/cultures as well as the target pop within each country/culture. If the study design does not change over time, strive to keep each target populati overall and within each participating country, consistent over time.

5.2 Prior to selecting sample elements or sampling units, provide the data collection staff with a list of all of the variables on the sampling frame, and ask which variables they would like and the format in which they would these variables delivered once sampling is complete for data collection purposes.

5.2.1 After sample selection, check that each selected sampling unit or element contains this information an the specified format.

5.3 If possible, use a responsive survey design to help achieve an optimal sampling design (see Survey Quality Paradata and Other Auxiliary Data, and Data Collection: General Considerations for more information about responsive survey designs). A responsive survey design uses prespecified paradata (quantitative indicators of collection process such as 'contact attempts' or 'interviewer success rate') for intervention during data collectic

5.3.1 Advantages of responsive survey designs are the prespecification of interventions instead of ad hoc de and the possibility to target efforts on hard to interview groups.

5.3.2 A disadvantage is that the survey designers walk a thin line between full probability and quota sample deviate from carefully predefined paradata-driven interventions.

5.4 After each stage of selection, generate frequency tables for key variables from the frame of sampling units t for the following:

5.4.1 Overall sample size and within stratum sample size.

5.4.2 Distribution of the sample units by other specific groups such as census enumeration areas.

5.4.3 Extreme values.

5.4.4 Nonsensical values.

5.4.5 Missing data.

5.5 Create a unique sample identification code for each selected sampling unit. This code will allow identifying information to be easily removed after completing data collection. The code value can then be used in place o values to preserve confidentiality when the original values may contain information that could lead to identify respondents.

5.6 Whether the participating country or the coordinating center is selecting the sample, assign a second sampli statistician within that organization to check the sample design methodology and the statistical software synta survey's primary sampling statistician.

5.7 Save all data files and computer syntax from the statistical software package used during the sample design in safe and well-labeled folders for future reference and use.

Lessons learned

5.1 The construction and maintenance of sampling frames constitute an expensive and time-consuming exercis participating country determines that no sampling frame meeting the specified coverage level of the target po exists, they can create a frame from sources such as census data collected by national statistics offices. Howe should be aware that official statistics differ greatly in accuracy, as well as availability, from country to count

5.2 As discussed in Guideline 2, Lessons learned, the decision to stray from full probability sampling reflects t conflict between standardization and flexibility in 3MC surveys. However, it bears repeating that without pro sampling, one cannot make justifiable inferences about the target population from the sample estimates.

⇡ Ba

6. Document each step of the sample selection procedure.

Rationale

Over the course of many years, various researchers will analyze the same survey data set. In order to provide these users with a clear sense of how and why the data were collected, it is critical that all properties of the data set be documented. In terms of the sample design and selection, the best time to document is generally shortly after sampl selection, when the information regarding sample selection is fresh in one's mind.

Procedural steps

6.1 Have participating countries document the sample selection procedure while selection is occurring or shortl thereafter. Ideally, set a deadline that specifies the number of days after sample selection by which each partic country must send sampling selection documentation to the host survey organization. Be sure to allow for app time to review and revise documentation when setting the deadline (see Tenders, Bids and Contracts and Stud Management.)

6.2 Include the following:

6.2.1 A clear definition of the survey population, as well as the differences between the target population an population.

6.2.2 The sampling frame:

- Both the sampling frame used and the date the frame was last updated.
- A description of the development of the sampling frame and the frame elements.
- A description of the how well the sampling frame is thought to cover to target population and the poten coverage error.

6.2.3 The data file of selected elements:

- A descriptive and distinct variable name and label.
- Unique variables that contain the selection probabilities at each stage of selection as well as the overall probabilities. If a participating country used a nonprobability method in at least one stage of selection, a therefore the selection probabilities are unknown, ensure that this is clearly documented.
- A clear description of all variables in the selected element data file, with all variable names, an accomp description, and a codebook which provides question-level metadata that are matched to variables in a d
- The statistical software syntax used for checking the dataset of selected sampling units or elements.

6.3 For each sample, indicate how many stages were involved in selecting the sample (include the final stage in the respondent was selected within the household, if applicable), and a description of each stage, including ho sampling units were selected at each stage.

6.3.1 Examples of different stages include:

- State/province.
- County or group of counties.
- City/town, community, municipality.
- Census/election district.
- Area segment/group of neighborhood blocks.
- Housing unit/physical address (not necessarily the postal address).
- Postal delivery point/address.
- Block of telephone numbers (e.g., by regional prefix).
- Telephone number.
- Person selected from a household listing or a non household listing (a list, registry or other source).

6.3.2 Examples of how sampling units were selected:

- All selected with equal probability.
- All selected with probability proportional to size; specify the measure of size used. (See Appendix C fo on probability proportional to size sampling methods.)
- Some units selected with certainty, others selected with probability proportional to size; describe the cri used for certainty selection.
- Census/enumeration (all units selected with certainty).
- Units selected using a nonprobability method (e.g., convenience sample, quota sample).

6.3.3 At each stage of selection, describe the stratification variables and reasons for choosing these variable

6.3.4 At each stage of selection, explain the allocation method used and the sample size for each stratum at e stage of selection. (See Appendix C for more on allocation methods in stratified sampling.)

6.4 If systematic sampling was used at any stage of selection, indicate whether the frame was sorted by any var prior to systematic selection in order to achieve implicit stratification. If this is the case, describe the variable(

6.5 Describe the time dimension of the design (i.e., one-time cross-sectional, fixed panel, or rotating panel desi

6.5.1 If a panel study:

- State how many previous waves or rounds of data collection there have been for this panel study.
- Describe the initial sample design for the panel study and any subsequent modifications to the design th important in documenting this study.

6.5.2 If a rotating panel design:

- Fully describe the rotating panel design for the study (e.g., fresh cross-section is drawn each month and respondents are interviewed once that month, and then reinterviewed once six months later).
- State the anticipated precision.
- Explain any problems encountered during the sampling process and any deviations from the sampling p during implementation.

6.5.3 Additional sampling documentation:

- Report any (additional) subsampling of eligible respondents, carried out in order to control the number interviews completed by respondents with particular characteristics (e.g., one in two eligible males was interviewed, one in four eligible persons with no previous history of depression was interviewed (descr protocol)).
- Describe any use of replicates (see Data Processing and Statistical Adjustment).
- Explain if releases (nonrandom subsets of total sample) were used or the entire sample was released to collection staff at the start of the study.
- Recount in detail any substitution.

Lessons learned

6.1 As the procedural steps outlined above show, selecting a sample can involve many detailed steps that may b recall after the fact. For example, the coordinating center for the World Mental Health Survey began gathering sampling documentation for weighting and other purposes after many of the participating countries had finish collection. They found that some countries had a difficult time recalling all the necessary details, such as the s size for each stratum at each stage of selection. It is wise to document sampling procedures in detail shortly af sample selection (see Data Processing and Statistical Adjustment for further explanation of the weighting pro

↑ Ba

References

↑ Ba