

Chapter 4

Collecting the Data



Abstract This chapter discusses aspects related to data collection. It focuses, firstly, on sampling issues and, secondly, on the survey mode. Sampling issues include sample size and the type of sampling that enable precise estimates to be obtained. Regarding the survey mode, discrete choice experiments can be implemented by mail, telephone, face-to-face or web surveys. Each of these survey modes has its advantages and shortcomings. They are described and compared in the course of this chapter, addressing an important decision in the planning process of a discrete choice experiment.

4.1 Sampling Issues

Most SP studies implicitly or explicitly aim for “representative samples” and generalisable results. This implies that the survey population, the persons, households, etc., which shall be generalised has to be known (Dillman et al. 2008). It further demands an appropriate sampling frame, a list from which the sample is drawn. Two well-known errors are the coverage error (a non-sampling error), referring to units in the survey population with a non-zero probability of being included in the survey, and the sampling error which refers to only collecting data from a subset and not all units of the sampling frame. The coverage error is present if, for example, all users of an environmental good comprise the population, but researchers sample from a household register that does not include all users, i.e. the sampling frame is not complete covering the intended population of interest. An error would occur if users who are not included in the household register have characteristics that differ from those included in the register. A sampling error is present if not all members of a population are included in the sample and figures such as mean values and willingness-to-pay estimates based on this sample differ from those based on the population. To some extent, all statistics based on a sample are biased, yet the precision of the estimates varies with the type of sample and sample size. Sample weights can be used to take sampling error into account; however, they will not overcome the weaknesses of a sampling approach (such as non-probability samples, see, e.g., Yeager et al. 2011).

Given that the survey population is known, a simple random sample can be drawn if lists of households, postal addresses or e-mail addresses are available. Then a computer program can be used that numbers respondents and randomly selects them. A stratified sample, separate and disproportionate samples for specific groups, can be employed if some groups of the population have a greater chance of being included in the survey. Coverage error can be especially problematic with web surveys (Couper 2000; Bonnichsen and Olsen 2016), as for example not all individuals in a population might have access to or use the Internet, or it is difficult to construct a list with all individuals with Internet and web access, from which a random sample can be generated. There are many survey organisations (panel providers), which offer web surveys and samples for web surveys based on so-called access panels. These panel providers differ in their sampling approaches and this can make a big difference in terms of survey quality and sampling error. While some providers work with opt-in panels, where individuals volunteer to take part in surveys, others recruit panel members “offline” using, for example, a random telephone sample design (or a mix of sample designs) to reduce sampling error. Clearly, the latter approach based on some kind of probability-based sample design results in better samples and survey quality (Yeager et al. 2011). In general, generalisations for a population are strictly speaking not possible from non-probability samples. This also applies to using social media like Facebook, Twitter, etc., to recruit survey participants for web surveys. Here respondents typically select themselves for the survey, and social media users can differ from the rest of the population, which can cause biased samples. Also, large web survey samples do not automatically mean that the data are more valid and generalisable (see, e.g., Savage et al. 2013; Mills 2014 for a web survey with over 160,000 respondents and a massive sample error).

With respect to users of an environmental good, the population (e.g. users of a national park) is often not known and DCEs may be conducted onsite (e.g. in the national park), or offsite by using a mail or web survey of the citizens in a region or a country. In this case, it might be advisable to collect data over different days and times of day and to work with quota (e.g. for gender, age and education) in order to obtain some control over the sampling process and to make sure that different user groups are represented in the sample. Respondents can be determined by a systematic approach such as asking every tenth person to take part in the survey.

Often the survey population, e.g. the market size, is not known and has to be estimated (see Glenk et al. 2020 for an overview). The market size refers to the distance between the environmental good/resource and the point where WTP drops to zero (e.g. Bateman et al. 2006). In many cases, this might not correspond with political jurisdictions. In general, the definition of the market can be challenging. For example, in research on the value of national parks it is important to differentiate between users of parks and non-users, where both groups can receive benefits from the park in terms of use and/or non-use values. Therefore, it has to be decided whether all citizens in a country belong to the study population (the “market”), citizens in regions close to the park, or only citizens who actually use the park, etc. Furthermore, some parks might attract visitors from different countries and, again, this can influence the market size. In order to test for market size, researchers can sample individuals living

in different distances to the environmental good/resource and then examine distance decay effects, i.e. to which extent WTP for the good decreases with distance, holding everything else constant (see Glenk et al. 2020).

A question that is often raised is the sample size that is needed in a DCE study. Here, two aspects have to be differentiated. The sample size question might firstly refer to the representativeness of the data collection, i.e. how well the sample represents the underlying population and its characteristics. This is important if DCE results shall be generalised to the population and a population's preference heterogeneity regarding characteristics such as gender, age, education, income and attitudes are of interest. Secondly, it might refer to the sample size needed to obtain statistically significant parameter estimates in the choice experiment. A practical problem might be that the sample size requirements for statistically significant parameter estimates might be different from those referring to data representativeness. For example, efficient experimental designs might suggest a low number of respondents (e.g. 300); yet, in order to analyse preferences for subgroups in the data (e.g. respondents with low or high environmental concern) larger sample sizes are needed to detect differences and to represent the population at hand.

In principle, focusing on the proportions of responses, the sample size, for example, required for representing the population in a two-alternative case with a specific certainty can be calculated (see formulas presented in Dillman et al. 2008, p. 56). In order to represent a country's population in terms of socio-demographics, a sample size of around 1,000 respondents should be sufficient and this number does not depend on the size of the country. Therefore, most cross-country surveys such as the World Value Survey include between 1,000 and 1,500 respondents per country. Similarly, the minimum sample size for estimating a proportion in a multinomial case can be determined (see Louviere et al. 2000). Some recommendations regarding the sample size requirements for stated choice experiments can also be found in Rose and Bliemer (2013) and de Bekker-Grob et al. (2015). Furthermore, it is important to stress that, once the experimental design has been generated, the sample variation for the model parameters can be analysed by simulation experiments like those presented in Sect. 3.3. Depending on the complexity of the experimental design and the type of model applied, sample sizes of 300–500 respondents might be sufficient to obtain valid estimates for stated preferences. But there are many situations and models for which this sample size may not be large enough.

In general, there is a trade-off between the number of respondents and the efficiency of the experimental design: the larger the sample size, the less important it is to have a very efficient design. For smaller sample sizes, such as 300 respondents, it is important to consider that sufficient data need to be collected to represent and analyse preference heterogeneity for subgroups in a population (e.g. regarding gender, age groups, education levels, use or non-use of the good). This can be achieved by over-sampling specific groups which are of interest. Moreover, small samples do not allow for precise estimation of more complex models.

Most researchers aim for a high response rate and see this as an indicator of a “good” survey. With respect to reporting response rate, the American Association for Public Opinion Research standards (AAPOR 2016) can be recommended. However,

high response rates should not be confused with non-response errors if those who do not take part in a survey differ from those who take part in the survey with respect to relevant beliefs, attitudes and socio-demographic characteristics. Surveys with high response rates might have a large non-response error and might not represent the population at hand well, and surveys with low response rates might have a low non-response error (Dillman et al. 2008). Furthermore, a high response rate is not beneficial if the questionnaire itself is problematic. Evaluating the quality of a survey can be a complex task, depending on different types of errors (sampling error, coverage error, nonresponse error and measurement error) and should not be related to a single measurement of quality.

Sampling involves many decisions and trade-offs. In any case and, if possible, a random sample of the population of interest is still the best approach to reduce sample-related errors. When working with web surveys and Internet panel providers, it is important to be aware of the type of access panel and to avoid opt-in panels. Some panel providers recruit their panel members based on probability samples, which is clearly preferable to non-probability samples. Probability-based samples are also needed if the aim of the study is to reveal generalisable findings for the population (of a region, country, etc.). While, given a very efficient experimental design, small samples (e.g. 300 respondents) might be sufficient to obtain valid SP estimates, it should be kept in mind that a larger sample might be needed to investigate preference heterogeneity regarding respondents' characteristics. On the other hand, if a sample is large (e.g. around 1,000 respondents representing the population of a country), the efficiency of the experimental design becomes less important. Finally, the estimation of the market size is a challenge in many environmental valuation studies. In this regard, it could be a good idea to sample individuals/households with different distances to the environmental good/resource and to test for distance decay effects, i.e. to what extent WTP for the good decreases with distance.

4.2 Survey Mode (Internet, Face-To-Face, Postal)

In principle, choice experiments can be implemented in any survey mode: mail surveys, telephone surveys, face-to-face surveys and web surveys. While some survey modes may have specific advantages over other modes, it has to be stressed that choosing a survey mode may also depend on the research context. For example, in development research, when collecting data in a remote area setting, face-to-face interviews might be the only option (Liebe et al. 2020). Likewise, an onsite survey is mostly conducted face-to-face or self-administered at the research site. While the research context can determine the survey mode, the survey mode can also affect the sampling approach. For example, if researchers plan to use a web survey they typically work with online access panels and not a random sample from the population, depending on the panel provider (see Sect. 4.1).

Face-to-face interviews: Computer-assisted personal interviewing (CAPI) is most often employed in face-to-face interviews: the questionnaire is in the form of a

computer program; the interviewer sees items on a screen (laptop or other mobile device), reads questions to respondents and enters the answers by pressing the corresponding keys (Loosveldt 2008). The presence of an interviewer can be an advantage for clarifying questions and surveying more complex issues, also complex DCE. However, it is important to consider that given the characteristics of a face-to-face interview, an interviewer could be a source of measurement error: Social desirability bias is an example of interviewer bias—the mere presence of an interviewer leads to a “systematic underreporting of undesirable attitudes or behaviour (e.g., drug use) and the systematic over-reporting of desirable ones (e.g., voting behaviour)” (Loosveldt 2008, p. 215). Such interviewer effects can be reduced by increasing the number of interviewers or decreasing the number of interviews for each interviewer, as well as reducing intra-interviewer correlation by providing additional interviewer training to standardise behaviour, and a follow-up of interviewers and feedback during field work.

Telephone survey: This survey mode is an interview survey (Steeh 2008) although technological innovations (answering machines, call blocking, wireless communication, Internet telephony) have changed the conditions for conducting telephone surveys over the last few decades. This has also affected response rates which have declined in most western countries. Since choice experiment tasks are often complex, telephone surveys have a disadvantage because they only contain auditory channels of communication and, hence, it is difficult to keep respondents involved, so interviews have to be shorter, questions should be relatively uncomplicated and only questions with a limited number of response categories can be employed (Steeh 2008). However, it has been demonstrated that multifactorial survey experiments such as (complex) vignette studies can also be integrated in telephone surveys (e.g. Emerson et al. 2001).

Mail survey: Mail surveys can be described as consisting of “questionnaires that are sent by postal mail to a sampled individual, who is requested to complete the questionnaire and send it back; no interviewer is present and the survey is completely self-administered” (Leeuw et al. 2008, p. 243). In comparison with face-to-face interviews, they can be implemented at low costs and respondents have less time pressure to answer the survey. Visual stimuli such as pictures and choice sets can be used and there is no interviewer bias. Furthermore, respondents have a greater degree of privacy compared with survey modes involving an interviewer. However, researchers cannot control who is answering the questionnaire and can also not control in which order respondents answer the questions. It might, for example, be a problem for a study if respondents can go through all the choice tasks provided before starting to answer them but they might check for the overall best alternative and choose the status quo or opt-out alternative on all the other tasks.

Web surveys: This is a computerised, self-administered survey mode without the presence of an interviewer. DCE and randomising questions can be easily implemented in web surveys; also paradata such as response time can be automatically collected. However, it should be considered that “[i]nternet users tend to read more quickly, are more impatient, and they scan rather than carefully read the text” (Lozar Manfreda and Vehovar 2008, p. 276). In web surveys, nonverbal aspects of the survey

have to be taken into account. It should be kept in mind that respondents use different web browsers, operations systems and hardware. Web surveys can be answered on different devices and with the increasing popularity of mobile phone usage around the globe respondents use mobile devices more and more frequently to answer web surveys. Recent research shows that there are systematic differences in response behaviour depending on whether the survey was answered on a personal computer or mobile device (Couper et al. 2016). This affects for example questions with an open answer format. However, the overall differences are rather small. This is also suggested in the first studies looking at mobile device effects on the results of stated choice experiments. For example, in a choice experiment on renewable energy expansion Liebe et al. (2017) do not find significant differences in WTP values for desktop and mobile device users.

Mixed-mode surveys: Combining different survey modes is a way of taking advantage of the strengths and compensating for the weaknesses of each mode (Leeuw et al. 2008). This includes having some respondents complete a questionnaire in a different mode than other respondents. Multiple modes can also be used in different stages of the survey process, e.g. in the screening and contact stage (e.g. first telephone followed by a mail survey), main data collection stage (e.g. combination of telephone survey and follow-up mail survey), follow-up stage (e.g. first mail survey followed by a telephone or web survey). Mixed-mode approaches are often employed in surveys on sensitive topics by combining face-to-face interviews and a self-administered questionnaire. This combination is also useful for DCE because more complex choice tasks can be better integrated in a self-administered mode.

Table 4.1 presents the comparison of survey models and demonstrates that no survey mode is better than all the others. For example, mail and web surveys do not suffer from interviewer effects which could be present in environmental valuation studies if the good at hand is highly socially desirable. Here mail and web surveys have an advantage over face-to-face and telephone surveys. Also mail and web surveys are less costly than face-to-face and telephone surveys. However, face-to-face surveys in particular allow for longer interviews, more complex questionnaires including choice experiment tasks, different ways of information transmission, etc. Here they have a clear advantage over all other survey modes.

While Table 4.1 implies some trade-offs when choosing a survey mode, there are aspects of DCE which suggest that web surveys have specific advantages over the other survey modes. First, randomisation of questions, choice tasks, and alternatives and attributes within choice tasks are easy to implement in web surveys, compared to mail and face-to-face surveys as well as telephone surveys (except the latter two are computer assisted). Second, visual elements such as images and short videos, that help to describe choice attributes or choice tasks, can be conveniently included in web surveys. Third, web surveys are self-administered and, hence, interviewer effects and socially desirable response behaviour are not present or less likely than in face-to-face and telephone surveys. At least compared to mail surveys, it can also be ensured that respondents in web surveys evaluate each choice set without knowing the subsequent choice sets included in the survey (i.e. they cannot screen the whole questionnaire before answering). Fourth, compared with other survey

Table 4.1 Comparison of survey modes

Aspect	Explanation	Face-to-face	Telephone	Mail	Web
Interviewer effects	“Interviewer as source of measurement error” (Loosveldt 2008, p. 214)	---	-	0	0
Media-related factors	“Social conventions and customs associated with media used in survey mode” (Leeuw et al. 2008, p. 116)	++	+	-	-
Information transmission	Presentation of information, channels of communication, regulation of communication flow	++	---	-	+
Complexity allowed	Choice task complexity, for example	++	---	+	+
Length	Interview duration	++	-	+	---
Paradata	Response time, for example	+ ^a	+	-	++
Costs	Assuming that a company is conducting the survey	---	-	+	++

Note ++ = very positive (advantage), + = rather positive (advantage), - = rather negative (disadvantage), --- = very negative (disadvantage); 0 = effect is absent

^aIn case of computer-assisted personal interviews (CAPI)

modes, valuable paradata such as response time can be collected in web surveys. Fifth, web surveys are less costly than face-to-face and telephone surveys. On the other hand, due to time constraints it is more difficult to implement very complex questionnaires online compared to face-to-face or mail surveys. Furthermore, web surveys are often not representative for a study population (e.g. citizens of a region or country) but only for the population with Internet access. Yet, this also depends on the survey panel provider, where some providers make more effort than others to represent the population at hand as closely as possible (see Sect. 4.1). Studies comparing different survey modes in stated preferences studies on environmental valuation indicate that web surveys reveal similar results as other survey modes, especially regarding willingness-to-pay values; however, once again, the presence of survey mode effects can depend on the Internet panel provider used (Olsen 2009; Lindhjem and Navrud 2011).

Choosing the survey mode for a DCE is an important decision that has to be considered in the planning process of a study and when applying for research funding. Often the survey costs are a (or the) main driver of this decision. While computer-assisted face-to-face interviews might have advantages in terms of sample representativeness, they are more expensive than web surveys. The latter have many advantages for DCE; yet, it is important to carefully select web survey panel providers and to examine how they select their panel members (e.g. whether they are recruited by telephone and it

is a managed panel, or whether it is an opt-in panel where everyone can participate and the panel provider does not have a clear overview of who is taking part, see Sect. 4.1). Lastly, as stated at the beginning of this section there might be research contexts such as in developing countries where computer-assisted face-to-face interviews are the only method that can be applied on practical grounds. A more detailed discussion of survey modes is provided for example by Dillman et al. (2008) and Leeuw et al. (2008). Menegaki et al. (2016) offer noticeable insights and guidance for web surveys in the context of DCE.

References

- AAPOR (2016) Response rates—an overview. Education/Resources, American Association for Public Opinion Research. <https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>. Accessed 8 May 2020
- Bateman IJ, Day BH, Georgiou S, Lake I (2006) The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. *Ecol Econ* 60:450–460. <https://doi.org/10.1016/j.ecolecon.2006.04.003>
- Bonnichsen O, Olsen SB (2016) Correcting for non-response bias in contingent valuation surveys concerning environmental non-market goods: an empirical investigation using an online panel. *J Environ Planning Manage* 59:245–262. <https://doi.org/10.1080/09640568.2015.1008626>
- Couper M (2000) Review: web surveys: a review of issues and approaches. *Public Opin Q* 64:464–494. <https://doi.org/10.1086/318641>
- Couper MP, Antoun C, Mavletova AM (2016) Mobile web surveys: a total survey error perspective. In: Total survey Error in practice: improving quality in the era of big data. <https://publications.hse.ru/en/chapters/191682818>. Accessed 13 Aug 2019
- de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA (2015) Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *Patient* 8:373–384. <https://doi.org/10.1007/s40271-015-0118-z>
- Dillman DA, Smyth JD, Christian LM (2008) Internet, mail, and mixed-mode surveys: the tailored design method, 3rd Revised. Wiley, Hoboken, NJ
- Emerson MO, Yancey G, Chai KJ (2001) Does race matter in residential segregation? Exploring the preferences of White Americans. *Am Sociol Rev* 66:922–935. <https://doi.org/10.2307/3088879>
- Glenk K, Johnston RJ, Meyerhoff J, Sagebiel J (2020) Spatial dimensions of stated preference valuation in environmental and resource economics: methods, trends and challenges. *Environ Resource Econ* 75:215–242. <https://doi.org/10.1007/s10640-018-00311-w>
- Leeuw EDD, Hox J, Dillman D (2008) International handbook of survey methodology, 1st edn. Routledge, New York, London
- Liebe U, Moumouni IM, Bigler C et al (2020) Using factorial survey experiments to measure attitudes, social norms, and fairness concerns in developing countries. *Sociol Methods Res* 0049124117729707. <https://doi.org/10.1177/0049124117729707>
- Lindhjem H, Navrud S (2011) Using internet in stated preference surveys: a review and comparison of survey modes. *Int Rev Environ Resource Econ* 5:309–351. <https://doi.org/10.1561/101.00000045>
- Loosveldt G (2008) Face-to-face interviews. Routledge Handbooks Online
- Louviere JJ, Hensher DA, Swait JD (2000) Stated choice methods analysis and applications. Cambridge University Press, Cambridge
- Lozar Manfreda C, Vehovar V (2008) Internet surveys. In: de Leeuw ED, Hox JJ, Dillman DA (eds) International handbook of survey methodology. Psychology Press, New York

- Menegaki AN, Olsen SB, Tsagarakis KP (2016) Towards a common standard—a reporting checklist for web-based stated preference valuation surveys and a critique for mode surveys. *J Choice Model* 18:18–50. <https://doi.org/10.1016/j.jocm.2016.04.005>
- Mills C (2014) The Great British Class Fiasco: a comment on Savage et al. *Sociology* 48:437–444. <https://doi.org/10.1177/0038038513519880>
- Olsen SB (2009) Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. *Environ Resour Econ* 44:591–610. <https://doi.org/10.1007/s10640-009-9303-7>
- Rose JM, Bliemer MCJ (2013) Sample size requirements for stated choice experiments. *Transportation* 40:1021–1041. <https://doi.org/10.1007/s11116-013-9451-z>
- Savage M, Devine F, Cunningham N et al (2013) A new model of social class? Findings from the BBC's Great British Class survey experiment. *Sociology* 47:219–250. <https://doi.org/10.1177/0038038513481128>
- Steeh C (2008) Telephone survey. In: de Leeuw ED, Hox JJ, Dillman DA (eds) *International handbook of survey methodology*. Psychology Press, New York
- Yeager DS, Krosnick JA, Chang L et al (2011) Comparing the accuracy of RDD telephone surveys and internet Surveys conducted with probability and non-probability samples. *Public Opin Q* 75:709–747. <https://doi.org/10.1093/poq/nfr020>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

