Sabine Häder
Michael Häder
Mike Kühne
Editors

Telephone Surveys in Europe

Research and Practice

Springer

Telephone Surveys in Europe

Sabine Häder, Michael Häder, and
Mike Kühne (Eds.)

Telephone Surveys in Europe

Research and Practice

Springer

Editors
Dr. Sabine Häder
GESIS – Leibniz Institute for
the Social Sciences
P.O. Box 12 21 55
D-68072 Mannheim
Germany
E-mail: sabine.haeder@gesis.org

Prof. Dr. Michael Häder
Technische Universität Dresden
Faculty of Philosophy
Institute of Sociology
Helmholzstr. 10
D-01069 Dresden
Germany
E-mail: michael.haeder@tu-dresden.de

Dr. Mike Kühne
Technische Universität Dresden
Faculty of Philosophy
Institute of Sociology
Helmholzstr. 10
D-01069 Dresden
Germany
E-mail: mike.kuehne@tu-dresden.de

ISBN 978-3-642-25410-9 e-ISBN 978-3-642-25411-6
DOI 10.1007/978-3-642-25411-6
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2011941490

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Part III: Weighting and Nonresponse

Part IV: Data Quality

Part V: Recommandations for Telephone Surveys

Introduction: Telephone Surveys in Europe

Sabine Häder, Michael Häder, and Mike Kühne

In the preface we would like to elaborate on three questions: First of all, what motivates us to publish a book dealing with telephone surveys? Secondly, why does this book on telephone surveys focus on Europe exclusively? And finally, which particular goals do we pursue with our book?

The Importance of Telephone Surveys

In the course of dealing with telephone surveys as a data collection instrument in the social sciences, two things attract one's attention: First, the great popularity of this method in the survey research of the past decades, and second, the immense need for more methodological research despite the widespread use of this method. Both phenomena will be briefly discussed.

Table 1 shows a comparison of the research turnover (in terms of money) across various European countries and across various modes, including telephone interviews. One should note that, in general, telephone surveys are considered to be a particularly inexpensive instrument, which in consequence makes it difficult to generate a high turnover.

Table 1 Total research turnover across European countries, only quantitative research, in percent, difference to 100 in line: non-quantitative research

Country	Mail	Telephone	Face to face	Online	Other
Austria	3	20	22	8	2
Belgium	1	32	32	9	3
Croatia	1	23	33	1	31
Czech Republic	1	18	44	5	14
Denmark	9	18	9	24	28
Finland	5	40	5	33	3
France	2	12	15	12	46
Greece	1	18	41	1	28
Hungry	4	16	45	14	5
Latvia	0	17	41	5	22
Netherlands	5	18	11	25	15
Norway	8	39	10	23	3
Poland	0	11	35	10	26
Portugal	0	19	21	3	38

Table 1 (*continued*)

Country	Mail	Telephone	Face to face	Online	Other
Russia	1	21	35	5	13
Slovak Republic	1	14	42	2	24
Slovenia	8	27	23	4	14
Spain	2	26	27	11	15
Sweden	12	36	7	22	12
Switzerland	5	49	21	10	1
UK	8	17	21	18	23
Weighted total	5	16	10	20	33

Source: ESOMAR (2008): Global Market Research, ESOMAR Industry Report, Amsterdam

This table demonstrates the appeal of telephone surveys in Europe. However, this table does not exhibit the methodological challenges associated with this survey mode; these challenges justify further research into the telephone method. A few of them are mentioned here briefly:

- First, there is the role of the interviewer. In telephone interviews, the interviewer only vocally communicates with the respondent. In contrast, in face-to-face interviews a more complex social relationship is established between interview partners with different consequences for the response quality.
- Efforts made to contact and motivate the selected persons to give an interview differ across survey modes. The relatively large number of contact attempts brings about the good quality of telephone interviews. However, the use of incentives is limited.
- Supervision procedures during field work differ across survey modes. In telephone laboratories there is strict supervision. This contrasts with a relatively independent and unobserved operation of interviewers in face-to-face fields.
- In face-to-face surveys the sample design is often regionally clustered for research-economic reasons. In contrast, telephone surveys are generally set up un-clustered and, consequently, the estimators are more precise. However, in landline surveys a random selection within the target household is necessary.
- In telephone interviews all communication is aural, which means that it is conducted through only one channel. Whereas in face-to-face interviews lists of answer options and other materials can be presented visually as well as read out to the participants, this option, which often makes the interview more interesting, is not available on the phone.
- The aural nature of the interview has consequences for the design of the questionnaire, which has to be carefully considered accordingly. The presence of third persons during the interview can affect the behavior of respondents. For telephone interviews this should be less the case than for face-to-face interviews, because in the latter case there is a danger of third parties interfering with the interview.

Due to the advantages mentioned and despite the presented problems telephone surveys are an especially attractive tool for the collection of social scientific information. Therefore, they are - in comparison to other designs for standardized interviews - also used very often. According to the spotlights in this volume this is true for almost all European countries. In some countries, telephone surveys are by far the most important way to collect data. The challenges faced in telephone interviews are not only very numerous, in addition, new problems arise. This is especially true if one compares them with other modes, such as mail surveys. Especially technical developments in the telecommunications sector such as the spread of mobile communication, changes in tariff structures across the various providers, and data protection concerns lead to a need for further developments in the telephone survey instrument. A book that reports both the state-of-the-art and the latest research findings on telephone surveys should thus be of great interest to the social science profession.

Research and Practice in Europe

In empirical social research the problem of non-response as well as rising costs seem to be omnipresent. The latter is especially a concern for face-to-face interviews. Europe is already united in searching for solutions to these problems. However, such solutions are difficult to find. Maybe in future the increased use of mixed-mode designs will offer such solutions. The trend shows that the popularity of telephone surveys, and of surveys in general, is still increasing, thus underscoring the importance of this instrument again. We view this as indication that telephone surveys – as sole survey mode or within a mixed mode design – will continue to occupy an important place in the social science methods toolbox.

As noted earlier, the conditions for telephone surveys have developed very differently across different European countries. The solutions for dealing with the (national) problems found in each country are very different. This may for example apply to strategies that were developed across Europe to respond to the spread of mobile-only households and the adoption of mobile phone surveys. Some countries may have a certain pilot function for coping with these problems. Problems for which solutions or routines already exist in one country, sooner or later become acute in other countries. Thus, solutions of one country could become a model for other European countries.

The interest of social scientists, politicians, and market researchers goes beyond the situation in individual European countries. Studies such as the European Social Survey, the European Values Study, and the International Social Survey Program suggest this. So researchers have to carefully analyze the conditions under which a specific solution in a specific country was found. This book makes a contribution to describing the empirical conditions faced across countries.

In the United States, researchers such as Lepkowski et al. (2008) and Groves et al. (1988) have dealt with telephone interviews in comprehensive books. However, naturally the center of attention in these publications is America and not Europe. We would like to close this gap with our book.

Objectives and Content of the Book

The chapters of the book provide insight into the state of the art in very different areas of telephone surveys. The first part of the book gives spotlights on the use of landline and mobile phone surveys across various countries in Europe. Researchers from national scientific institutions and market research give an overview of sampling frames, the use of mobile and landline phones, response rates, as well as the resulting strategies for telephone surveys in their countries.

In Russia we can currently observe an extremely quick growth in the popularity and penetration of mobile phones. This delayed development in comparison to other European countries and the current penetration rate of landline phones also has an impact on survey practice. It is not assumed that the penetration of landline phones will ever reach a level where national surveys by landline phones could be justified. It is more likely that in Russia other “post-telephone” methods of data collection will be implemented, especially mixed mode designs. Anna Andreenkova reports on the situation in Russia.

The size of the Netherlands and the population density play a major role in the data collection process. In the Netherlands, as a very small country where the people live closely together, face-to-face interviews were affordable as a mode of data collection until the recent past. This led to a relatively late use of CATI. At the moment telephone surveys are carried out predominantly via landline phones. This is mainly due to a lack of quality sampling frames for mobile phones and the interview situation, which is often not suited for long interviews. This spotlight is presented by Dirkjan Beukenhorst.

For telephone surveys in Switzerland topographical features are relevant. Compared to other European countries, Switzerland has excellent landline telephone provisions, because of a large amount of hardly accessible areas. Telephone surveys are the most reliable and cost-effective method for both market and scientific research. Nevertheless, not only coverage errors but also over-surveying are seen as the two major problems. Michèle Ernst Stähli wrote this report.

Compared to the vast majority of European countries in Finland mobile phones started to become popular relatively early. For this reason, the methodological research on mobile phone issues has a long tradition. One of the most important results is the finding that mobile phones do not seem to disturb interviewing more than landline telephones. The spotlight of Finland is presented by Vesa Kuusela.

In Great Britain the acceptance of telephone surveys differs across market research and public sector research. Although the telephone penetration in the UK would be adequate for telephone interviews, the UK government continues to prefer face-to-face interviews. Concerns about the data quality of telephone interviews helped launch quality standards like interviewer quality control schemes and the guidelines for conducting telephone surveys of the American Association for Public Opinion Research. Liz Nelson wrote the spotlight on Great Britain.

Today coverage error is the biggest issue for phone surveys in Italy. The resulting data quality of telephone surveys is often a cause for concern. Because of that future developments are of great interest. The worst-case scenario describes an expansion of low quality low cost phone surveys. The more optimistic scenario considers a combination of methodological and logistics advantages of phone surveys on probabilistic grounds and the use of mixed mode designs. Teresio Poggio and Mario Callegaro present the state of the art concerning telephone surveys in Italy.

In Portugal a parallel trend of decreasing landline phone coverage and increasing mobile phone ownership can be observed. This is an important stimulus for survey companies using mobile phones either as a supplement or as an alternative to landline phones. Because the mobile phone operators do not provide the necessary information for probabilistic sampling frames the current challenge in Portugal is building a sampling frame. This spotlight is written by Paula Vicente and Elizabeth Reis.

The second part of the book addresses the issue of sampling for telephone surveys. Survey researchers present results of current research projects, especially problems with under-coverage as well as findings on new sampling frames.

The results presented in the chapter by Dominique Joye, Alexandre Pollien, Marlène Sapin and Michèle Ernst Stähli show clearly that people without landline phones are different from those with landline connections in terms of age, education, and their personal situation measured through household composition. However, the differences are even stronger for opinion variables, e.g. political interest or party preference. These findings are stable across different surveys and seem not to be influenced by nonresponse bias.

The sampling frames and parameters for five European countries presented in the chapter by Christiane Heckel and Kathrin Wiese show how important combining landline and mobile phone samples has become. Otherwise the coverage error because of missing mobile-only households can lead to biased samples. However, this happens at the expense of regional stratification, since mobile phone numbers do not indicate in which region the subscriber lives. Moreover, mobile phone surveys are more expensive than landline phone surveys. Thus, costs and benefits must be weighed up with the client for each CATI survey.

In their chapter Götz Schneiderat and Tino Schlinzig argue that response rates are an essential but not sufficient indicator for the quality of a telephone survey. The quality of the realized sample may be even more important. Differences with respect to demographic characteristics between subpopulations using mobile phones and/or landline telephones require separate data collection methods to diminish possible coverage errors. Altogether, the data of the CELLA survey suggest that the use of a dual-frame approach leads to a better coverage than the traditional landline phone survey.

The third part of the book addresses the issue of weighting and nonresponse. Sampling experts discuss problems such as refining the dual frame approach and adjustment strategies for nonresponse in telephone surveys.

The chapter by Siegfried Gabler, Sabine Häder, Iris Lehnhoff and Elisabeth Mardian explains the importance of design weighting and an additional adjustment weighting for dual frame telephone surveys. It can be demonstrated that weighted data of socio-demographic variables are frequently closer to the population distributions than those without weighting. Thus, the most problematic drawbacks of biased samples can be avoided if adequate weighting is applied.

In the next chapter by Fannie Cobben, Barry Schouten and Jelke Bethlehem the bias caused by non-observation errors in a telephone survey is analyzed. Various weighting techniques that aim to adjust for errors caused by under-coverage or nonresponse, and for under-coverage and nonresponse errors simultaneously are compared. These techniques are partly based on the generalized regression estimator, and partly on the use of propensity scores.

The chapter by Oliver Lipps and Kathrin Kissau analyzes which socio-demographic subpopulations are over- or underrepresented in a telephone survey where the sample was drawn from a population register in Switzerland. First, they examine which bias occurs if different sources of telephone numbers for matching names and addresses to the sample are used. Second, the authors check socio-demographic differences in the final response status and different reasons for refusals.

The fourth part contains chapters on data quality. For instance, this part discusses whether or not different modes of data collection (mobile versus landline phone surveys) have an impact on the response quality.

The likelihood of measurement error depends on characteristics of the respondent and on characteristics of the survey, such as the mode of data collection, the survey procedures, and the design of survey questions. Peter Lynn and Olena Kaminska look into factors affecting such measurement errors in mobile phone interviews. In their chapter they provide a structured framework that identifies the set of features of a mobile phone interview that can influence measurement error.

Mike Kühne and Michael Häder emphasize the need for dual frame designs that combine the use of telephone surveys via landline phones and mobile phones. The focus lies on particular aspects of the interview situation. Subsequently, existing differences are presented. On the basis of their analysis, it can be said that controlling for relevant aspects of the interview situation the parallel use of landline phones and mobile phones in the same survey does not lead to major differences in response behavior.

In the subsequent chapter Michael Häder first proposes a theoretical model that can be used to explain response quality. This is followed by considerations on mode effects, particularly those that may be expected in telephone interviews in comparison to other modes of administration. Finally he turns to methodological aspects influencing the response quality in a study using landline and mobile phones in parallel. He presents recent findings of empirical research using the combined landline and mobile phone survey CELLA 2.

The last part of the book gives recommendations based on the theoretical considerations and empirical findings presented in the previous chapters. It offers suggestions and advice on how to conduct surveys and polls in both academic and market research.

Response rates in market and social surveys have been continuously declining. With this development in mind, Gerd Meier looks for effective techniques to stop this decline. In two experiments with different designs he shows that a question or information within the introduction of an interview leads to a higher cooperation rate and hence reduced refusals.

Vasja Vehovar and Ana Slavec argue that dual frame surveys combining fixed and mobile telephone frames can be a solution for the growing problem of under-coverage due to “mobile-onlys”. To take full advantage the optimal mixture parameter of the mobile and fixed telephone subsample needs to be formulated. An analytical optimization approach is used to maximize accuracy and minimize costs in order to evaluate the behavior of this mixture parameter.

The aim of the study by Patrick Schmich and Franziska Jentsch is to discuss selected parameters for the fieldwork management of a telephone survey. Several factors are presented that influence the composition of the completed sample. The authors concentrate on criteria that are relevant and important in daily fieldwork. Especially factors such as the number of calls, the effect of the calls after a “soft refusal”, and the choice of contact times are analyzed and described.

This book was written for all scientists and practitioners who deal with theory and application of telephone surveys in academic and market research. Furthermore, it is a valuable reader for students who are interested in survey methodology.

We would like to express our sincere gratitude to Dr. Annelies Blom (Survex – Survey Methods Consulting[1]) for her impressive work as copy-editor and her valuable methodological advice. Without her help this book would not be the same.

We are also very grateful to Sandra Herrmann (Technical University Dresden) who made fundamental contributions to the technical completion of this volume.

Lastly, we offer our regards and blessings to all those who supported us in any way during the course of this book project.

References

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls II, W.L., Waksberg, J.: Telephone Survey Methodology. Wiley, New York (1988)

Lepkowsky, J.M., Tucker, C., Brick, J.M., de Leeuw, E., Japec, L., Lavrakas, P.J., Link, M.W., Sangster, R.L.: Advances in Telephone Survey Methodology. Wiley, New York (2008)

---

[1] see www.survex.org

Part I
Spotlights on Selected European Countries

Chapter 1
Russia

Anna Andreenkova

1.1 Introduction

For the last twenty years face-to-face interviews in respondents homes have been the dominant method of interviewing the general population on a national level in Russia. Furthermore, the importance of this method is not diminishing yet, even when taking into account all difficulties and complications related to it: increasing costs, relatively long interviewing periods, respondents' concerns regarding security (allowing strangers into their home), decreasing response rates, the need of travelling on long distances and high communication costs in such a large country. Face-to-face interviewers and telephone survey methodology have both been introduced in Russia at almost the same time – at the end of 1980s. This coincided with the start of social-economic reforms, the "Gorbachev era" of openness and free circulation of information ("glasnost"). Since then surveys of general population became an important part of public life and a widely-used method in the social sciences.

Regardless of all the limitations, the number of telephone interviews is growing exponentially each year. Nowadays telephone surveys are one of the most widely used method for regional surveys, mainly in largest metropolitan areas of Moscow and St. Petersburg but also in other large cities in the country. It is also used as the main method of interviewing of special groups of the population – elite, professionals in different areas (health providers, businessmen, experts, NGOs, etc.) or the population with only telephone contact information, for example consumers of particular products, users of particular services, members of some organizations.

The main limitation of telephone interviews as a data collection method for national surveys in Russia is the low telephone penetration. About half of the total population would be excluded from the sampling frame for a national telephone survey because of this reason. And even this penetration level was reached only in the last ten years. Of all countries of the former Soviet Union (with the exception of the Baltic countries) only Belarus and Moldova had the telephone penetration of above 80 % in both urban and rural areas. In all other countries the telephone penetration is below 50 % nationally and below 40 % in rural areas.

The landline telephone penetration is not just low but also unequally distributed within the population, and this has major impact on the results of most surveys. The main predictor of telephone penetration is the level of urbanization of the place of residence which in turn is explained by the history of the development of telephone communications in Russia: large cities were the first to receive telephone connection. Telephone penetration in the two largest cities of Russia – Moscow and St. Petersburg – has exceeded 90 % since the beginning of 1990s. In the past ten years it has reached 80 % in almost all large cities of the country (defined as cities with population of one million people or more – 13 of such cities in Russia accounting for 18 % of total population of the country). However, only 35 % of the households in rural areas have a landline telephone. Access to landline phones are also related to many other factors: education, socio-economic status, type of housing and area of residence, composition of the household, and also attitudinal characteristics of respondents, exposure to mass media, usage of Internet, interest in politics and even electoral participation.

1.2 Historic Overview

The end of 1980s was the time of liberal reforms in the USSR starting from "glasnost", a policy of openness and freedom in the flow of information proclaimed by Gorbachev. Before this time the number of surveys among the general population was very limited. Surveys were used for academic purposes only or as information support for ideological divisions of the state authorities. The prevailing method of data collection was self-administered questionnaires in public places (for example, in working place), sometimes mail surveys.

Political and ideological liberalization coincided with a growth of expertise, knowledge and technical skills of scholars and also high enthusiasm and devotion of a new generation of researchers to obtain reliable information from the population based on probability samples. The major methodological work on designing the national and regional samples was concentrated in the Institute of Sociology USSR's Academy of Sciences. The interest to public opinion in Russia was very high not only within the country, but also from outside. Western academic institutions, the international press and mass media of such countries as the USA, Germany, and France played an important role in stimulating the introduction and usage of population surveys based on random samples. The New York Times, the German public television channels ARD and ZDF, German periodical Der Spiegel and the French newspaper Le Figaro were initiators of the early telephone surveys on public opinion in Russia. The reputation and positive attitudes of these media contributed to a gradual decrease of skepticism and distrust in telephone surveys among the Russian academic community, the Russian media and even the state authorities. Most early concerns were not as much about technical difficulties attributed to telephone surveys (coverage, sampling issues, etc.), but concerns about measurement biases: lack of sincerity, fear of respondents to provide true information by phone (in a country with strict state control over means of communication and the threat of being overheard and recorded). Russian respondents surprised researchers being in relatively open and even radical in their answers and readiness to give their opinion, at least on the most commonly discussed public issues. Russian media followed the example of respected Western media rather quickly and after period of reprinting Western articles on telephone surveys on Russian public opinion, they became initiators and clients of their own polls.

The growth of population surveys in the USSR – for both personal and telephone interviews – is closely related to the development of comparative projects and comparative methodology. The first telephone survey which we can find in data archives and publications was a comparative study conducted in October 1987 in Russia and France. The purpose of the project was to study public opinion of in the USSR and France on different international issues and bilateral relations of USSR and France[1]. The survey was also designed as an experiment of data collection modes – telephone interviews were conducted with random sample of residents of Moscow city and face-to-face interviews with respondents in the small town of Inzhavino in the Tambov region (Central Russia). The success of this project broke the wall of prejudices around usage of telephone surveys in Russia, because the level of cooperation of respondents was relatively high (above 70 %) and sampling issues had been solved rather effectively[2].

The number of telephone surveys on public opinion in Russia grew very fast. Only by the end of 1980s more than a dozen of different telephone surveys had been conducted. Most of them were focused on studying public opinion on issues of international relations. The new politics of openness and the later breakdown of the Soviet Union led to reconsiderations in Russia's relations with many countries, the need to find a new place for Russia as an independent state in the world arena[3].

---

[1] In Russia the survey was conducted by the telephone surveys group within department of survey research methodology of the Institute of Sociology, USSR Academy of scientists. Principal investigators were Dr. Vladimir Andreenkov and Dr. Valery Mansurov. In France the survey was conducted by the company IPSOS.

[2] At the time, there was no publicly available data on private telephone numbers in the USSR (the telephone number of a particular person could merely be obtained in the public Information Center, if the name and some other information about the person were known). However, the Moscow state telephone company, the monopolist in the area of telephone services in Moscow in those years, was persuaded to cooperate with Russian scientists. A special software was developed for drawing the random sample of private telephones from the general city database of telephone numbers. This system survived till today, although the state telephone company is not a monopolist in telephone services anymore and covers only 80 % of private telephones in Moscow city. Unfortunately, in other cities all attempts to create similar systems failed because of financial, organizational and other reasons.

[3] The examples of such telephone polls which were broadly broadcasted in the international media are listed below. The topics of the projects remind us of the major events in the modern history of Europe which we all witnessed and have been part of. "Public opinion of Russians about the results of the meeting between M. Gorbachev and R. Reagan" December 1987, principal investigator- E. Bashkirova; "Public opinion of Muscovites before M. Gorbachev's visit to Great Britain"; "Muscovites about the results of the visit of R. Reagan"; "Public opinion about the visit of Helmut Kohl to the USSR"; "Muscovites about the Olympic Games in Seoul", 1988; "Public opinion of Muscovites about international relations before the visit of M. Gorbachev to FRG", 1989 (principal investigator of all these projects – Dr. Vladimir Andreenkov). The field work for all these surveys was conducted by the department of survey research methodology of the Institute of Sociology, USSR's Academy of Sciences.

From studying public opinion on international issues, researchers shifted to the study of opinion on internal politics and social issues[4]. Most telephone surveys at the end of the 80s were conducted in Moscow city, but some also in the capitals of other Soviet republics: in Tbilisi, Tallinn, Almata, Kiev etc.

From the end of the 1980s and beginning of the1990s onwards we witnessed the formation of organizational structures for conducting mass surveys – face-to-face and telephone – first in academic institutions. In 1988 the field group for telephone surveys was organized in the Institute of Sociology in the department of survey research methodology headed by Dr. V. Andreenkov. Later there emerged specialized survey organizations (like the state-run VSIOM founded in December 1989) and independent survey companies (like CESSI in Russia and Saar Poll in Estonia founded in 1988 - 1989, ROMIR and Vox Populi in 1990, Baltic Data House in Latvia, BRIF in Kazakhstan, FOM and Comcon in Russia followed by Baltic Surveys in Lithuania, Novak in Belarus, KMIIS and SOCIS in Ukraine in 1991 - 1992).

The next decade was the time of the gradual introduction of telephone surveys into everyday research life, shifting from one-off specialized surveys to regular and routine research. Even if telephone interviews were not used as a mode of data collection in national surveys, the field of application for this interviewing method was very broad. Telephone surveys were accepted as legitimate method in marketing research to study consumer behavior and preferences in different cities and metropolitan areas, measuring the effectiveness of advertising on the local TV and radio, as well as poster advertising. The second important field of application of telephone studies since the 1990s is the study of public opinion on local issues and electoral studies on regional elections (before the ban of direct elections of the head of the regions by popular vote in 2004) and in national elections, measuring the effectiveness of political advertising and public opinion on different city events. Already from the beginning of the 1990s telephone surveys became one of the main data collection methods for the measurement of media preferences (surveys of TV and radio audiences) in Russia.

During early introductions of telephone surveys in Russia, the telephone was mostly a means of relatively short interpersonal communication and a way of obtaining short factual information from different sources – government agencies, clients, work colleagues, etc. Telephone interviews for survey purposes entered the life of Russians as a continuation of this, for communication on important issues and for providing information. In recent years, the variety in telephone surveys substantially increased. In addition, they are used for marketing purposes (sales, advertising of products and services, information about new services). This led to a much higher exposure of telephone calls from strangers to home telephones and consequentially lowered the response rate in telephone interviews. Broader functions of telephone calls in 2000s led to establishing new forms of organizational structures – multi-function call-centers. Initially, the main purpose

[4] Survey "Residents of Moscow about issues of perestroika", 1988, principal investigator – V. Andreenkov; several telephone surveys in Moscow city before and during the I Congress of People's Deputies (reformatted Soviet Parliament) in 1989 - 1990 with participation of a number of leading social scientists of that time.

of call-centers was accepting calls as phone social services agencies, client services of different companies, large national information and advertising campaigns by phone or sales by phone. However later, some of them also offered the conducting telephone surveys. Currently, a substantial proportion of the telephone surveys in the country are conducted not by specialized survey research or marketing research companies, but by multi-function call-centers. The advancement of technology also played an important role in shaping the methodology of telephone surveys.

The following trends can be named in the current status of telephone surveys in Russia:

- Increase of geographical coverage of telephone surveys – from surveys just in Moscow and St. Petersburg to other large cities in the country; now many projects are done based on a sample of the urban population of middle-size cities and larger.
- Switching from single-city surveys to broader geographical coverage; also the decrease in the costs of inter-city telephone calls in recent years made it possible and effective to conduct telephone interviews from a central location in all geographical units. Recently, regionally-based interviewing prevails, which limits the advantage of centralized quality control during the actual field work and the usage of CATI systems.
- Nowadays most telephone surveys are conducted using different CATI-systems, although the introduction of such systems in regular research practice has taken place much later than in many other European countries.
- Telephone surveys finally got wide public acceptance as a trustworthy and reliable method of data collection. State authorities have admitted telephone surveys as valid mode of measuring the effectiveness of different government programs[5].
- Telephone surveys have become a “norm”, standard method of interviewing of special business groups, elites and professional groups. The culture of telephone interviews instead of personal conversation face-to-face has not come quickly. For many years getting cooperation of experts and business groups by phone has been a very difficult task. Personal visits played the symbolic role of respect and attention and individual approach to a particular respondent which is so important for busy high-ranking people. After some intermediate years of using mixed mode methods for such interviews (respondents were asked for their preferred type of interview – telephone or personal), now B2B surveys are mostly conducted only by telephone with business groups. Also a majority of interviews with expert groups are conducted by phone. Nevertheless, even today state authorities in some regions prefer to be visited and asked for the interview personally rather than on the phone.
- One of the major barriers to a wider use of telephone surveys was the argument that telephone interviews can be used only to obtain information on rather

[5] For example, in Moscow and in St.-Petersburg the city government conducts monthly telephone surveys among residents of these cities to study general social-economic situation in the city, public attitudes toward government’s policies, satisfaction with city programs and actions.

simple, straightforward questions and cannot be used for any complicated or sensitive topic. In recent years we have seen more and more publications on surveys on such sensitive topics as health, AIDS, etc. based on telephone surveys although the methodological background of such surveys are not commonly discussed.

- Regardless of all these advancements of telephone surveys, they are still not used as a method of data collection for national population surveys. They are hardly used for any sample survey of State Statistical Agency (mixed mode – face-to-face and telephone interviews – are used in the Survey of Small and Middle-Size Business, but not for population). They are also not used in any large academic projects.

Several basic requirements should be available and met when conducting reliable high quality telephone surveys. First of all, technical requirements – adequate telephone coverage among the target population, adequate quality of telephone connection and the lack of financial pressure to respondents for in-coming phone calls. Second type of requirements is availability and access to information for drawing a random sample from all units of the target population. And the last are cultural and normative requirements which do not prevent open conversations with a stranger on the phone for respondents and general trust in telephone surveys among data users. All these background requirements cannot appear immediately but tend to develop and change with time, sometimes stimulating and sometimes preventing the development of telephone interviews. Let us consider one of the major issues for telephone surveys in Russia as well as in many other countries: the telephone coverage of the target population.

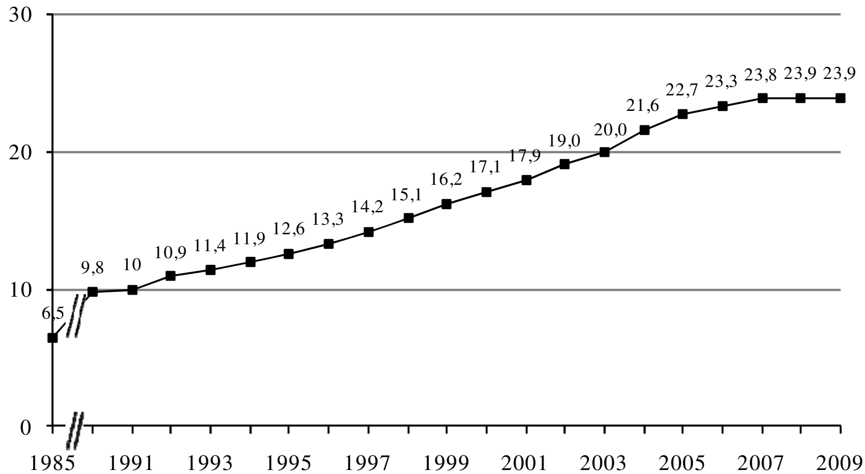
1.3 Technical Requirements for the Organization of Telephone Surveys in Russia

Low telephone penetration is the major barrier for using telephone surveys in Russia as well as in almost all other countries of the former Soviet Union.

Telephone communication in the USSR was not a priority for Soviet state, whose policy was to first develop postal and telegraph connections. (Until the last year of the Soviet Union it was the world leader on number of telegrams sent each year, which tells a lot about the importance of this means of communication in the Soviet Union.) Therefore, the telephone system was far behind most advanced European countries and the USA. For many years telephone connections were only local and not connected into inter-city or international lines. The first automatic international connection with countries in Eastern Europe was introduced in 1987 only, with the USA in 1991 and with Great Britain in 1992. The country went through a shortage of telephone lines. The demand for private telephones was much higher than the supply, and people had to wait for their telephone connection for years, even in large cities.

In the middle of 1980s there were only 6.5 private telephones per 100 people in Russia. Unfortunately, the State Statistical Agency does not provide any estimates of telephone penetration in households. Therefore, we have to rely on survey data to obtain such numbers.

For the last 25 years the number of private telephones and therefore telephone penetration in households grew dramatically (see Figure 1.1). While in the year of the breakdown of Soviet Union (1991) it amounted to ten telephones for 100 people, in the next ten years it doubled. However, in the last five years, for the first time in Russian history the number of private landline telephones stopped increasing and in 2009 it even decreased (although to a very small extent)[6].



Source: data of Russian State Statistical Agency and Ministry of Communications, 2010

Fig. 1.1 Number of private stationary phones for each 100 people in Russian Federation

Across all countries of the former Soviet Union (except for the three Baltic countries, where we do not have comparable data), Russia occupies the third place regarding the number of telephones per capita. (Table 1.1 shows the number of all phones including non-private, but private phones constitute the prevailing majority. Unfortunately, other comparative estimates are not available).

Belarus is the leader regarding telephone coverage: 43.6 telephones for 100 people and according to survey data household penetration exceeds 80 %. Such results were reached in the last ten years, as the number of landline phones doubled in Belarus and continues to grow. Moldova is in second place with 32 telephones for 100 people, which is also almost double compared to ten years ago. In other countries the growth in the number of telephones is not so impressive. The Ukraine reached only 28.5 telephones for 100 people, although in 2000 they

[6] We cannot say whether this is the beginning of a stable trend of decline in the number of landline phones in Russia or just a temporal fluctuation. The reasons for this are a constant increase in landline telephones for private people (and an even stronger increase for companies), a gradual switch from a monthly fixed fee to paying per minute and also the growth of mobile phone usage.

started only slightly lower than in Belarus or Russia (see Table 1.1). In Kazakhstan we find 24.1 phones per 100 people due to its low starting position. In Trans-Caucasian and Central Asian countries telephone penetration has been extremely low ever since Soviet times and increased only moderately in recent years to less than twenty telephones for 100 people in Trans-Caucasian countries and less than ten telephones per 100 people in Central Asian countries.

Table 1.1 Number of telephones (private and non-private) for each 100 people

	2000	2005	2007	2008	2009
Russia	22.6	30.0	31.8	32.1	32.0
Azerbaijan	9.9	12.8	14.4	14.7	15.6
Armenia	16.4	18.8	19.3	19.3	19.0
Belarus	29.6	35.9	38.5	41.2	43.6
Georgia	13.0	13.0	13.9	15.4	…
Kazakhstan	13.8	17.8	20.8	22.8	24.1
Kyrgyzstan	7.9	8.6	9.2	9.4	9.3
Moldova	16.6	26.2	30.3	31.3	32.0
Tajikistan	3.5	3.9	4.0	4.0	3.9
Uzbekistan	6.7	6.8	6.9	6.9	NA
Ukraine	20.0	25.4	27.9	28.7	28.5

SOURCES: Inter-state Statistical Committee of CIS, 2009. Data on Turkmenistan is not available, but it is close to the lowest numbers.

An adequate quality of the telephone connection was also only reached recently by most of post-Soviet countries. The telephone equipment used in the beginning of 1990s was very much out of date and automatic intercity connections were very limited. All of this led to a low effectiveness of telephone calls. Telephone lines were breaking down frequently, such that it was difficult to differentiate numbers that could not be reached from non-existing (empty) numbers. Such problems still play a role in telephone interviewing even today.

General low coverage of telephones is only one problem, but it is intensified by the unequal coverage of the different units. The major dividing factor is the level of urbanization of the area (see Table 1.2a and 1.2b in the appendix). In 2009 the number of telephones per 100 people in Moscow was 35.7, in St. Petersburg 38.8 and in other urban areas 28.2, but in rural areas it was with 11.7 almost half.

Although the level of urbanization is the key factor in access to a landline telephone, non-coverage of landline phones is also related to several important socio-demographic and economic factors: education, type of housing and household composition. In Russia access to landline phones at home is related not only to some socio-demographic factors and area characteristics, but also to some attitudinal and behavior factors which should be taken into account when planning a survey on a particular topic and when choosing the data collection method. People with access to landline phones show higher exposure to traditional media (radio, newspapers and to some extent also TV) than groups without landline phones. They have more interest in politics and participate in national elections.

In the last two years, the penetration of mobile phones has been much higher than that of landline phones and it continues to grow. In addition, mobile phone penetration is much more equally distributed among socio-demographic groups and does not depend on the type of residence as much as landline phone penetration does. The only group which is very strongly underrepresented among mobile phone users are older-age people (60 years and older). All this makes the sampling of mobile phones a very attractive approach for selecting and reaching a respondent, but unfortunately not necessarily for stimulating respondent to cooperate or to conduct full interview.

1.4 Sampling for Telephone Surveys in Russia

Not only coverage is an important issue in telephone surveys in Russia. Access to a sample frame from which the same could be drawn is another major point. Although the management of telecommunication systems in the Soviet Union was centralized, the actual telephone lines and registers of home telephones were maintained on the city or regional level. A single directory of private telephone numbers never existed in Russia and this is true today. The directories of private telephones were held at the city level, but they were rarely published and could not be legally obtained. Only some cities have open telephone directories of private households, but even they are usually out of date and usually belong to a single major provider. Recently, a new privacy law forbids publishing such data as the private address or telephone number in any directory. The Russian sector of landline phones consisted of 89 regional state operators and one national operator of inter-city and international connection (Rostelecom) till recently. During the restructuring of this section in 2000 - 2002 most regional state operators were united into seven large companies in different regions. Some share of the private telephone lines belongs to independent smaller operators (for example, 20 % of telephones in Moscow city).

Because of such a disintegrated system of telephone providers and inaccessible information about private telephone numbers, the most common option for constructing the sample of a telephone survey is a three-stage sampling design with a primary selection of geographical units (cities) based on population figures. Subsequently, independent samples are constructed in each selected city by generating random numbers taking into account the theoretical capacity of each telephone station.

As is the case in many other countries, a much debated and difficult issue in the sample construction in Russia is the selection of a respondent within a household. A random selection process based on the household composition (for example the Kish procedure) leads to a substantial increase of the refusal rate. Questions about the household composition are regarded as very sensitive and even threatening in the Russian context. It is primarily due to safety concerns and criminality threats resulted in under-reporting of single households and also to legal concerns - requirement to register all residents in large cities, which is very often violated, and uncertainty about the law on letting private housing which lead to under-reporting of some household members

Another issue is biases in the interviewed sample due to underrepresentation of some demographic groups in the randomly selected samples. We observe a consistent underrepresentation of certain socio-demographic groups in telephone surveys: men, middle-aged people of 35 - 45 years old, sometimes young people and people with low education. The reason for this is either a low accessibility of these groups by phone or a higher refusal rate. Direct refusals are an important reason for under-representing people with low education. Lower accessibility is major reason for gender and age biases – because of differences in the at home patterns and therefore lower probability of being reached by landline phone for people aged 35 - 45 years. In general, men are underrepresented both because of less time spent at home and also because of a cultural effect due to which the telephone is usually (in two thirds of the cases) picked up by women and women often serve as a gatekeeper in the household (Rogozin 2006). Mostly, the problem of a complicated selection procedure within a household and underrepresentation of particular groups is solved by introducing a quota sample. And use of quota samples has been increasing dramatically. Unfortunately, little methodological work has been done in the country to offer a better alternative to quota samples based on random selection.

The estimation of and trends in response rates in telephone surveys in Russia is difficult. Information about the response rates in telephone surveys is very rarely published. We found only few mentions of it in open sources among hundreds of surveys we went through. The calculation of response rates differs very much. The main differences are in the inclusion or exclusion of non-contacts which could be due to the non-existence of telephone numbers in the RDD system or a failure to reach anybody at home. There is no ways to distinguish the two and usually they can account for 50 - 80 % of all selected telephones. Many surveys use a quota approach to select the respondent and make very few call-attempts, such that estimating response rates does not make much sense. Usually, researchers report about 40 - 60 % direct refusals and interrupted interviews which is higher than direct refusals in face-to-face interviews in Russia nationally but similar to refusal rates in large cities (Yudin 2008). In the recent years there has been a trend to higher percentages of interrupted interviews if average length of the interview exceeds twenty minutes.

The fast growth of telephone surveys in Russia regardless of all sampling issues and complications can be explained by the large benefits in the speed of the survey compared to the long period of interviewing face-to-face in very large geographical territories and other organizational advantages. There are also generally positive attitudes toward telephone conversations with strangers on different important issues, non-technical factors (telephone penetration and sampling) related to effectiveness of telephone interviews in general and positive feedback from respondents. Nevertheless, telephone surveys are usually still considered and most often used for the interviews on "easy", non-sensitive, non-threatening topics which do not require establishing a minimal level of interpersonal trust. Telephone surveys offer a quick way of getting information when the demands on the respondents' memory, attention and answering precision are minimal. In Russia, examples of such topics are the consumption of typical consumer goods, media preferences, but also political questions including evaluations of the government's performance and different actions and policies, electoral choice, etc. (In contrast, in Belarus or in Middle Asia telephone conversations on these topics would be considered threatening and undesirable).

In the recent years we have found more and more attempts at using telephone interviews for studying a broader range of topics including such sensitive issues as AIDS and other health issues. At least the researchers report that these attempts were rather successful[7]. Some methodological experiments with telephone interviews in Russia showed higher social desirability effects on some questions (for example, participation in elections), a lower level of direct refusals to particular questions, bias toward the positive side of the scales (for example, in evaluation of political leaders or government performance) and an avoidance of extreme answers on the negative side of the scale. But the amount of methodological work in this field is very limited in Russia, since practical experience in conducting such surveys has not yet been accumulated or elaborated on in methodological research and analysis in the last twenty years.

1.5 Recent Trends

The introduction of mobile phone connections and the extremely quick growth in popularity and penetration of this means of communication do not leave much hope that penetration of landline phones will ever reach an adequate level for national surveys. It is more likely that Russia will skip this stage and will come directly to developing other "post-telephone" methods of data collection.

We are aware about some attempts at conducting national surveys in Russia by mobile phones, but there are serious limitations to this. One of them is the limitation on the length of interview. Higher accessibility of respondents of different social groups and skipping one sampling step of the respondent selection within a household is compensated for by higher direct refusals. Reasons for refusals are manifold - common for all survey modes (distrust to the surveys, security reasons, lack of interest) and specific to mobile phone method. The latter includes contact call in inappropriate place or time (in public transport, driving, in the midst of a private conversation, etc), lack of time to persuade respondent to participate, privacy concerns, financial reasons (costs of roaming applied in some regions for in-coming calls), interrupted interviews.

It is not likely that mobile telephone surveys will dominate national surveys in Russia in the near future. Some mix-mode experiments are conducted mostly by commercial research companies to find the solution for difficulties related to different interviewing methods including interviewing face-to-face in rural areas and small cities and conducting telephone interviews in large cities; using landline telephone calls for screening purposes, the allocation of respondents or making appointments for face-to-face interview; using mobile phones for screening purposes and obtaining contact information for web self-administered questionnaires. All these or new ideas could be part of our future, but at least for some time face-to-face interview will stay the main data collection method for large population surveys in Russia.

---

[7] AIDS knowledge, attitudes, and behavior in Russia: results of a population-based, random-digit telephone survey in St Petersburg. Amirkhanian YA, Kelly JA, Issayev DD. St Petersburg State University, Russia. yuri@mcw.edu.

References

Andreenkov, V.G., Sotnikova, T.N.: Telephone surveys of population. Methodological recommendations on conducting sampling mass surveys. Institute of Sociology, Academy of Sciences of USSR, Moscow (1985)
Myagkov, A.Y.: Influence of the method of data collection on the verbal behavior of respondents. Sociological Journal 1(2) (1999)
Rogozin, D.M.: Effectiveness of telephone survey depending on the restrictions on the selection of respondent within a household. Sociological Journal 3 (2005)
Rogozin, D.M.: Impact of respondent's gender on the effectiveness of telephone interview. Sociology: Methodology, Methods, Mathematical Modeling 22 (2006)
Yudin, G.B.: Area localization and the level of non-response in mass survey. Sociological Journal 1, 49–72 (2008)

Appendix

Table 1.2a Coverage of socio-economic and demographic groups by different means of communications in Russia

	Stationary telephone		*Mobile phone*		*Access to Internet**		*% of group*
	available	Not available	available	Not available	Yes	No	in total population
Age							
15-24	51	49	88	12	72	28	19
25-29	40	60	88	12	65	35	8
30-44	53	47	83	17	50	50	24
45-59	55	45	75	25	29	71	26
60 +	51	49	36	64	5	95	22
Education							
Low	41	59	40	60	7	93	16
Medium	49	51	72	28	25	75	55
High	62	38	89	11	69	31	29
Urbanization of the place of residence							
Cities with population of 1 mln people +	82	18	86	14	58	42	17
Cities of 500 th - 1 mln people	72	28	79	21	67	33	8
Cities of 100 - 500 th people	57	43	76	24	40	60	20
Cities of 20-100 th people	52	48	74	26	32	68	14
Urban areas with less than 20 th residents and rural areas	32	68	62	38	24	76	41

Table 1.2b

	Stationary telephone		*Mobile phone*		*Access to Internet**		*% of group*
	available	Not available	available	Not available	Yes	No	In total population
Type of housing							
Detached house for single family	29	71	54	46	21	79	27
Multi-flat building less than 3 floors (old type)	34	66	64	36	31	69	15
Multi-flat building more than 3 floors (modern type)	66	34	76	24	48	52	56
Communal apartment in multi-flat house (several facilities in one flat)	36	64	68	32	NA	NA	1
Dormitory	13	88	72	28	NA	NA	1
Area of residence							
Houses in the area in very or good shape	57	43	73	27	NA	NA	44
Satisfactory conditions	48	52	67	33	NA	NA	49
Bad or very bad conditions	22	78	51	49	NA	NA	7
Number of people in the household							
One	38	62	47	53	NA	NA	12
Two	50	50	66	34	NA	NA	28
3 and more	55	45	79	21	NA	NA	60
Household composition							
Single man - 40 years old	23	77	88	12	NA	NA	1
Single woman - 40 years old	33	67	90	10	NA	NA	1
Single man 40 +	29	71	44	56	NA	NA	2
Single woman 40 +	43	57	37	63	NA	NA	8
Two adult people, oldest person is younger than 35 years old (young families)	35	65	87	13	NA	NA	9
Two adult people where oldest is 35 - 50 years old	52	48	87	13	NA	NA	14
Two adult people where oldest is 50 + years old	53	47	56	44	NA	NA	19
HH with 3 + adults	58	42	77	23	NA	NA	44
Other types	54	46	82	18	NA	NA	2

* from home

Source: ESS data wave 4 (2008) on stationary phones and mobiles phones. Information about access to Internet from homes – is from “Monitoring of Social-Economic Situation in Russia” – National Survey of population of 18 years old and over by face-to-face interviews on national sample, conducted in fall 2010 with 2000 respondent.

Chapter 2
The Netherlands

Dirkjan Beukenhorst*

2.1 Introduction

The Netherlands are a very small country where 16 million people live closely together. This situation is important for understanding the history and use of telephone surveys for research or marketing. In countries like, for example, Finland conducting survey research using only face-to-face interviews is out of the question: many regions are so thinly populated that traveling costs and traveling time for face-to-face interviewers would be enormous. In the Netherlands, on the contrary, face-to-face interviews were affordable as a data collection method for scientific or statistical research until the recent past. During the last two decades, however, the costs of face-to-face interviews have been rising sharply, among other things due to rising transportation costs and to a change in the laws prohibiting hiring interviewers on piece-wages. Paying interviewers by the hour implied that traveling time had to be reimbursed. As a consequence of these increasing expenses for face-to-face data collection telephone interviews became more popular. This increase in CATI surveys was so sharp that by 1996 the CATI share in marketing research had risen to two times the European average. In fact, the Netherlands came immediately after the thinly populated Scandinavian countries in CATI use for market research (Bronner 2000). Data collection by telephone, however, has some drawbacks, as will be explained in this spotlight. Especially sampling frame problems are serious in the Netherlands.

In this chapter we will describe the (best) practices as employed by high quality market research organizations and survey organizations like Statistics Netherlands. Practices of low standard organizations and marketeers who want to sell, not to collect representative data, are not covered here.

We will show that at the moment in the Netherlands mobile phone surveys do not seem to be suitable for traditional surveys with rather long interviews. In the first place this is because a good sampling frame is lacking; in the second place, because contact is often made with respondents at moments that are not very well suited for long interviews. Nonetheless, innovative, quick data collection will

---

* The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

become feasible when 3G or 4G technology becomes widely spread among mobile phone users.

Subsequently, I will contend that in the years to come traditional landline surveys will have a smaller share in the total amount of research that collects data by interviewing persons and households. This is caused by shrinking coverage and especially an ever diminishing quality of the sampling frame. At the same time, web surveys become more popular as a cheaper alternative, often as web access panels. CATI as a data collection mode has to be combined with other modes of data collection, if the survey is intended to deliver data of high quality.

2.2 Mobile and Landline Coverage, Sampling Frames and Response

In 2009 about 20 million mobile phones were used in the Netherlands, that is 120 for every 100 inhabitants (TNO 2009). People have sometimes more than one mobile phone, for example one for private use and one for work related calls. 28 percent of the population had two or more mobile phones. Twenty percent of all mobile phones were smart phones using 3G technology. This is changing very quickly as the modern devices are becoming ever more popular.

The number of landline connections is quite stable at around 5,8 million (household and enterprise connections). A landline phone is a typical household asset, compared to mobile phones, which are predominantly personal possessions. According to the Eurobarometer survey 91% of Dutch households possessed a landline phone in 2008, which is the highest percentage in Europe. Only 9% are mobile-only households (having no landline telephone), which is very low, compared to, for example, the Czech Republic with 64% or Finland with 61%. Around 85% of Dutch households had a landline and one or more mobile phones, which is also the highest percentage in Europe.

In theory market or survey research could nowadays use mobile phones as a communication channel for reaching almost anyone for an interview. However, in the Netherlands nearly no one has his mobile number listed in a telephone directory. People consider their mobile phone as a private communication channel and give their number only to selected family members, friends or business acquaintances. Thus it is impossible to draw a random sample of persons and approach them for an interview by mobile phone because it is impossible to find telephone numbers for the sampled persons. Another possibility would be to start not with a sample of persons but with a sample of cell phone numbers. To achieve such a sample one needs a frame of mobile numbers in use, or create banks of (possible) mobile numbers. This approach is difficult because providers do not release the numbers in use or are very reluctant to do so. One could call a Random Digit Dialing (RDD) sample for conducting interviews. Because people have more than one mobile phone, the inclusion probability would be unknown. Asking them during the interview how many mobile phones they possess would only give some clues about the respondents, not about how many phones the nonrespondents possess.

Although coverage is more than complete and people could easily be contacted – carrying their phone everywhere in public places and even in the bathroom (as research has shown) – this medium is not the first choice for surveying a sample of the population. Just because contact can be easily established, people often are called at moments that are not well-suited for giving an interview. People who pick up the phone will probably often not be able or willing to answer a possibly long questionnaire. A relative high refusal rate among very active people would bias the results. Correcting for this bias seems impossible. If respondents ‘on the move’ do answer the questions, a lot of satisficing is to be expected. This means that they will not spend much energy on comprehending the question and finding and formulating the right answer (Krosnick 1991). This would also lead to biased outcomes. Therefore, in comparison to interviews where the respondent is at home, some fears about data quality seem justified. Of course, if people are at home when answering their mobile phone, there will be little or no difference in data quality between data collected through landline or mobile calls. Important are only the circumstances of the interview, not the technical aspects of the call.

So it seems as if the advantage of calling cell phones – reaching respondents always and everywhere – may be offset by disadvantages concerning biased results and data quality. This holds especially for long and difficult questionnaires. For mobile-only persons or households there is no alternative for interviewing them on their mobile phones, but for people owning both a landline and a mobile phone it seems preferable to do as many call attempts as necessary to reach even the very mobile population at home.

The situation may change in the near future. More and more people in the Netherlands cancel their landline phone. For them their cell phone is not a private communication medium anymore but their only telephone. Therefore they will be more inclined to give their number to other people and businesses, or even let it be registered in a public directory. This would in part solve the sampling frame problems. Short interviews would deliver data of acceptable quality. One can imagine very useful applications, for example, a panel of mobile phone users who give short interviews on selected moments on selected days for time use surveys or mobility research.

Coverage of landlines in the Netherlands is – as we have seen – still very high with more than 90% of households having such a device. However, problems with sampling are growing. We will illustrate these with the example of the practices of Statistics Netherlands. As a national statistical institute Statistics Netherlands aims at the highest standards in representativeness and data quality. Furthermore, Statistics Netherlands has the possibilities to do a lot of research on coverage and sampling frames because a lot of data about unlisted and listed persons are available in the Social Statistical Database (SSB) which contains much information about the Dutch population. This SSB mainly consists of administrative data on persons, households, jobs, benefits and pensions. It covers the entire Dutch population. Sources are among others insurance data, tax data, data on education and data gathered from surveys. In the SSB several characteristics of persons are available, such as gender, age, ethnic group, region, position in the household and type of household, (household)income, position in the labor market and earned wages.

When preparing a CATI survey Statistics Netherlands draws a representative sample of persons or addresses using the national population register as sampling frame. Then, telephone numbers for these persons or addresses are looked up in the telephone directory by private parties; a first search is automatic and a second one is done manually. This method of drawing a sample is done, among other reasons, because it is Statistics Netherland's policy to send all sampled elements a letter inviting them to participate in an interview and announcing the call by an interviewer. RDD surveys are therefore out of the question.

The success rate of this search is diminishing since many years. At the moment 65 to 70% of sampled elements can be matched with a telephone number. The fact that one in three persons cannot be approached by telephone is evidently detrimental to data quality. The risk of biased outcomes is great. This situation has been existing already for many years.

Comparing the population distributions with the distributions of those people for whom a number is traced shows that the prevalence of unlisted telephone numbers is generally above average for non-western immigrants, young people and inhabitants of urban areas. Even the differences between metropolitan agglomerations are large: in Amsterdam 40% could be matched in 1998, in Rotterdam 53% and in The Hague 60%. The sample can be weighted for this uneven distribution because population margins are known. For other characteristics this is not possible. Other research (Engbersen et al. 1990) found for example that also jobless people and persons on social security benefits frequently have an unlisted number. More recent research conducted at the neighbourhood level of a medium-sized city in 2003 found relatively more unlisted numbers in underprivileged and socially disorganized neighbourhoods (Jansma et al. 1990).

Apart from these biases caused by the different success rates of finding a telephone numbers for various subpopulations response rates differ markedly between subpopulations. In a perfect world these response rate differences would counterbalance the biases from the incomplete frame. This is, alas, not the case. Those groups that are underrepresented in the sample frame are also less likely to respond to a survey request. This is not only the case in the CATI mode but also in other modes (Beukenhorst and Wetzels 2009). On average the response rates at Statistics Netherlands fluctuate around 65%; but among non-western migrants, young people and in metropolitan areas response rates are much lower. These response rates are calculated by dividing the number of interviews by the number of persons (with a known telephone number) called. Frame errors are included in the denominator. These are negligible because the frame is built by combining the population register with telephone directories. If the response rate is calculated over the originally drawn sample from the population register, the average response rate is only 70% traced numbers *65% response rate = 42% and for some subpopulations this can drop to around 20%.

Another way of looking at the effects of the large portions of unlisted numbers is studying the differences between respondents with a listed number and respondents with an unlisted number, instead of the differences between the population segments with or without a known telephone number. In this kind of analyses one finds differences between these groups that cannot be inferred from the Social Statistical Data Base or the population register. In 1987 respondents of Statistics Netherlands' Health Interview Survey (HIS), a face-to-face interview with around 5.000 respondents, were asked if they had a telephone and if the number was listed in the directory. 95% had a telephone, 88% had a listed number and only seven percent an unlisted number. Possession of a telephone was, however, not evenly distributed, for example ten percent of young people and the same percentage of divorced people did not have a telephone. They were also more likely to have an unlisted number. The combined effect was that only 82% of young people and only 6 % of divorced people could have been approached by telephone (van den Berg 1987). In this research differences in target variables were not studied.

In 1998 Statistics Netherlands conducted a CATI experiment with a sample of respondents with a known number and respondents in an RDD sample in the city of Rotterdam. A sample of numbers was generated with prefixes that were overrepresented in neighbourhoods where relatively many persons have an unlisted number. The percentage of listed numbers in these neighbourhoods varied between 44 and 82%. The sample of random numbers was generated in such a way that unlisted numbers would be oversampled. All these numbers were called and eligible persons were asked to participate in the Housing Survey. Comparisons of four important target variables showed large differences between respondents with a listed number and those with an unlisted number. Both groups consisted of 519 respondents. 41% of respondents with unlisted phone number planned on moving within a year in comparison to 33% of respondents with listed numbers, 39% were migrants (an important target group) compared to 18% in the sample of listed numbers. The number of people who owned their living quarters was overestimated and the number of people receiving rent allowances was underestimated in the sample of listed phone numbers (Vousten 1999). In another large city, The Hague, the CATI responses to the same Housing Survey of listed people has been compared to the CAPI responses of unlisted people. Differences between the two groups were significant, probably due to the combination of mode- and selection effects (Beukenhorst 1999).

A very recent research concerned the respondents of the Labor Force Survey (LFS). For this face-to-face survey of addresses 67% of the respondents could in theory have been approached by telephone. Differences were analyzed for the group of respondents with a known telephone number and the group who could only be reached face-to-face. For all background variables and all target variables of the LFS significant differences were found between respondents, who potentially could have been interviewed using CATI, and the total sample. For example, of the listed people thirty percent lived in cheap houses compared to 38% of the total sample of 19.000 cases used in the analysis (Banning et al. 2009a). The same results were found in a similar analysis of the Health Interview Survey, using 5.700 cases in the analysis (Banning et al. 2009b)

The conclusions of all these studies using different analysis methods are that the thirty percent of sampled persons who cannot be approached by telephone differ in a significant way from the persons who can be approached in a CATI survey. These differences sometimes lead to strongly biased results.

As a consequence of this strong bias in the telephone-only data collection mode Statistics Netherlands and other research organizations are abandoning this type of survey and substituting it with mixed-mode designs that use CATI for people who can be approached by telephone and other modes like face-to-face or mail surveys for people without a known telephone number. Such strategies are often preceded by a request to participate in a web survey. These combinations of different data collection modes, on the one hand, aim to reduce coverage and response problems and, on the other, aim to design affordable surveys (Cuppen et al. 2010).

2.3 Survey Climate and the Future of CATI Surveys

It is tricky to write about 'the survey climate' because it is a rather vague notion. It is difficult to study, too. The only way to study 'survey climate' is by conducting a survey. The drawback is that we only survey the opinions of respondents on this topic. Calling back nonrespondents with a short questionnaire usually is not very illuminating: the 'not interested' and 'no time' answer categories are used most often.

What we do know is that response rates have been falling over the years. This trend can only be countered by putting ever more energy and money into getting responses. This holds true for CAPI as well as CATI surveys. It is difficult to judge to what extent 'survey climate' contributes to this decline. The concept of the 'survey climate' implies an underlying willingness or unwillingness to participate in a survey and thus has to do with the refusal rate. Other changes in society in general may also contribute to a decline in response rates. People are at home less frequently and consequently more difficult to get into contact with, by phone or in person.

Sometimes it is suggested that the survey climate is changing because people are getting tired of being called 'all the time' for surveys or telemarketing. In 1999 van Rooy and Steenis (1999) conducted a survey about CATI surveys and telemarketing. Their conclusion was that only a small percentage of all calls to private households concern surveys or telemarketing. Since then telemarketing has been increasing very much and generating negative reactions. Since 2009 people can have their telephone number (mobile or landline) registered in the Belmeniet register (Don't call me register). Telemarketeers are forbidden by law to call numbers on this register. Serious survey and market research organizations are, however, allowed to call numbers in this register. Already more than a million numbers are registered. This suggests that, at least for telemarketing, the climate is worsening.

One may wonder if this development will create a more positive attitude towards serious survey research, because pseudo-research aiming at selling products or services will decrease. It is also possible that people perceive telemarketing and market or statistical research as the same thing and get all the more annoyed, if called by a survey organization.

Rising costs, diminishing coverage, fewer listed numbers, falling response rates and possibly a more negative CATI survey climate call for other modes of data collection; for example in mixed mode designs. They solve many of the quality problems mentioned but are expensive compared to pure CATI research, albeit less expensive than pure face-to-face surveys.

Concurrently, the alternative of web surveys is getting more popular and offering more technical possibilities. As yet, low response rates and non-representative panels raise doubts about data quality. It is, however, doubtful if telephone surveys in future will survive the competition of web surveys answered on a personal computer or laptop.

The present trend is that mobile phones are becoming more similar to personal computers, so maybe in the future there will be no difference between a phone and a pc or notebook, thus making the above question an academic one. Already quite a few respondents answer web questionnaires on their cell phone.

References

Banning, R., Wetzels, W., Schouten, B.: Dekking van geregistreerde telefoonbezit en web in de EBB 2009. In: Coverage f Listed Numbers and web Access in the Labour Force Survey. Internal CBS-report, Den Haag (2009a)

Banning, R., Wetzels, W., Schouten, B.: Dekking van geregistreerde telefoonbezit en web in de Gezondheidsenquête 2008. In: Coverage f Listed Numbers and Web Access in the Labour Force Survey. Internal CBS-report, Den Haag (2009b)

Beukenhorst, D.J.: Vertekening door non-respons in het WBO. In: Non-response Bias in the WBO. CBS en Ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer, Den Haag/Voorburg (1999)

Beukenhorst, D.J., Wetzels, W.: A Comparison of Two Mixed Mode Designs of the Dutch Safety Monitor: mode effects, costs, logistics. In: Paper Presented at the 3rd European Survey Research Association Conference 2009, Warsaw (2009)

Bronner, A.E.: De praktijk van het marktonderzoek (market research in practice). Van 1975 via 1984 naar 2000. MOA jaarboek (2000)

Cuppen, M., van der Laan, P., van Nunspeet, W.: Challenges of redesigning household surveys and maintaining output quality. In: Paper Presented at the European Conference on Quality in Official Statistics (Q 2010), Helsinki, Finland, May 4-6 (2010)

Engbersen, G., Jaap, T., van der Sluis, J.: Unlisted Numbers: the making of a Dutch Underclass. Working paper no 3, University of Leiden. Department of Sociology (1990)

Jansma, F., Goor, H., van, V.R.: Verschillen staswijken in onderdekking en nonrespons. In: Do neighbourhoods differ in undercoverage and non-response?. MOA-jaarboek (2003)

Krosnick, J.A.: Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology 5, 213–236 (1991)

van der Laan, P., Nunspeet, W.V.: Modernizing Household Surveys in the Netherlands: Design, Efficiency Gains and Perspectives. Discussion Paper 09044, Statistics Netherlands, The Hague/Heerlen (2009)

Van den Berg, J.: Telefonische bereikbaarheid respondenten face to face enquêtes, interne notitie. In: Telephone Accessibility Respondents Face to Face Surveys, International Control Notes. Centraal Bureau voor de Statistiek, Voorburg (1987)

Van Rooy, C., van Steenis, J.C.: Bellen & gebeld worden: fabels en feiten. In: Calling and Being Called: Fables and Facts, MOAjaar-boek (1999)

Vousten, R.: Random Digit Dialing in Rotterdam. Internal CBS-report (1999) (in Dutch)

Chapter 3
Switzerland

Michèle Ernst Stähli

3.1 Introduction

Compared to other European countries, Switzerland has historically had and still has excellent landline telephone provisions, in spite of a large amount of peripheral zones that are hardly accessible. As long as the national telecom operator had a monopoly on the market, the survey agencies had access to a unique telephone directory with good coverage. Even though there were no reliable figures about this directory, it was generally admitted that it contained almost one hundred percent of Swiss households. Therefore, starting in the 1980s, telephone surveys were seen as the most reliable and cost-effective method for market and scientific research. This was the beginning of a real telephone survey culture.

However, for approximately the past ten years, landlines have declined in favor of mobile connections, an evolution that has also been observed on an international level. Moreover, for more than a decade, Swiss citizens are no longer obligated to list their phone numbers in the public directory. To compile their samples, the survey agencies therefore relied massively upon private address management companies. Public administrations also tried to obtain exclusive contracts with operators for non-listed numbers and academic and scientific-driven institutes have turned more and more to face-to-face interviews in order to guarantee better coverage. Very recently, the Federal Statistical Office also questioned its own code of practice as the directory it works with contains less and less landline numbers. Strategies about new paths, such as mixing modes, have begun to be undertaken, at least for official statistics and high-quality social research.

3.2 History

Telephone surveys have a strong tradition in Switzerland. Since the early 1980s, survey organizations have invested in CATI systems. The two survey agencies in Switzerland that started using CATI surveys as early as 1982, pioneers even on a

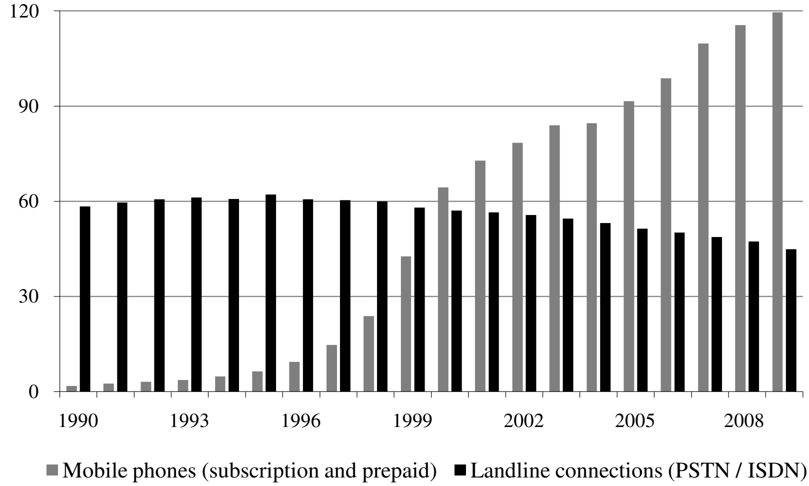
European level are, today, the major institutes for science-driven surveys. Telephone interviews developed quickly and were established as the best mode for high-quality, cost-effective surveys. As of 1989, 47% of all interviews of members of the vsms (Swiss association of marketing and social research) were conducted by telephone (SMS 2002). The survey institutes developed their know-how on questionnaires, design, interviewer briefings, fieldwork management, etc., specifically in the field of CATI surveys. The predominance of telephone surveys is such that, for the past two decades, almost all personal and household surveys done by the Federal Statistical Office (FSO) are carried out by telephone if possible. The FSO does not have its own telephone laboratory, so it relies upon external survey agencies to do the fieldwork.

This situation was satisfying for all involved up to the turn of century. In fact, Switzerland had one of the best landline provisions in the world (KIG 2001), and until 1998 it was compulsory to register one's landline number in the public directory, except for special and rare instances for which approval had to be given. Switzerland, therefore, presented very propitious conditions for high-quality telephone surveys.

The abolishment of the obligation to register all telephone numbers and the development of mobile phones raised, however, the question of coverage. Do available landline telephone numbers still represent a satisfying selection frame for population and household surveys? A survey agency established that, in 2000, over 12% of Swiss households had no registered landline telephone number (Schmugge and Grau 2000)[1]. Meanwhile, the situation has deteriorated noticeably regarding the usability of landline phone numbers for high-quality surveys.

Let us first look at the details of the mobile phone expansion, probably the principal cause of the deterioration of landline coverage. The use of mobile phones has increased rapidly in the last several years. In effect, between 1990 and 2009, the number of mobile phone connections (prepaid and subscription) has increased consistently: from 2 mobile phones per 100 inhabitants to 120 mobile phones per 100 inhabitants (see Figure 3.1). As of 2007, each resident of Switzerland had more than one mobile phone connection on average, and the trend is still growing. At the same time, the number of landline phone connections tended to diminish (from 58 per 100 to 45 per 100 inhabitants between 1990 and 2009). As of 2000, the number of Swiss mobile connections exceeded landline connections.

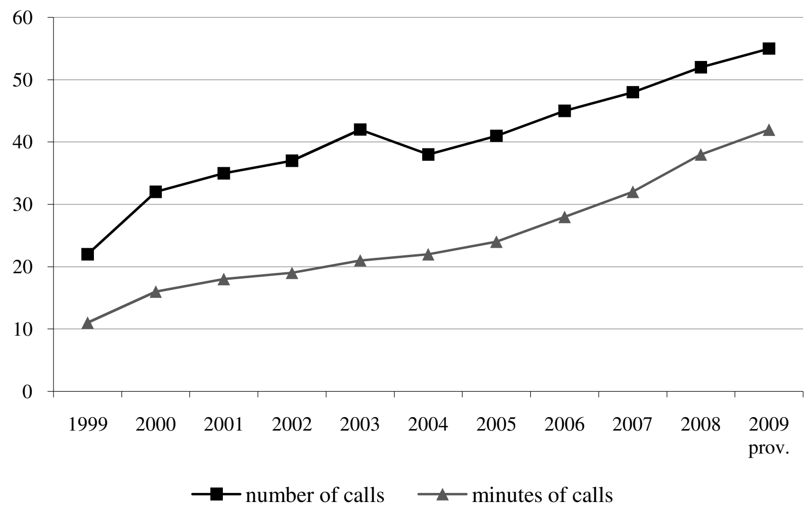
[1] This figure might have been underestimated, as the study was carried out by telephone (landline and mobile) and was corrected for by an estimation of households without any telephone connection, an estimation the authors admit might be too low.



Source: OFCOM/BAKOM © 2011 OFS-BFS-UST/SUKO

Fig. 3.1 Evolution of landline and mobile infrastructure in Switzerland, 1990–2009 (number per 100 inhabitants)

The use of this infrastructure evolved in the same manner. Not only has the number of mobile connections increased rapidly, but the number of calls from mobile phones has recently exceeded those operated from landline phones, even though their length is clearly shorter than calls from landline phones (see Figure 3.2).



Source: Swiss Official Statistics on Telecommunication 2009, OFCOM; Graphical representation by the author

Fig. 3.2 Rate of mobile over total telephone communications (landline + mobile)

Mobile phone calls are not only more expensive than landline calls, except for certain calling plans, but are also less comfortable for both conversational partners. The connection is less stable, especially when travelling, the quality of sound is not always optimal and privacy is not available most of the time. All these dimensions explain why the majority of call minutes are still conducted on landline phones, even if the number of mobile connections is higher than the number of landline connections. These considerations have to be kept in mind when discussing the convenience of surveys by mobile telephone.

What do these provision and call figures mean for the distribution and use of mobile phones in the population?

According the survey “Mobilfunkmonitor” (Bieri et al. 2008), 80% of the resident population (aged 16 years and older) used a mobile phone in 2004 and 87% in 2008[2]. The most recent data have been published by Public Data 2010 based on the “KommTech” survey. They indicate that 91% of the resident population in Switzerland (aged 15 years and older) used a mobile phone at the end of 2009 (5% a smart phone), and 47% used it every day. Even though the data are not strictly comparable because of methodological differences[3], the trend of rapidly growing mobile phone use is confirmed.

None of these figures inform us, however, about the actual rate of households and individuals still reachable by landline phone. To begin with, we have to know how many households do have a landline telephone.

A first indication is given by a survey carried out for the Federal Office of Communication indicating that, in 2007, 92% of Swiss households still had a landline phone (see Table 3.1). The report gives also some interesting insights (M.I.S. Trend/OFCOM 2007): households with a mobile phone only are mainly one-person and low-income households. Households with no phone at all are mainly households with very low income. Among the households that are not listed in the directory, the rate of mobile only is 26% and the rate of no telephone at all is 10%. Classic telephone surveys therefore necessarily imply bias related to income and household size. Unfortunately, the report does not give any information about the rate of households for which the agency found a phone number listed in the directory.

Even though the telephone provision of Swiss households was and probably still is among the best in Europe, the situation is not perfect for telephone surveys. In fact, the problem has two root causes: the growth of mobile connections on the one hand, as we showed previously, and telephone numbers that are not listed in the directory, on the other. In addition, the two factors interact: mobile numbers are listed much less often than are landline numbers. Hence, all the data about phone provisions do not tell us how many households can actually be reached for a telephone survey.

---

2 The results give some insight into mobile phone use by socio-demographic characteristics. In 2008, the rate of people using mobile phones was only 45 % among senior citizens (65 years and older); However, this rate is increasing as people that are accustomed with mobile phones are growing older. Working people, the higher educated and high-income people, as well as those aged less than 40 years have the highest rate of mobile phone use.

3 An OFCOM study, quoted previously, that was conducted in 2007 showed 90 % of the population aged 15 years and older using mobile phones (M.I.S. Trend/OFCOM 2007).

There are no official figures about the rate of connected phone numbers that are listed in the directory. The European Social Study and the European Value Study, which were carried out in Switzerland in 2008, asked this issue directly.

Table 3.1 Coverage of households in Switzerland in percent, 2008

	OFCOM 2007	ESS 2008	EVS 2008
With landline telephone	*93*	*90*	*87*
- and at least one listed number	-	-	80
- but no listed number	-	-	7
Mobile only	*5*	*9*	*11*
- with at least one listed number	-	-	1
- but no listed number	-	-	10
No landline and no mobile phone	*2*	*1*	2
Total number of interviews	1,002	1,819	1,272

Sources: M.I.S. Trend/OFCOM (2007) and European Social Survey and European Value Study 2008,[4] design-weighted data

From this table we can see that about 90% of the private households that have a landline connection have their number listed in the public directory, whereas only 10 % of the only mobile phone households do so (see Table 3.1). If we look at individual mobile phones, only about 6 – 7% of their owners say that their number is listed in the directory (ESS 2008). This is not surprising, knowing that 43% of mobile phone clients have a prepaid phone and no subscription (OFCOM 2008).

This table also confirms the assumption that, in 2008, more than 20% of households could not be sampled out of the public directory. A survey agency comes to a similar conclusion based on other evidence[5]. The problem of insufficient coverage arises at least as much from 'mobile only' households as from missing listings of landline phone numbers. As mobile connections continue to grow at the expense of landline connections and mobile numbers are rarely listed in directories, the problem of coverage will increase. In only one year, between 2007 and 2008, households with mobile phone only doubled. Very recently, the loss of landline connections seems to grow by several percentages per year. If we take this informal information seriously, by 2011 Switzerland has to face landline number coverage of as low as 70%, with a trend towards increasing rates of households without a known landline number.

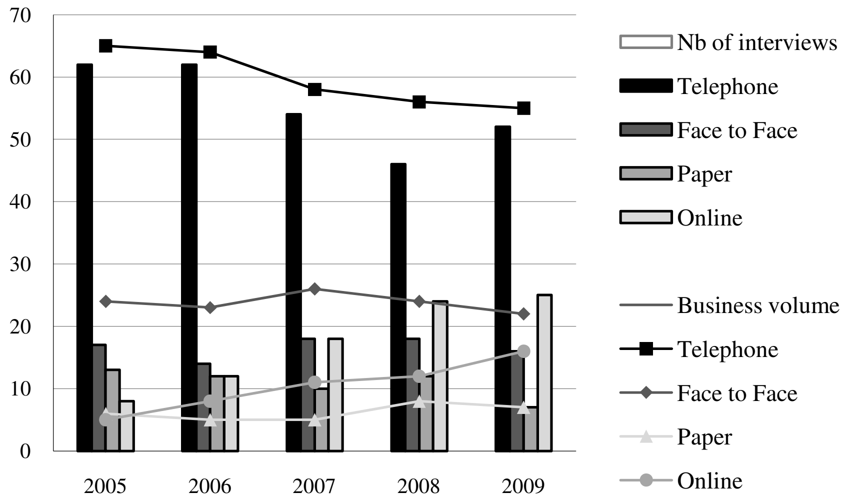
[4] The ESS and EVS were realized in Switzerland as CAPI face-to-face surveys, with a sample based on the complete register of buildings and mailboxes in the country. The OFCOM study had a similar sampling, but when landline phone number was known, the interview occurred by telephone.

[5] Based on face-to-face surveys with innovative sampling, the survey concludes in 2009 that about 20 % of the Swiss households do not have a publicly accessible telephone number. This is also the rate of households without a telephone number contained in the database the agency works with (M.I.S. Trend/BAKOM 2009).

Despite this evolution, telephone surveys still represent the majority of the interviews conducted in market and social research.

The annual statistics of the association of market and social research show that even though the amount of interviews conducted by telephone has slightly decreased in the last years, telephone surveys continue to be dominant (see Figure 3.3). In fact, online interviews begin to expand and already represent a quarter of the interviews. However, the know-how for high-quality surveys with this new mode is still developing, whereas the background for telephone surveys is well consolidated.

Unfortunately, this association does not provide differentiated figures for landline and mobile telephone interviews. As we know, some agencies use mobile phone numbers in the same way as landline phone numbers; others only use mobile phone numbers if they have been communicated by the respondent himself or if they feature in databases supplied by a client.



Source: Data from the vsms Swiss Interview Institute®; Graphical representation by the author

Fig. 3.3 Part of interviews and business volume by mode in Swiss survey agencies[6]

The fact that the Federal Statistical Office carries out most of its surveys by telephone contributes to strengthening the legitimacy of this mode in Switzerland (see Figure 3.3). We will see, however, that not only does deficient coverage constitute a challenge but also over-surveying of the population. Too many telephone surveys and marketing strategies conducted by telephone worsen the climate for this mode (see section 3.3).

[6] The main survey agencies in Switzerland all are member of this association. The figures in this table do not include panel interviews, even though they represent over a third of the business volume!

3.3 Selection Frame

Since the end of the Swisscom monopoly, there is no longer a single database that covers the phone numbers of all suppliers, not even with listed numbers. There still is the official Swisscom directory, in the form of books, CD-ROMs, and online versions. There are also other suppliers of directories on CD-ROM and the Internet. Some online directories even offer the possibility to add, modify, and suppress the entries, so that the databases may vary substantially. While some call centers continue to use these directories, especially for marketing purposes, the main survey agencies work with databases from address management companies. They offer continuously updated databases that gather information from a multiplicity of sources, like different directories and commercial customer databases. Following their own promotion, the best of them have information on approximately 3.2 million households and 5.8 million individuals, but they do not have a phone number for all of them. The previously cited survey agency, which conducts their CATI interviews only with those who have landline phones, estimates that such databases contain information on landline telephone numbers for almost 80% of the households[7]. Their experience shows, however, that the databases have a lower landline number coverage for households with young members. If we consider that in 2008 Switzerland had 3.4 million households and 7.7 million individuals (official estimates based on the 2000 census), such databases cover at best 95% of the households and ¾ of the population. Moreover, landline phone information is available for, at most, ¾ of Swiss households.

For market research, such databases are regarded as sufficient to produce reliable results. For surveys with administrative and, most notably, scientific aims, they begin to be clearly insufficient.

The strategies of Swiss agencies that want to offer highly representative samples are diverse. One option is to complete the telephone sample by a sample of addresses without known telephone number. Some survey agencies try to build own databases with numbers of ‘mobile-only’ households[8]. Another option is a (RDD) procedure.

Since technology offers this opportunity, some survey organizations invested in RDD. The telephone numbers in Switzerland are structured in a way that makes this possible, though laborious.

The structure of landline and mobile numbers does not differ. The first three numbers, though, indicate if the number belongs to a landline or a mobile phone (see Table 3.2). For landline numbers, the first three digits indicate the region (these digits have to be composed, even on a local level, since 2007). Since 2002, however, the numbers are transferrable across regions, so that the regional code is no longer reliable information about a person’s location, but however few people use the opportunity to transfer their number to another region. For mobile phones, the first three digits originally indicated the operator. This is still true with new

[7] The presence of mobile phone numbers is marginal.

[8] For instance, by asking interviewees if they know people who no longer have a landline connection and getting their numbers. Such a procedure might produce severe bias because it relies on people’s willingness to communicate other people’s numbers.

numbers, but since the numbers are transferable (people can keep their number even when changing operators), operators are no longer necessarily identifiable through the mobile phone number.

Table 3.2 Structure of telephone numbers in Switzerland. Emergency, service and free numbers have another structure (at least one digit less)

Type	Code for CH	Specific code	Block of numbers	Individual number
Landline	0	YZ[9]	abc	xx xx
Mobile	0	7X[10]	abc	xx xx

The numbers are attributed to the operators by blocks. Far from all the blocks are attributed and far from all numbers in the attributed blocks are active. With RDD huge amounts of non-attributed numbers are dialed, which requires enormous resources. Furthermore, numbers for private households are indistinguishable from business numbers and other institutions. The consequence is a high effort in human resources for telephone centers.

The Federal Statistical Office (FSO) undertook another strategy. They could negotiate a special contract giving them the complete list of the numbers, also those that are not listed in the directory. This database can be used for FSO surveys only, but not for academic surveys, which are funded by federal sources. The first survey carried out with this database (the survey on cultural activities in 2008) showed a higher rate of non-contacts than with the former framework, most likely because of a high amount of unlisted phone numbers belonging to secondary residences. The FSO estimates that in 2010 this database covered about 80% to 85% of Swiss private households, although unlisted telephone numbers are included.

Starting in 2011, the Federal Statistical Office will use a newly centralized register of administrative data about the resident population. This database, however, does not contain any phone numbers, so that a link between this register and its telephone number database is being constructed. If this does not give a perfect coverage for 'telephone only' surveys, all necessary information will be available to build methodologies that take into account the whole population. The register data will also be available for some other surveys outside the FSO, but under very restrictive conditions.

[9] The actual regional codes are: 21, 22, 24, 26, 27, 31, 32, 33, 34, 41, 43, 44, 52, 55, 56, 61, 62, 71, 81 and 91. There are also some specific codes (51 and 58) that indicate large public enterprises. Otherwise, commercial and professional connections cannot be distinguished.

[10] The actual mobile codes are 75, 76, 77, 78 and 79.

3.4 Participants' Reactions and Response Behavior

The survey agencies essentially apply quota-based sampling, not only in market research. Only strictly scientific surveys work with random samples and have higher expectations for response rates.

Response rates in telephone surveys differ widely according to the principal, previous information, subject, interview duration, incentives, and contact procedures. A survey that is announced by letter by the national statistical office can achieve up to a 55% response rate for long and complicated topics or even up to 65% for interesting topics, incentives, and short duration. Such response rates, however, can be attained only with intensive contact procedures (no limitation in contact attempts, variation in time of contact, long field duration, intensive refusal conversion procedures, etc.). Without previous announcement, response rates fall drastically. If the principal is not an official entity or not recognized as of public interest, response rates will be far below 50% (more around 25 – 40%), even with short interviews. For example, the Swiss Electoral Studies (Selects), a repeated scientific survey not conducted by the FSO, reveals response rates around 50%[11]. However, the published response rates are rarely strictly comparable, as the definition of ineligible units can differ strongly across studies and organizations in charge[12].

In 2006, an experiment was conducted with the European Social Survey in Switzerland, which is usually only conducted face-to-face (M.I.S. Trend 2007). Overall, 859 households were contacted for a CATI version of ESS. The experiment had a double dimension: testing the length of interview (30, 45 and 60 minutes) and testing the amount of conditional incentive given (20 or 30 CHF, cash, or a donation at choice). Both dimensions turned out to be effective. The gross response rates varied from 34% for a 60-minute/20CHF incentive interview up to 53% for a 30-minute/30CHF incentive interview. The effect of the incentive was stronger on short interviews than on long interviews (11 percentage points response rate increase for 30 minutes compared with a 3 percentage points increase for 60 minutes). In conclusion we can say that CATI surveys are not a suitable mode in Switzerland for interviews taking longer than 30 minutes if response rate is relevant (see also discussion of these data by Eva et al. 2008). Survey institutes often recommend a duration of no longer than 20 minutes and feel uncomfortable with interviews lasting more than 35 minutes.

[11] 1995: 40 %, 1999: 51.3 %, 2003: 46.9 %, 2007: 49.3 % ; these rates are calculated on the basis of valid units (excluding business numbers, collective households, households with no voting rights, and invalid numbers). The sampling frame has so far been the public directory, which means households with a registered landline phone only (Lutz 2008).

[12] There is no common practice in Switzerland and no recommendation by the vsms, for instance, concerning the calculation of response rates. Reasons for non-interview such as language problems, physical or mental inability, and absence during the field period are sometimes considered ineligible, against the AAPOR standard definition. The exact formula is rarely communicated; in some cases, the response rate is not even published.

The Swiss population is accustomed to telephone interviews and generally feels confident with this mode. Even sensitive subjects such as sexuality and violence in couples can be surveyed with success in this mode in Switzerland.

However, if people are asked which interview mode they would prefer for a survey, the telephone is not necessarily in the first position, despite or maybe partly because of its predominance in the Swiss survey landscape. In effect, respondents often prefer the mode in which they have been interviewed. On one extreme, we have the case of the MOSAiCH2009 survey (which includes ISSP 2008 and 2009), a face-to-face CAPI survey of about one hour at the respondent's home, where an additional paper questionnaire was self-administered with, among others, a question on the preferred mode. People predominantly prefer the known face-to-face and paper modes; telephone is chosen only by 2% (and by 7% as second choice, see Table 3.3). On the other extreme, in the Swiss Household Panel (SHP), a CATI panel survey that is in its 11th wave, people predominantly prefer the known telephone mode, and a face-to-face interview at one's home is chosen by only 6% of the respondents.

In these two surveys, the response rate is fairly comparable: 50% for MOSAiCH and 47% for first-time recruited respondents to the SHP 2004[13] (Graf 2009) plus about 7% attrition per year (Graf 2010). But, as pointed out in the previous chapter, the telephone sample does not cover the entire Swiss population.

Table 3.3 Mode preference by survey mode in percent, an indicative example

Survey	MOSAiCH 2009	Swiss Household Panel 2009
Mode of survey	Self-administered paper after 1 hour CAPI (Face to face)	Telephone (CATI), Panel
Question	"Generally, for an interview of 45 minutes, would you prefer to answer…" (first choice out of two)	"If you could choose, would you prefer to answer our interview…"
By telephone	2°	66*
Face to Face	42°	6*
Online	10°	28*
Paper (selfadm.)	46°	- Choice not offered as not feasible for this survey -

°Sample design weighted data, n = 796; *Unweighted data, n = 7,109.

Even though the Swiss population is accustomed to telephone surveys and the survey agencies generally have the necessary know-how to treat the respondents well, the success of this mode begins to show its limits in this small country.

In 2008, for example, the members of the vsms carried out as much as 4.2 million interviews, excluding panel interviews, and almost half of them were conducted on the telephone. This means that one adult out of three responds to a telephone survey per year. If we consider the relatively weak response rate, the numerous contact attempts, the agencies who are not members of the vsms and,

[13] This rate is based on the number of households, where at least one person completed the interview.

above all, the calls from direct marketing[14], we can understand why serious survey agencies complain about falling cooperation. Another indicator is the fact that, overall, the number of CATI interviewers grows with the constant number of telephone interviews, despite technology that minimized the demands on human resources such as automatic dialing. The number of calls has to be multiplied, which increases the burden on the population with a landline telephone and reduces the cost-effectiveness of telephone surveys. All CATI interviewers can attest to the growing amount of effort needed to get people involved in basic exchange and to convince them to participate.

These results are only some elements of a beginning discussion that has to be conducted about the optimal mode or mix of modes for science-driven surveys in Switzerland.

3.5 Outlook

As we have seen, Switzerland was and still is a telephone survey country. CATI interviews are cost-effective, and owing to large experience, the field management and the fieldwork itself are accomplished with professionalism by most agencies. There is no legal restriction upon calling published numbers even for marketing purposes, until recently coverage was comparatively satisfying and the structure of the numbers makes Random Digital Dialing reasonably feasible.

The telephone survey culture is well established, so that searching for alternative paths is rare, at least in the circle of survey agencies. CATI will probably continue to be a dominant mode for surveys that have to be very reactive (like opinion polls), have lower quality requirements (like market research) and/or low budgets. The only apparent trend is the development of online surveys. Online surveys have slowly gained weight and the major survey institutes in Switzerland invest massively in this mode. The low cost of this mode is very attractive for their clients, but there still are considerable problems in drawing truly representative samples.

However, decreasing coverage of the sampling frame and decreasing cooperation and response rates push scientific and administrative institutions in charge of large-scale surveys to look for solutions in other modes. Sampling on the basis of buildings and letterboxes is one possibility for a new sampling frame of the Federal Statistical Office; merging the whole resident population into one database, could be another way for a few selected surveys. These two sampling procedures assure good or even almost perfect coverage, but only for the very expensive face-to-face interviews.

Another direction is to consider solutions that combine different modes of data collection. Mixing modes probably will be the most promising path to assure high-quality surveys in the future. Even the Federal Statistical Office is beginning to think about such a way. Nevertheless, sufficient research has not yet been done in the national context to rate the impact of different designs of mixing modes on data quality and make sufficiently advised choices. Several dimensions have to be

[14] It is possible to indicate in the directory that one does not wish to be bothered by commercial calls, but most direct marketing enterprises do not respect this. They often masquerade their call as a simple survey.

taken into account: the population coverage of the sampling frame, response rates, nonresponse bias (from a socio-demographic but also attitudinal point of view; Stoop et al. 2010), measurement error related to subsequent modes and costs. The next few years will surely be challenging.

References

Bieri, U., et al.: Schlussbericht zur Studie. Mobilfunkmonitor 2008. Final Report on the Study. Mobile Monitor 2008, Gfs. bern, Bern (2008)

Eva, G., Roberts, C., Lynn, P.: Measuring the Effect of Interview Length on Telephone Survey Cooperation. In: American Association of Public Opinion Research (AAPOR) 63rd Annual Conference, New Orleans, USA, May 15-18 (2008)

Gabler, S., Häder, S.: Machbarkeit von Random Digit Dialing in der Schweiz. Random Digit Dialing feasibility of Switzerland. Bundesamt für Statistik, Neuchâtel (2007)

Graf, E.: Pondérations du Panel Suisse de Ménages. Weights Swiss Household Panel. OFS, Neuchâtel (2009)

Graf, E.: Etude empirique de l'attrition du Panel suisse de ménages. Empirical Study of Attrition of the Swiss Household Panel. OFS, Neuchâtel (2010)

Joye, D., Ernst Stähli, M., Pollien, A., Roberts, C., Sapin, M.: Survey climate in Switzerland. Paper Presented at the Annual International Workshop on Comparative Survey Design and Implementation, Lausanne (March 2010)

Joye, D., Schöbi, N., Kaenel, C.: Nonresponse and nonrespondents in Switzerland. Paper Presented at the ESRA, Prague (2007)

KIG Bericht der Koordinationsgruppe Informationsgesellschaft (KIG) an den Bundesrat vom 30. Report of the Information Society Coordination Committee (ISCC th the Federal Council of the 30th April). BAKOM (April 2001)

Schmugge, S., Grau, P.: Sind die SchweizerInnen überhaupt noch zu erreichen? Are the Swiss people to achieve at All?, Link Focus 14 (2000)

Link Institute. In: Hofte-Fankhauser&, K., Wälty, H.F. (eds.) Sind Telefonumfragen noch repräsentativ (Are telephone surveys still representative?), vol. 2, p. 66. Compendio Bildungsmedien, Marktforschung (2009)

Lutz, G.: Eidgenössische Wahlen 2007. Wahlteilnahme und Wahlentscheid. Federal elections in 2007. Electoral participation and electoral vote. Lausanne: Selects-FORS (2008)

M.I.S. Trend. ESS Telephone Experiment: Technical Report for SIDOS. Lausanne (2007)

M.I.S. Trend/OFCOM. Étude auprès des usagers des services mobiles de télécommunications. Survey of users of mobile telecommunications, vol. 1. Rapport final pour, OFCOM (2007)

M.I.S. Trend/BAKOM. Studie über den Internetanschluss und dessen Nutzung in der Schweiz. Schlussbericht für das Bundesamt für Kommunikation. Study on the Internet and its use in Switzerland. Final Report for the Federal Office of Communications (2009)

OFCOM. Statistique officielle des télécommunications 2008. Official telecommunications statistics (2008)

Publica Data AG. Die KommTech-Studie 2010. The study KommTech 2010, Bern (2010)

Schallenberger, P.: Branchenstatistik. Industry Statistics 2008. In: Vsms Jahrbuch 2009, pp. 8–12. Verband Schweizer Markt- und Sozialforscher (2008)

SMS. Jahrbuch 2001. Verband Schweizer. Markt-und Sozialforscher 2010 (2002)

Stoop, I., Billiet, J., Koch, A., Fitzgerlad, R.: Improving Survey response: Lessons learned from the European Social Survey. Wiley, West Sussex (2010)

vsms swiss interview institute®. Vsms-Jahrbuch, Markt- und Sozialforschung (2010)

Chapter 4
Finland

Vesa Kuusela and Matti Simpanen

4.1 Introduction

Knowing the telephone coverage is nowadays more difficult than it used to be. Up to mid 1990s, it was sufficient to know how many households own a landline telephone but the increasing proportion of mobile phones has changed the situation radically. Solving the methodological problems is not straightforward because of the different nature of mobile and landline telephones, and because many households use both types of telephones concurrently.

In Finland, as compared to many other countries, the popularity of mobile telephones started to soar relatively early (Kuusela et al. 2007). In the beginning of this trend, there was some concern that interviewing on mobile telephones would influence the survey results. Therefore, Statistics Finland conducted several surveys in the 1990s to monitor the changes in telephone coverage and to assess potential influences of mobile telephones on survey quality.

4.2 Telephone Coverage in Finland

The history of telephony in Finland has distinctive characteristics: Private local telephone companies have been dominant since the emergence of telephones. This set the foundations for competitive markets which later yielded a versatile supply of services for mobile telephones. Another significant factor has been the early development of mobile technology (most notably by Nokia) that brought affordable devices within nearly everybody's reach. The cost of using mobile telephones has always been competitive, and unlike for example in the North American countries receiving a call is free of charge. Only when a Finnish mobile telephone is used outside of Finland, one has to pay half of the roaming cost to receive a call, like in all European countries.

A specifically Finnish phenomenon has been that various services for mobile telephones emerged in an early stage. For example, already with the second generation of mobile phones it was possible to pay for the car wash or parking or to order and pay for movie tickets. Another example is the quickly spreading application for secure personal identification via mobile telephones. The 3G mobile phone systems with access to the Internet have dramatically increased the services provided. For example, social media, such as the Facebook and Twitter, can now be reached with a 3G mobile telephone.

In Finland, the attitudes towards telephones are and have been slightly different to those in many other countries. This can be observed in some aspects relevant to telephone surveys. Mobile telephones became popular in the mid 1990s and already then they were often listed in directories. Currently, nearly all telephones, both mobile and landline telephones, are listed in a common database. Listing includes the owners' names and addresses. Less than 5 % of all telephones have secret numbers and they are found mainly in large cities. In the countryside, secret numbers are rare. In addition, prepaid mobile phones are rare in Finland: In 2004 when this was last studied, only 3.4 % of mobile phones were prepaid (Kuusela et al. 2007). Since then, little has changed.

4.2.1 Changes in the Structure of Telephone Coverage

Since the mid 1990s, Statistics Finland has carried out four similar surveys (in 1996, 1999, 2001 and 2005) that mainly focused on the structure of telephone coverage and the usage of mobile telephones (Kuusela 1997, 2000)[1]. In addition, the telephone coverage in Finland has been followed continuously in connection with surveys carried out in the CATI centre. Already the first survey in 1996 indicated that the use of mobile telephones was growing fast.

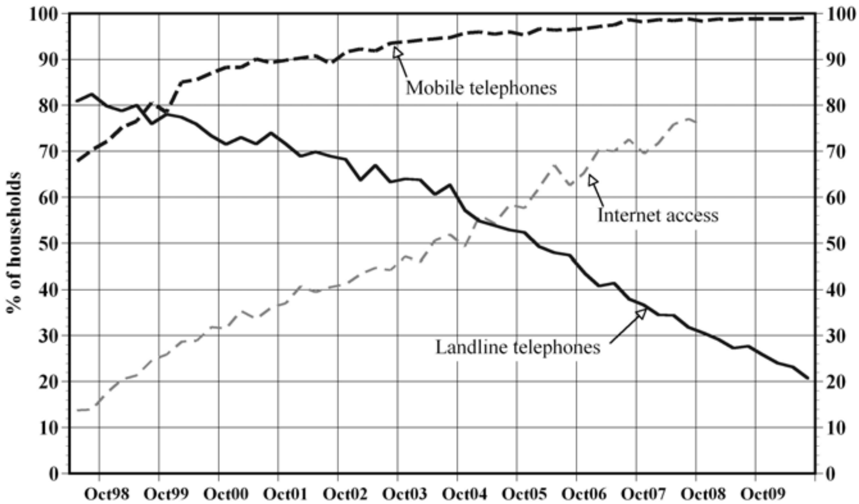
The rapid change in the structure of telephone coverage started when the second generation mobile telephones were introduced in the mid 1990s. The new generation models were considerably smaller and their batteries lasted longer than the earlier models. The structure of telephone coverage changed considerably already in the first half of the 1990s but it accelerated in the second half of the decade. The coverage of landline telephones has been decreasing since 1993 (see Figure 4.1). Already in 1999, the popularity of mobile telephones surpassed that of landline telephones. Interestingly, the growth in the percentage of households with Internet access shows similar patterns as mobile phones did earlier.

The fear that the rapid change in the structure telephone coverage may have some impact on the results of the telephone surveys was the most significant reason to carry out surveys in late 1990s. The results of these surveys and the influence of mobile telephones on survey quality have been reported by Kuusela and Notkola (1999), Kuusela and Vikki (1999), Kuusela and Simpanen (2002),

---

[1] Telephone coverage and coverage structure and the way telephones are used in a modern society can only be reliably estimated through a survey designed for this purpose. Telephone coverage and its structure are so complex that general purpose surveys may be biased. The accuracy of coverage estimates based on service provider data is also questionable because they do not reflect the actual usage of telephones. Only very few countries have reliable surveys focused on telephone coverage. In some cases the published data are based on methods whose accuracy may be questioned, and these results may lead to discrepancies between estimates published by different sources for the same country.

and Kuusela (2003). The conclusion was that the device itself, i.e. mobile or landline telephone, had no noticeable impact on the length of the interview or on the survey results. However, some differences were observed, which were due to differences in the population segments interviewed by mobile telephone or by landline telephone (Kuusela et al. 2007).

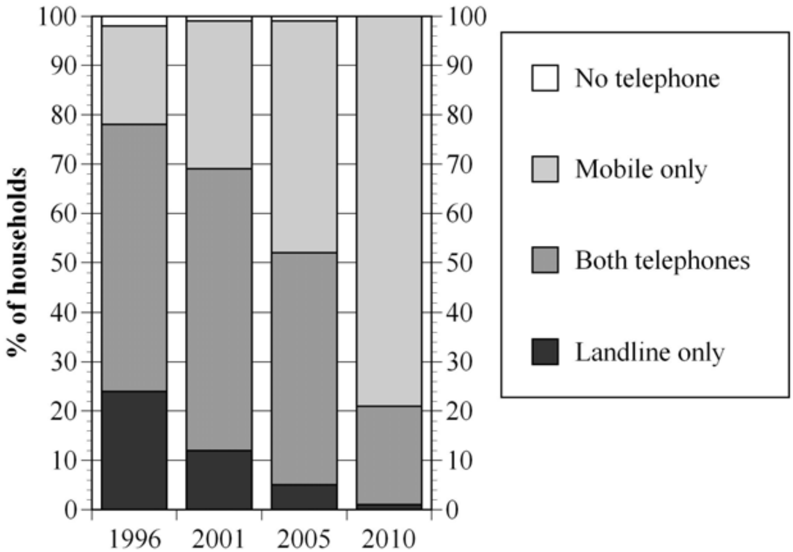


Source: Statistics Finland

Fig. 4.1 Development of mobile and landline telephone coverage in Finland from May 1998 to November 2010. The popularity of Internet access via landlines was not measured anymore since 2008 because of the fast growth of the wireless Internet access using mobile technology.

The rapid increase of mobile telephone coverage slowed down at the end of the previous millennium, and the market was more or less saturated by 2007. In 2010, more than 99 % of households had one or more mobile telephones, while only 20.7 % had a landline telephone; and a little less than 20 % of households had both types of telephones. In addition, 79.1 % of households had only mobile telephones and less than 1 % only a landline telephone. Figure 4.2 shows how the structure of telephone coverage has changed in Finland during the last fifteen years.

The remarkable change in the last decade has been the steady decrease of landline subscriptions in Finland, and this is still continuing. It is possible that in the future landline telephones will disappear completely. The reasons for the downward trend are two-fold: In the 1990s, newly established households characteristically did not acquire a landline telephone in their new apartment but existing subscriptions in



Source: Statistics Finland

Fig. 4.2 The change in the structure of telephone coverage in Finland from 1996 to 2010

"old" dwellings were kept. Since the beginning of the new millennium, more and more households have given up their landline subscriptions even when they did not move. Probably, the most important reason is the pricing: In landline subscription there is an initial fee that must be paid even if the telephone is not used, while with mobile telephones one pays for the use only.

Currently, nearly all of the landline telephones in Finland are owned by elderly people. For example, 38 % of pensioners' households had a landline telephone in 2010 and 2.6 % of them had only a landline. The other typical population segments owning landline telephones are farmers and entrepreneurs: 28 % of farmers had a landline and 2.3 % only a landline; 32 % of entrepreneurs had a landline telephone but they all have also mobile telephones. In all other population groups, landline ownership was below 20 % and virtually all had also a mobile telephone.

The stability of people's life situation is the most significant predictor for landline telephone ownership. For example, 30 % of households that own their dwelling also have a landline telephone, while among those households that rent their apartment only 9.2 % have a landline telephone. According to telephone operators, new landline telephone subscriptions are acquired only occasionally.

Apart from the landline telephone ownerships, also its use is declining. According to the telephone operators in Finland, landline telephones are used less than would be their relative share among all telephones. The telephone communication habits are changing, probably because the mobile phone is available nearly all the time regardless of user's whereabouts. In addition, the fact that telephone numbers (including those of unanswered calls) are in the memory of the mobile telephone makes using them convenient.

A mobile telephone is a personal appliance, like a wrist watch. Therefore, looking at the coverage at the household level may be slightly misleading. Single adult households are the most revealing population group in this respect. In Finland, the single adult households are the most common household type with 41% of all households at the end of 2009. Figure 4.3 shows the structure of telephone coverage of single adult households: Only in the oldest age group do a few households own landline telephones only (3.6 %). In younger age groups, everyone has a mobile telephone and landline telephones are rare: 2.6 % in the age group 15 - 34 years and 7.5 % in the age group 35 - 54 years.

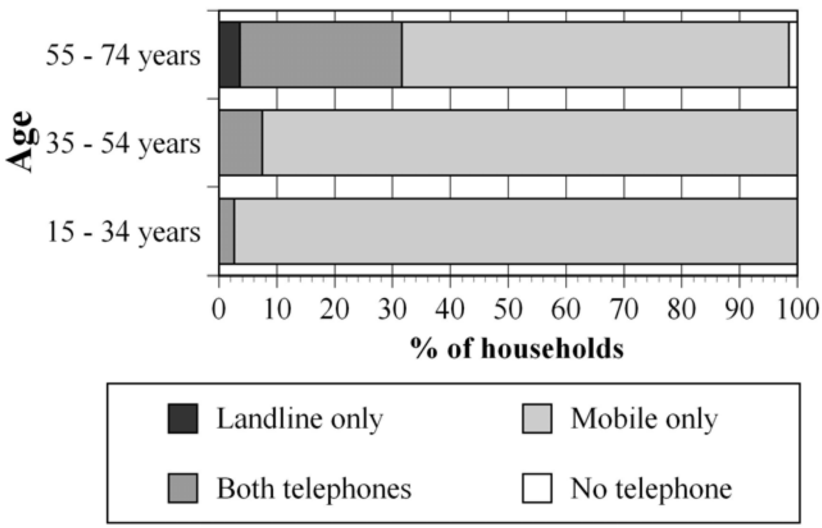


Fig. 4.3 Structure of telephone coverage of single adult households in Finland in 2010

Practically all working-age people in Finland have a mobile telephone (more than 99 %), and its popularity is roughly at the same level as with teenagers. In nearly every household, each member has a mobile telephone of his or her own. Children usually get their own mobile phone when they go to school. Only elderly people, especially elderly women, have limited mobile telephone coverage (10 %).

Mobile telephones have also become mundane in Finland and they are used as everyday appliances throughout society. In 2004, nearly all (98 %) persons aged 15 - 34 reported that they get frequent calls on their mobile telephones, and 91 % receive frequent SMS text messages (Nurmela et al. 2004). Not only households and individuals give up landline telephones. In addition, companies and establishments are increasingly dismantling their landline telephone systems and are relying only on mobile telephones in the office use. Companies may still retain the old landline number but the incoming (and outgoing) calls are automatically directed to mobile telephones.

4.3 Telephone Surveys at Statistics Finland

In Finland, the constantly updated population register makes sampling relatively easy. All the samples for the surveys that Statistics Finland conducts are drawn randomly among the eligible population in the register. The register contains information about an individual's name, age, gender, address, and many other data, but does not contain telephone numbers. However, telephone operators maintain a common database that is publicly available at a reasonable price. The database includes all post-paid telephone numbers (i.e. no pre-paid phones), both for mobile and landline telephones. In standard survey practice, the sample file of individuals (drawn from the population register) is merged with the telephone number database. Through this process, a telephone number is found for approximately 65 % of sampled persons. Next, numbers are searched for other people in the same household. This roughly produces a number for 15 % of the remaining sample. After this, telephone numbers are searched for the rest of the sample by manual methods using the same database. On average, at least one telephone number is found for 95 % of a random sample.

Interviewers are instructed to first call on the landline telephone, if such a number is available, because it is slightly less expensive. If a sampled person does not have a landline phone or cannot be reached on the landline,, the interviewer calls on the mobile telephone. After each contact, including interviews and refusals, interviewers code the type of phone used for the contact.

As mobile telephones have become everyday devices, it is quite natural to be interviewed on the mobile telephone, as well. The proportion of interviews conducted on mobile telephones has been increasing continuously. Currently, more than 90 % of all telephone interviews are conducted on a mobile telephone.

A problem in using mobile telephones for interviewing is that the respondents can be virtually anywhere. This is especially a problem if respondent is abroad (and paying for half of the roaming costs). If interviewer observes that a respondent is abroad, he or she is instructed to make an appointment for the interview when the respondent returns. However, it may be very difficult to notice that the mobile telephone is outside the country unless the respondent points this out. If the respondent does not mention to the interviewer that he or she is abroad, the additional roaming costs are considered as their responsibility.

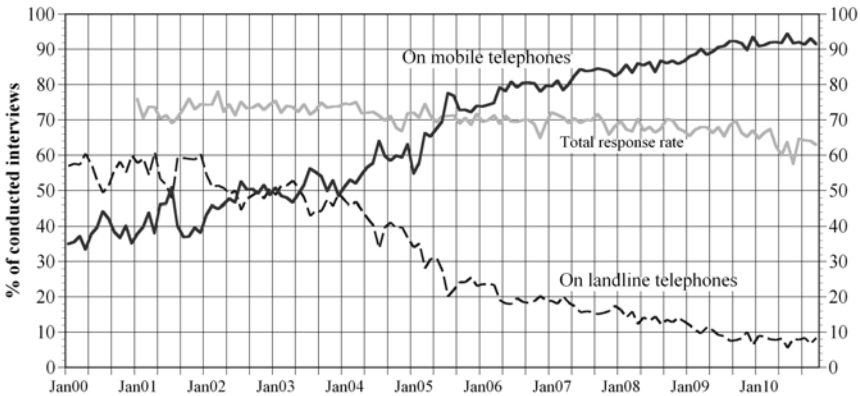


Fig. 4.4 The proportion of interviews conducted on mobile and landline telephones in the CATI centre of Statistics Finland, from January 2000 to November 2010. In the beginning, the proportion of mobile telephone interviews was noticeably greater during the common holiday months. Response rate is calculated from the whole sample and it includes also those respondents for whom no telephone number was found. The response rate is calculated according to the guidelines given by Callegaro et al. (2007).

In some CATI surveys, interviewers have also recorded where the respondents were when they were interviewed. The results are given in the Figure 4.5. Mostly the interviews were conducted at home (79.0 %) and occasionally at work (8.2 %) or in a private car (4.8 %). 3.0 % of interviews were conducted in a public place, including busses and trains.

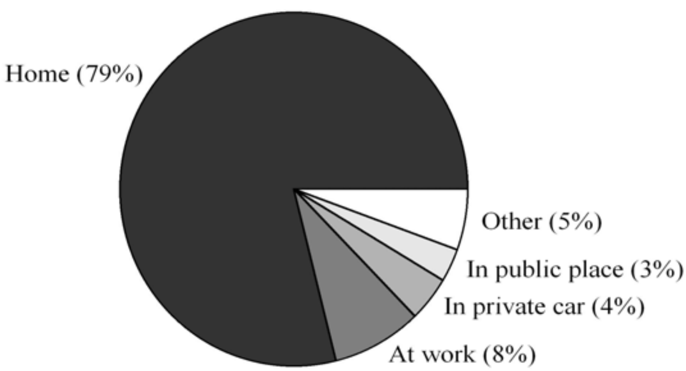


Fig. 4.5 Distribution of places where respondent were when interviewed. (Based on 47.500 interviews at the CATI Centre of Statistics Finland in the period starting in March 2007 and ending in May 2009.)

Another question that may arise is how quiet the place is where the interview takes place. In most cases respondents were alone (61.8 %) or with the other members of the family (31.1 %). Very rarely the interview was carried out while respondent was within hearing distance of strangers (1.3 %).

As can be seen from the Figure 4.4, response rates stayed fairly stable up to 2005 but then they started to decline. It is unknown what has caused this. The same amount of telephone numbers have been found for respondents and no special reason causing nonresponse has come up. On the other hand, the same trend has also been observed in face-to-face surveys.

4.4 Discussion

In Finland, mobile telephones have not had any noticeable influence in telephone surveys. Obviously this is due to the fact mobile telephones have been regarded as "normal" telephones since their emergence. The infrastructure around mobile telephony has actually facilitated telephone survey operations. The possibility to draw samples from the population registry and to merge telephone numbers to these samples makes telephone surveys relatively simple to carry out.

Mobile telephones do not seem to disturb interviewing more than landline telephones. In addition, respondents have been reached most often at home and then it has been of minor importance which type telephone was used. An interesting feature was observed earlier when mobile telephones were not as common as they are now: During the most common holiday months the proportion of mobile telephone interviews was markedly higher than at other times. This, in turn, kept the nonresponse rates during the holiday season at the same level as before and afterwards.

RDD is not used at Statistics Finland because it is not necessary and it would probably also be inefficient. RDD was developed for landline telephone systems in which the number space has regular patterns which can be used to increase the efficiency of RDD. However, the number space of mobile telephones is not regular in the same way. Therefore, it is difficult to find an algorithm that would increase the frequency of valid telephone numbers with random number generation (see also Callegaro et al. 2007).

In addition, it is not clear how to draw an unbiased RDD sample from the current number space. In Finland, there are three major networks and few minor ones which all operate countrywide. The numbers of users and the user profiles are different across networks: one is favored by business men and women; one is popular among younger users; in one network users live outside the metropolitan area more frequently. Sampling with equal inclusion probabilities would probably yield biased samples and finding a valid weighting method would require a lot of work.

Even if RDD were possible, its use would bring up problems because the inclusion probabilities of large households would increase because most members also have mobile telephones. A different problem is posed by the fact that roughly a quarter of mobile telephone users are younger than 15 (Kuusela et al. 2007). In most surveys they are not part of the target population and if they were interviewed their guardian should be present. In addition, RDD would hinder the standard practice of Statistics Finland of sending an advance letter to all sampled persons.

References

Callegaro, M., Steeh, C., Buskirk, T.D., Vehovar, V., Kuusela, V., Piekarski, L.: Fitting disposition codes to mobile phone surveys: experiences from studies in Finland, Slovenia and the USA. J. R. Statist. Soc. A 170(3), 647–670 (2007)

Kuusela, V.: Puhelinpeittävyys ja puhelimella tavoitettavuus Suomessa. In: Telephone Coverage and Accessibility by Telephone in Finland. Statistics Finland, Helsinki (1997)

Kuusela, V.: Puhelinpeittävyyden muutos Suomessa. Change of telephone coverage in Finland, Helsinki, Finland (2000)

Kuusela, V.: Mobile phones and telephone survey methods. In: Banks, R., et al. (eds.) ASC 2003 - The Impact of New Technology on the Survey Process. Proceedings of the 4th ASC International Conference, pp. 317–327. Association for Survey Computing (ASC), Chesham Bucks (2003)

Kuusela, V., Callegaro, M., Vehovar, V.: The Influence of Mobile Telephones on Telephone Surveys. In: Lepkowski, J.M., et al. (eds.) Advances In Telephone Survey Methodology, J.W. Wiley and Sons, Inc., New York (2007)

Kuusela, V., Notkola, V.: Survey quality and mobile phones. In: Paper presented at the International Conference on Survey Nonresponse, OR, Portland (1999)

Kuusela, V., Simpanen, M.: Effects of mobile phones on phone survey. Practices and results. In: Paper Presented at the International Conference on Improving Surveys (ICIS), Copenhagen, Denmark (2002)

Kuusela, V., Vikki, K.: Change of telephone coverage due to mobile phones. In: Paper Presented at the International Conference on Survey Nonresponse. OR, Portland (1999)

Nurmela, J., Melkas, T., Sirkiä, T., Ylitalo, M., Mustonen, L.: Finnish people's communication capabilities in interactive society of 2000's. Statistics Finland, Helsinki (2004)

Chapter 5
Great Britain

Elizabeth Nelson

5.1 The Beginnings in the UK

5.1.1 Telephone Surveys in a Country with a Face-to-Face Survey Tradition

"Compared with the 1950's …we are suffering from falling response rates, through fear of crime, selling under the guise of research and generally less confidence in the value and confidentiality of surveys, and the sheer difficulty of finding people at home during social hours…" It was "socially acceptable for a lone woman to walk in darken streets on her own. In those halcyon days 50 % of interviews were achieved on the first call: there was less out of home entertainment, the shops closed early and the evening meal, eaten early, was a family event." (Mitchell 1995: 63).

By the 80s researchers were beginning to see, that in some situations, telephone had advantages over face-to-face. As market research continued to spread outwards from fast moving consumer goods, to take in new areas of business such as finance, healthcare and business-to-business global markets, telephone interviewing was seen as a cheaper and quicker method (Saxton-Howes 1981).

Face-to-face interviewing remains the gold standard in government surveys. The adoption and acceptance of survey methods by government was kick-started by World War II. In later years for many government departments and for the Central Office of Information, the largest commissioner of survey research, face-to-face interviews are the preferred survey mode (McDonald and King 1996).

Likewise Social and Community Planning Research (now called the National Centre for Social Research) which exclusively conducts social surveys, uses random sampling and predominantly face-to-face interviewing and is known to be in a position to carry out government surveys to the level of detail and accuracy required. The best known example is the British Social Attitudes (BSA) Survey, which was run by Prof Roger Jowell up from 1983 until 2001. Prof Jowell, now Sir Roger Jowell, at City University heads up the Central Co-ordinating Team responsible for the ESS (European Social Survey). The same rigorous face-to face interviewing techniques with highly trained interviewers are used in this biennial survey. In all 30 countries have been surveyed.

Much government survey work demands these high response rates often over 50%, which involve many return calls to the same address at different times of the day and different days of the week until the desired response rate has been achieved. (Park 2010)

The National Readership Survey (NRS) is the gold standard in the media research world, following many of the rules of government surveys. NRS is a continuous survey of adults aged 15+ in Great Britain designed to measure the readership of newspapers and magazines. Face-to-face interviewing has remained the interviewing mode since the 1930s. The survey now covers over 350 titles. 36.000 interviews with adults aged 15+ in March - April are conducted every year and the average interview takes 27 minutes. Response rates have decreased from 77 % in the 60s but remain above 50 %.

Telephone interviewing was never seriously been considered for the NRS because its readership measurement method is dependent on visual aids (mastheads), particularly important in relation to magazines and occasional readers.

On the other hand, very recently Northcliffe publishers have piloted the use of telephones and it is anticipated in the industry, that because of reduced budgets, more government and other prestigious surveys will be conducted via telephone interviews.

5.1.2 The Market Research Development Fund (MRDF)

A seminal event occurred in 1981. The Market Research Development Fund (MRDF), set up by The Market Research Society (MRS), launched a project examining telephone interviewing. Its findings were that after many years in which the face-to-face interview had dominated UK market research, there was an increasing interest in telephone research methods among both research suppliers and buyers (McDonald 1981).

The 1973/1974 National Readership Survey showed that telephone ownership penetration declined by social grade[1] (Miln 1989). This was supported by Wilson et al. in 1988, who reported that 96 % of senior managers and professionals compared with 23 % of unskilled and semi-skilled manual workers. UK telephone ownership was also low among the unemployed, in manual occupations, low education, in rented properties and where the head of household was young and single (Wilson et al. 1988). On the other hand insignificant differences were reported in patterns of consumer behavior in telephone vs. non-telephone households and researchers concluded that weighting for corrections in representation is acceptable (Miln 1981, Hyett and Allen 1976).

[1] Social Grade includes detailed questions about the occupation of the Chief Income Earner (CIE) to establish social grade. Questions include not just what the CIE's job is, but also details such as their qualifications and the number of people they are responsible for. A = Higher managerial, administrative and professional (4 %); B = Intermediate managerial, administrative and professional (22 %); C1 = Supervisory, clerical and junior managerial administrative and professional (29 %); C2 = Skilled manual workers 21 %; D = Semi-skilled and unskilled manual workers (15 %); E = State pensioners, casual and lowest grade workers, unemployed with state benefits (8 %).

By 1997 UK household telephone ownership was around 85 % and by 2006 the ownership was between 90 % and 95 %. 45 % of research agencies thought that telephone surveys (compared with 16 % face-to-face interviews) were more representative (Taylor 1997, Wilson and Bradley 2006).

However, an important distinction is made about the situations where telephone interviewing is suitable and not suitable. When the research is designed to consider the opinions, attitudes and habits of particular user groups, the samples of respondent with access to the telephone are satisfactory. Where the research is required to measure accurately the occurrence in the total population of a particular factor, such as readership or voting intentions, then telephone interviewing can produce different estimates from face-to-face interviewing (Weitz 1990). Intimate subjects, however, can be asked and answered more easily in a phone interview (Weitz 1990: 177, Hyett 1982: 56-71).

Skeptics were highly critical of the representativeness of telephone ownership and raised objections to the selection of telephone numbers and the quality of interviewing. The latter concern helped launch the Interviewer Card Scheme (subsequently renamed IQCS, Interviewer Quality Control Scheme) in 1977. Its aim was to improve consumer acceptance of surveys.

Invasion of privacy and surveys carried out by persons not bound by the MRS Code of Conduct led to significant numbers of respondents refusing to be interviewed. One of the biggest problems was the abuse of public cooperation known as “selling under the guise of market research” (SUGging) (IQCS 1976 - 2010, Weitz 1990).

5.2 Telephone Interviewing Becomes Fully Established

5.2.1 Unlisted Numbers

For many years British researchers used the Directory Plus 1 method, selecting a sample of numbers from directories and adding the integer ‘one’ to those numbers in order to generate a sample of both listed and unlisted numbers. By the end of the 1980s many agencies were moving over to RDD: generating random numbers, rather than choosing from directories (Foreman and Collins 1991). RDD was effective in lessening the biasing effect of unlisted numbers on the selected sample.

Noble et al. successfully used a real probability sampling approach using a two stage RDD sample. Compared with other research firms, they produced closer predictions of 1997 election results (Noble et al. 1998).

5.2.2 Selection of Respondent within a Household

The choosing one household member to answer question by means of the next birthday method (Hyett 1981) overcame problems of identifying the respondent. Asking for the details of every household member could seem both irrelevant and suspicious. The CATI systems can now effectively select the person to be interviewed from any household.

5.3 Response Rates in the UK

Members of MRS use the guidelines for conducting telephone surveys of the American Association for Public Opinion Research (AAPOR 2008, 2010). In 1981 Market Advisory Services Ltd quoted response rates of 95 % for telephone interviews, but in the 80s, response rates fell and researchers developed ways of improving response including repeated call-backs. Making three calls before declaring a number unproductive proved the most cost-effective procedure (MAS 1981). The first few seconds of a call are crucial for persuading the respondent to continue; after that dropping out is rare (McDonald 1981).

Nevertheless in 1981, 41 % of the general population found telephone interviewing the least acceptable form of contact (O'Keefe 1981). Other methods have been used to improve response rates e.g. leaving messages on answering machines, including in the introduction 'I am not selling anything' (de Leeuw and Hox 2004) and offering incentives at the beginning of the interview (Brennan et al. 2005). Guidelines are available in the Interviewing Quality Control Scheme (IQCS) scheme (ICQS 1976 - 2010).

Response rates for telephone interviewing have been decreasing steadily since their inception in the 80s. Rates now typically range from 10 % to 40 %. A number of factors account for the decline in response rates, including caller ID, voicemail and the Telephone Preference Service (TPS). The TPS is a central opt out register whereby individuals can register their wish not to receive unsolicited sales and marketing telephone calls. It is a legal requirement that companies do not make such calls to numbers registered on the TPS, although market research is exempt.

5.4 Best Practice

5.4.1 Professional Standards

Monitoring systems continue and the validation methods have stayed very similar to those developed in the last two decades of the 20th century. The IQCS is now an independently run scheme, which requires members to adhere to a set of benchmark market research industry standards, including:

- Providing the MRS's Freephone Service to respondents
- Guidelines on training interviewers and supervisors and on back-checking interviews
- Monitoring of interviewers using languages other than English must be carried out. Interviewers are tested to ensure that they understand the briefing and are able to translate open responses accurately. Link et al. report that real-time interpretation during a survey expands the number of languages in which surveys are offered.
- 5 % of calls on any survey are listened to in real time or retrospectively (Link 2009, IQCS 1976 - 2010)

5.4.2 Incentives

The majority of researchers conducting telephone interviews do not offer incentives to encourage respondents to take part in telephone interviews. There are exceptions, for example, when the list of phone numbers has been supplied by the client. Entry into prize draws might be offered.

Many business-to-business (B2B) surveys and telephone surveys among professionals, very small universes and difficult to reach groups will involve offers of donations to favorite charities, prize draws, playing a game ‘gamification’ or a straight fee. Incentives are also offered where the interview is long e.g. over 30 minutes in length or where the respondents are required to complete an additional task such as testing a product or completing a diary.

In these cases in the 80s and 90s, respondents might be sent a client’s voucher on receipt of their completed extra survey element and/or be entered into an overall prize draw (Wilson et al. 1988). In the 2000s giving clients’ vouchers came under scrutiny by the MRS which ruled that clients’ products or vouchers cannot be used as incentives.

5.4.3 Questionnaire Design

Up to 25 - 30 minutes is an acceptable interview length for members of the public, but this depends on the subject and how well the questionnaire is designed. Excellent advice on questionnaire design including “unfolding” can be found in Weitz (Weitz 2002). Evidence comparing results from face-to-face and telephone interviewing is confused. It is impossible to generalize since the subject of the questionnaire can affect the response using different modes.

It is recommended that respondents must hear essential words more than once in a conversational fashion. The conversational repetition of key words helps the respondent mentally to organize the entire question before attempting to answer (Locander and Burton 1976, Wheatley 1973).

Questionnaire layout is important to ensure that the interviewer is never forced to break the natural flow. Explanatory material tends to be fuller (than in face-to-face) and is fully written out. Interviewers should not have to hunt for fillers or sequences (Market Research Development Fund 1981).

It is not safe to rely on the respondent writing down the items with pencil and paper. Stimulus material including advertising materials and mastheads can be sent ahead to the respondent by fax, or viewed online.

5.5 Interviewers

The pool of face-to-face interviewers typically middle-class, middle-aged women who were looking for ‘pin money’ was drying up in the 1970s. Employment opportunities for women were developing quickly and certainly less arduous, even less dangerous work could be done in other occupations.

From the 80s onwards it was recognized, however, that a different type of person is required for telephone interviewing. Voice, phone manner and competence on the phone are all important (van der Vaart et al. 2006).

Training for telephone interviewing is short because interviewers work directly under supervision in a controlled environment. Young people, ‘resting’ actors, students, and those looking for evening jobs are suitable as it was usually done in a shift work environment.

5.6 Is the Gold Standard Changing?

As mentioned above. the UK government continues to prefer face-to-face interviews. However, Sport England, a huge project of the Department of Culture, Media and Sport, uses telephone interviews. TNS, and previously IPSOS MORI, conduct this survey which is thought to be the largest survey among consumers in Great Britain. The number of interviews in the most recent and third survey was 166,000. The sample size is not the only unique characteristic of this survey. In the future, interviewers will, via random digit dialing, also contact those households which have no landline home phone but have a mobile phone (estimated to be a quarter of the general population). The mobile device will be used in the same way as a landline hand set, i.e. interviews will ring the mobile number and conduct the survey with the appropriate household member.

5.7 The Impact of Online Research on Telephone Interviewing

5.7.1 Telephone Interviewing as a Proportion of the Total

By 1997 telephone interviewing had become 20 % of market research turnover (Weitz 2002). In the same year some research agencies were looking to develop online research. The acceptance of online was slow. ESOMAR did not begin to collect data on the proportion of interviews done online until 2004.

According to Tony Cowling[2], this is in contrast to the US where National Family Opinion Panel (NFO) report a faster growth in new methods of collecting data. NFO, one of the largest panel operations, conducted 100 % of their interviews via mail in the 1960’s. By the 70s nearly 50 % had moved to telephone research. By mid-1995, online had moved from 0 % to 70 %.

Table 5.1 shows the impact of online survey modes in the UK confirming Cowling’s prior statement that in the UK online research had an impact on face-to-face not on telephone or postal interviewing.

[2] Interview with Tony Cowling, President of TNS July 2010.

Table 5.1 Interviews carried out in the UK via different modes in percent

Year	Postal	Telephone	Face-to-face	Online
1988	10	18	72	-
1995	10	23	67	-
2004	10	24	44	22
2009	12	22	32	34

By the 90s the refusal rate to participate in a telephone survey before determining eligibility was 30 - 40 % (Nobel et al. 1998). Nevertheless, in spite of increasing numbers of unlisted telephone numbers and the blurring of telemarketing and telephone research (Weitz 2002), telephone research has maintained its share of the market.

5.7.2 Why Telephone Interviewing Maintains Its Strength

Multi-mode Research

Multi-mode research is now common (de Leeuw and Hox 2004). Global surveys and panels collect data via online, mail and phone. Many of the short-listed and winning papers in the MRS Awards 2009 use telephone interviewing as part of a multi-mode research project (Pugh and Warrington 2009).

Other Factors in Favor of Phone Research

While the skepticism about online research continued into the new millennium, online and phone surveys were producing similar results. The predictability of behavioral measures was essentially equivalent but online produced significantly lower cost per contact (Roster et al. 2004).

Similarity in results was found in qualitative data across online survey, telephone and postal surveys (Coderre et al. 2004).

On-the-fly translations during a survey can expand the number of languages in which surveys are offered (Link et al. 2009).

Cowling maintains that the strategy of research companies is to continue to use telephone interviewing as a good balance of the cheaper and representative; and to use telephone interviewing for tracking where a new sample is needed in each time period.

Furthermore, it is the writer's contention that respondents will increasingly expect incentives for answering surveys by landline and mobile phone.

Hodson argues that market research is being pushed towards nano-surveys. Very short cheap and quick surveys will be carried out on iPhone and other smart phones. These will be directed towards groups of respondents about whom enough is known when the sample is drawn for the questionnaire not to be the sole source of demographic data (Hodson 2008). This leads naturally into mobile telephone interviewing which is often associated with nano-surveys.

5.8 Mobile Market Research

5.8.1 The Beginnings of Mobile Internet Research

In 1947 the cellular concept was discovered at Bell Telephone, but commercialization of analogue cellular networks did not take place until the 1980s.

The subsequent growth in usage and ownership of mobile phones has been phenomenal. It is estimated that in 1998, there were around 200 million mobile customers worldwide and that at the close of 2004, these had dramatically increased to about 1.6 billion, and likely to reach 2.6 billion by 2006 (Steinbock 2007).

The market research industry has been slow to respond, noting low penetration of use of Internet mobile facility, biased ownership and buying in samples leading to unrepresentative lists of mobile users, while failing to recognize that a mobile Internet survey is just another form of an online survey and the mobile screen is perfectly readable, and uses the same software as a pc.

In the first mention of mobile market research in the International Journal of Market Research (Okazaki 2007), the author extols the advantages of Internet mobile as an interviewing mode. The author does not mention Short Message Service (SMS) which requires one question to be sent and the answer sent back, with this process being repeated for all questions. An online survey on a mobile screen is restricted in terms of the number of questions appearing but it uses exactly the same software as online surveys on a computer.

Okazaki recognized that limitations are due to micro–browser screen space and scrolling-down capacity but these are being overcome. Researchers now invite respondents by SMS to click on to the questionnaire on the Internet site. The cost of receiving the survey, accessing the mobile Internet and answering must be compensated by incentives. The incentives are typically higher than for online pc surveys but other costs make the mobile survey cheaper, quicker, and often a better medium for full responses to open-ended questions and qualitative research.

Wilke[3] of Global Park is a mobile Internet survey pioneer, who became involved in 2005, but surveys via mobile Internet were first introduced by Q Research. In the UK Wilke has overseen the expansion of software used for online panels to mobile panels. Moreover he is working on a project, which will bring back the power of face-to-face interviews. Global Park UK is currently working on a large project (over 100,000 interviews per annum) with a client market research company on store exit face-to-face interviews where the interviewers are using mobile phones as CAPI devices over mobile broadband connections. Market research agencies have expressed interest since they visualize that they can get face-to-face data in real time. This was not what Wilke anticipated but believes that market researchers using mobile phones for this type of data collection, will discover the benefits of the mobile phone as a self-completion methodology in due course.

[3] Interview with Wilke, A., Head of Global Park UK, July 2010.

5.8.2 The Advantages of Mobile Internet Research

Capturing data in real time by mobile phones is powerful: those attending an event, i.e. attending a football match or shopping can be asked what they see, which promotions/advertisements interest them.

Cowling sees that the QR Bar Code, the unique bar code developed by Google, will be a motivator in increased use of mobile phone interviewing. It lets customers and potential customers instantly respond to a new concept/product and can be scanned by the camera on a mobile phone. The code can then link the customer to a questionnaire about a specific product or promotion.

Online mobile and pc surveys share several characteristics: lower costs and speedier results. They will be increasingly used for collecting sound and visual clips from respondents, and therefore leading to more extensive ethnographic analyses.

The powerful argument for mobile Internet research is that it is a personal possession. Younger people feel more confident about answering surveys via their mobile phone than using a pc. Older respondents carrying the device can choose to answer a survey when it best suits them.

Fly Research reports that among a mobile panel of eleven to 35 year olds, the response rate is never less than 40 % and averages 55 % (Nelson 2009).

5.8.3 Internet Access

A recent survey by CCS Insight, the digital analytics company, shows that a quarter of young people now access the mobile Internet every day. The launch of the iPhone, has significantly increased the proportion of mobile owners accessing the Internet on their mobile and downloading apps (Bennison and Davidson 2010).

Whereas pc online access is biased towards upper social classes, mobile usage is similar across demographic groups and is well spread across educational level and age (up to 55 years of age).

Males are more active users of mobile phones but a higher percentage of females answer mobile surveys. (Nelson 2009)

A 2006 survey among members of The Market Research Society found that

- Online mobile response is faster: 30 % of surveys answered within two hours.
- Mobile is the preferred device for 18 - 25 year olds: 48 % mobile/SMS, 28 % email, 20 % instant messages.
- Young people are more active on their mobile phone than on their pc.
- 95 % of mobile texts are opened vs. 25 % of emails.
- Mobile phones allow easier access to business people and access to developing countries.
- There are increasing numbers of mobile internet sites and telecommunications companies are less inclined to limit data access to data providers.
- Clients can communicate directly with the consumer – it is the most personal device (Wilson et al. 1988, Noble et al. 1998).

To conclude, telephone research whether landline or mobile will continue to thrive.

Acknowledgements. I would like to thank all those who helped me with the history of telephone interviewing and its present-day status in market and government surveys: Dawn Mitchell, Tony Cowling, Alex Wilke, Bill Blyth, Simon Burrows, Mark Edwards, and Phyllis Vangelder.

References

AAPOR Cell Phone Task Force (The American Association for Public Opinion Research). Guidelines and Considerations for Survey Researchers when Planning and Conducting RDD and Other Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers (2008), `http://www.aapor.org` (accessed December 30, 2009)

AAPOR Cell Phone Task Force (The American Association for Public Opinion Research), New Considerations for Survey Researchers when Planning and Conducting RDD Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers (2010), `http://www.aapor.org` (accessed January 24, 2011)

Bennison, C., Davidson, I.: Getting onside with consumers by getting inside their mobiles. Market Leader Quarter 2 (2010)

Brennan, M., Benson, S., Kearns, Z.: The effect of introductions on telephone survey participation rates. International Journal of Market Research 47(1), 65–74 (2005)

CCSInsight. Private survey conducted by Fly Research on Social Networks, `http://www.ccsinsight.com/`

Coderre, F., Mathieu, A., St Laurent, N.: Comparison of the quality of qualitative data obtained through telephone, postal and email surveys. International Journal of Market Research 46(3), 347–357 (2004)

de Leeuw, E.D., Hox, J.J.: I am not selling anything: 29 experiments in telephone introductions. International Journal of Public Opinion Research 16(4), 462–473 (2004)

Dillman, D.A., Gallegos, J.G., Frey, J.H.: Reducing refusal rates for telephone interviews. Public Opinion Quarterly 40 (1976)

Dillman, D.A.: Mail and telephone surveys: the total method. John Wiley & Sons Inc., New York (1978)

Foreman, J., Collins, M.: The viability of random digit dialling in the UK. Journal of the Market Research Society 33(3), 219–227 (1991)

Glasser, G.J., Metzger, G.D.: National Estimates of Non-listed Telephone Households and their Characteristics. Journal of Marketing Research 12(8) (1975)

Hodson, M.: The Market Research Society. Research, London (2008)

Hyett, G.P.: Telephone interviewing. Telephone Research Proceedings of a one day Conference. MRDF, London (1981)

Hyett, G.P., Allen, G.M.: Collection of data by telephone (CDT): its viability in consumer research. MRS Conference. Market Research Society, London (1976)

IQCS. Interviewing Quality Control Scheme papers 1976-2010. The Market Research Society, London (1976-2010)

Link, M.W., Battaglia, M.P., Osborn, L., Frankel, M.P., Mokdad, A.H.: Measuring the quality of real-time telephone survey interpreters. International Journal of Public Opinion Research 21(3), 271–292 (2009)

Locande, W.B., Burton, J.P.: The Effect of Question Form on Gathering Income Data by Telephone. Proceedings of a one day Conference, p. 33. MRDF, London (1976)

Market Research Development Fund. In: Proceedings of a one day Conference, p. 33. MRDF Market Research Development, London (1981)

MAS. In: Contribution in Telephone Research Proceedings of a One Day Conference, p. 22. MRDF, London (1981)

McDonald, C.: Telephone Surveys. Telephone Research Proceedings of a One Day Conference, pp. 3–44. MRDF, London (1981)

McDonald, C., King, S.: Sampling the Universe. NTC Publications Ltd. (1996)

Miln, D.: Contribution in Telephone Research Proceedings of a one day Conference. MRDF, London (1981)

Mitchell, D.: Association of Market Survey Organisations (AMSO). In: McDonald, C., King, S. (eds.) Forty years On Conference Paper Quoted in Sampling the Universe, p. 63. NTC Publications Ltd., Oxon (1995)

Nelson, E.: The potential of mobile research: Implications for the future and the role of industry standards. In: World Association of Public Opinion Research Annual Conference, WAPOR, Lausanne (2009)

Noble, I., Moon, N., McVey, D.: 'Bringing it all back home...' – Using RDD telephone methods for large-scale social policy and opinion research in the UK. Journal of the Market Research Society 40(2), 93–120 (1998)

O'Keefe, J.: Telephone Surveys. In: Telephone Research Proceedings of a one day Conference, p. 51. MRDF, London (1981) (quoting a survey by Bowen, J.)

O'Shaughnessy, J.: Competitive Marketing; a Strategic Approach. Unwin Hyman Ltd, London (1995)

Okazaki, S.: Assessing mobile-based online surveys: methodological considerations and pilot study in an advertising context. International Journal of Market Research 49(5), 651–675 (2007)

Park, A. (ed.): British Social Attitudes 26th Report. Sage (2010)

Pugh, A., Warrington, J.: Education is the best provision for old age – Why we couldn't agree more with Aristotle. Westminster City Council, London (2009)

Roster, C.A., Rogers, R.A., Albaum, G., Klein, D.: A comparison of response characteristics from web and telephone surveys. International Journal of Market Research 46(3), 359–373 (2004)

Saxton-Howes, R.: Contribution. Telephone Research Proceedings of a one day Conference, p. 53. MRDF, London (1981)

Steinbock, D.: The Mobile Revolution: the making of mobile services worldwide. book review. International Journal of Market Research 49(2), 275–276 (2007)

Taylor, H.: The very different methods used to conduct telephone surveys of the public. Journal of the Market Research Society 39(3), 421–432 (1997)

Van der Vaart, W., Ongena, Y., Hoogendoorn, A., Dijkstra, W.: Do interviewers voice characteristics influence cooperation rates in telephone surveys? International Journal of Public Opinion Research 18(4), 488–499 (2006)

Weitz, J.: Telephone Interviewing. In: Birn, R.J. (ed.) The International Handbook of Market Research Techniques, pp. 147–163. Kogan Page Limited, London (1990)

Weitz, J.: Telephone Interviewing. In: Birn, R.J. (ed.) The International Handbook of Market Research Techniques, pp. 177–201. Kogan Page Limited, London (2002)

Wilson, A., Bradley, N.: Survey conducted for The Market Research Society (MRS) involving MRS research agency members (2006)

Wilson, P., Blackshaw, N., Norris, P.: An evaluation of telephone interviewing on the British Labour Force Survey. Journal of Official Statistics 4, 385–400 (1988)

Chapter 6
Italy

Teresio Poggio and Mario Callegaro*

6.1 Introduction

CATI has been widely used in Italian surveys since the 1980s. As a matter of fact telephone interviewing, compared to traditional face-to-face data collection methods, allows to save time and money. It also offers further advantages: better control over the quality of the interview through a centralized facility, continuous monitoring and supervision, and the opportunity to include experiment-like variables in the survey (Groves et al. 1988, Piazza and Sniderman 1998). Institutional data producers, academic researchers and companies in the market research and polls industries have been taking advantage of all these opportunities.

Sampling designs adopted in phone surveys and polls range from non-probabilistic quota sampling to more valuable probabilistic sampling approaches, similarly to what happens in face-to-face surveys.

Voting lists and population registers represent the typical frames for face-to-face surveys. The former are publicly available and cover all Italian citizens having political rights while institutional data collectors may use the latter in order to cover foreign residents, too. Public phone directories have traditionally been used for the same purpose in landline phone surveys on the general population. In this last case, once a household owning a sampled telephone number has been contacted, through randomization techniques such as the nearest birthday method or the Kish grid, a single individual may be selected for interview within the household.

Dedicated phone directories – clients' records, for instance – may be used for specific sub-populations as in the case of customer surveys, instead of the white pages. To the best of our knowledge, the former also represent the only adopted sampling frame for mobile phone surveys and polls. The setting up of a comprehensive phone directory (for which the Italian acronym is DBU) – listing both landline and mobile phones, irrespective of the different operators – has been promoted by the Authority for Communications in 2002. Various types of printed

---

* The contents of this paper do not commit the University of Trento and Google Inc. Responsibility for the ideas and opinions expressed are entirely under the authors.

and electronic phone directories are produced and distributed on the basis of available DBU data. For privacy reasons, citizens may choose to register or not in the DBU. In practice, almost all private mobile phones are unlisted in phone directories.

In both landline and mobile phone surveys in Italy RDD is rarely used as a technique to select the sample units. It is considered inefficient and expensive. Landline phone directories have been favored as a sampling frame, given their reasonable coverage. During the 1990s households with no landline phone at home were about the 9.5 % of all households (Callegaro and Poggio 2004: 485). They were also a rather homogeneous group: mainly elderly families, with a low level of education, residing in rural areas. Public phone directories used to be rather comprehensive at that time. Having one's number not listed was only an issue for VIPs or people with serious security problems. Only 6 % of landline phones were not listed in the white pages in 2001 (Callegaro and Poggio 2004: 503). However, this relatively favorable context has rapidly worsened during the last decade (see section 6.3).

6.2 The Reality of Phone Surveys in Italy

How are phone surveys used by the different organizations in the field?

The Italian National Institute of Statistics (ISTAT) is the main institutional data producer in the country. It still collects most of its survey data via personal interviews, either PAPI or CAPI. This is especially the case when all household members have to be interviewed, when major measurement issues are involved or – in any case – when the questionnaire is particularly time-demanding.

However, there are some exceptions where phone surveys are used instead of personal interviews. All the three studies surveying employment outcomes after the completion of – respectively – secondary schools, higher education, and doctoral studies are carried out through phone surveys. The availability of dedicated alumni's phone directories, as a sampling frame, is probably a big reason for this choice.

Another example of ISTAT taking advantage of phone surveys is when individuals are re-interviewed. The Labor force survey design implies four subsequent interviews to the same individual within the same year. The first one is carried out with CAPI techniques, on a sample obtained from population registers, while CATI is used for further interviews. The same approach is used in a few dedicated phone surveys whose sample is extracted from respondents to a larger previous survey conducted through personal interviews.

Finally, phone surveys have also been used by ISTAT in another few surveys: the Women's Safety Survey, the Citizens' Safety Survey, the Business Trips and Holidays Survey and the Birth Survey. A phone survey has been preferred for the former two studies because of measurement considerations, given that some of the questions were particularly sensitive. Phone directories represent the sampling frame for these two surveys and for the third study. The Birth Survey is an interesting methodological experiment: the households sample is obtained from the population registers. It is then possible to link parents' full names to public phone directories in order to obtain the contact number. Only the matched households could be interviewed: in the 2002 survey this linked “factual sample” represented only 60.4 % of the sample originally obtained from the population registers. Furthermore, another 21 % of the households in this restricted factual sample were excluded from the survey because the matched phone number was either wrong or non-existent[1] .

The quality of the data from these last four surveys is problematical. They rely – either directly or indirectly – on phone directories as a sampling frame while it is clear that these lists have serious coverage problems (see section 6.4).

ISTAT phone surveys do usually obtain a good response rate. However, this is also the result of a “mandatory answer” policy for major public surveys.

Phone surveys are very common in the polls market where the timing of data delivery is a crucial matter. The CATI approach has actually allowed the flourishing of this industry.

The research design of opinion polls, whose results are reported by the media, are - according to Italian law - to be publicly available for auditing purposes. Similar special provisions exist for political and electoral polls. Two open archives for methodological information on polls used by the media have been established by the Authority for Communications and by the Government Department for Information and Publishing[2]. Despite the law aiming to make public auditing possible, the quality of the methodological information provided in the two archives is generally poor and ambiguous, to the point that it is often impossible to understand even key issues: whether or not a sample has been defined on a probabilistic basis, or which sampling frame has been used (Billari and Rosina 2006, Di Gioia 2009, Gasperoni and Callegaro 2007, Pisati 2008). In fact it is impossible to assess the quality of the polls. For this reason, the Authority for Communications, who controls both the depositing procedures, established stricter regulations in December 2010.

Despite many deficiencies, the two archives still provide valuable information on how phone surveys are carried out in the market, under the assumption that the methodologies adopted for polls used by the media are not different from the ones used in surveys that have not been made public[3].

Gasperoni and Callegaro (2007, 2008; Callegaro and Gasperoni 2008) investigated the accuracy of polls before the general elections of 2001, 2006 and 2008. They scrutinized the methodological information available on these polls in the electoral polls archive. They focused on the pre-electoral period polls and on national surveys targeted to the whole electorate of the Chamber of Deputies. Overall they considered 69, 73 and 72 different polls respectively, for the three years. To our purposes, their main findings may be summarized as follows.

---

[1] Own calculations, or reported data, from figures presented in ISTAT (2006: 65, 72), that provide a detailed and very interesting discussion of the Birth Survey methodology.

[2] Bill number 249, 31st July 1997, article 1, and bill number 28, 22nd February 2000, article 8, respectively. The two archives are available at the following web sites: http://www.agcom.it and http://www.sondaggielettorali.it.

[3] See also Gasperoni (2007) for a detailed discussion on electoral pollsters in Italy.

The quality of the polls' documentation is poor and it is possibly getting worse, but available indications allow to reasonably assume that quota sampling is the industry standard and that phone directories represent the typical sampling frame.

CATI is the common mode adopted for data collection. Personal interviews have definitely become marginal while web surveys have a limited importance: nine polls have been carried out in the latter mode in 2001, five in 2006 and six in 2008; twelve mixed mode surveys – mostly CATI and CAWI – have been reported among the ones conducted in 2008. Unfortunately, no reliable information is available for an assessment of the response rate and of item nonresponse due to an undefined proportion of undecided voters and non-voters.

As a way to complement this inventory we have scrutinized the methodological reports available in the general polls archive (all but the political ones) for all the 233 surveys – targeted to general population – that have been carried out in the first semester of 2010 (registered in the archive as to December 31th 2010). Most of the polls were carried out with CATI methods; 31 were web-based and seven used personal interviews; five used dedicated telepanel devices and 14 a mixed-mode approach. Typically, no information on the sampling design and frame is provided in a methodologically appropriate way. Instead, it is only claimed that the sample is representative of the population of interest. Quota sampling is rarely explicitly acknowledged but the usual reference to random and stratified – but not probabilistic – procedures and the typical length of the field period (1 - 2 days in 64 % of the phone surveys) allows us to reasonably assume that quota sampling is the standard.

How phone surveys are used in market research is less visible. In order to get some insight, we have reviewed the information available to potential buyers – on their web site – for their main services. We have considered major members of ASSIRM (Associazione tra Istituti di Ricerche di Mercato, Sondaggi di Opinione, Ricerca Sociale), the main association for market, social and opinion research professionals. Our focus is on general population surveys rather than on specific targets surveys.

Almost all companies offer a wide range of services as regards data collection modes – CAPI, CATI, CAWI, postal questionnaires, experiment-like panels and other computer-assisted methods using dedicated remote meters, smartphone-like – depending on their area of expertise and on the clients' needs and budget.

All the companies offer ad hoc solutions. However, their "multiclient" or "omnibus" surveys, their panels and – in general – their studies that are repeated over the time, probably offer a clearer picture of common practices and standards.

CAPI surveys and panels are generally offered, apparently based on samples obtained from the voting lists, as a standard top product. This is especially the case for established companies. Corresponding CATI surveys are also generally offered as a standard product which allows fast data delivery and a low budget alternative. No information is usually available on the sampling design and frame of the latter. Most of the companies offer CAWI solutions, at least for some areas of interests. According to the information available, they are based on opt-in panels. In some cases, the adoption of solutions aimed at controlling the effects of self-recruitment is declared but not explained.

Some of the companies adopt and offer telepanels, apparently established on the basis of a sample from the voting list, through dedicated communication devices.

Data on the industry expenditures provide some further indications on the role of different types of surveys, even if the differences in the typical costs should be taken into account. According to the Esomar industry report (2010: 102), Italy's market research industry spends 85% on quantitative research, 12% on qualitative and 3% on other kinds of research. Within the quantitative research segment, CATI accounts for 19% of the expenditures, while face to face surveys account for 27% and online surveys for 7%. The automated, digital and electronic meters segment accounts for the largest quota: 43%. Other quantitative research including postal surveys accounts for only 4% of total expenditures.

Finally, it has turned out to be impossible to carry out a systematic inventory of academically-led surveys. To our limited knowledge, they have a high variability in the mode of data collection and in their sampling design. Data collection is often out-sourced to companies in the field (see above) and the survey strategy can usually be agreed with them, according to their common practices. We can reasonably assume that phone directories represent the sampling frame adopted in phone surveys for the general population. We can also reasonably assume that landline – rather than mobile – phones are targeted the most and that these studies share the same methodological problems as examined in this chapter for other market and institutional data collectors.

Important exceptions subsist, but in general it should be considered that academic research funding is a big issue in Italy. For this reason, low cost – and often low quality – solutions are more likely to be preferred. In this sense it is not surprising that Italy is currently not participating in the European Social Survey and that one of the main reasons for that seems to be the costs of national face-to-face surveys.

To conclude, we would like to remark that mobile phones seem to be rarely used in Italian surveys.

6.3 Main Recent Changes in the Technological and Social Context

The technological context of personal communication has been rapidly changing over the last two decades, the same as in the other European countries. The diffusion of mobile phones, the use of the Internet and, more recently, the progressive convergence between information and communications technologies – Internet on mobiles, smartphones, VOIP solutions – are setting new opportunities and constraints for survey activity.

On the consumers' side, the diffusion of mobile phones across all social strata of the Italian society has led to an overall theoretical telephone coverage of 97.5% of all individuals (see Tables 6.1a and 6.1b) - never seen before – considering both landline and mobile phones. Mobile phones have become affordable to all pockets as has the cost of calls, if one considers the myriad of available promotional

Table 6.1a Personal phone arrangements, listing of the landline phone number in the white pages and use of Internet connection, by selected individual and household characteristics - 2009. Percentage values – 18 y. o. individuals or older.

	Phone arrangements			
	no phone (1)	landline only (2)	mobile only (3)	both type (4)
Sex				
Male	1.9	9.7	26.1	62.3
Female	3.0	15.8	22.8	58.4
Age				
18-30 y. o.	0.5	0.8	32.7	66.0
31-40 y. o.	1.0	1.0	41.1	56.9
41-50 y. o.	1.3	2.7	26.2	69.8
51-60 y. o.	1.6	7.4	19.2	71.8
61-70 y. o.	2.9	20.5	13.9	62.7
older	8.2	50.4	8.0	33.4
Formal education				
Low level or none	7.5	38.5	16.9	37.1
Lower secondary	1.6	7.5	32.1	58.8
Upper secondary	0.5	3.5	25.0	71.0
University level	0.3	2.4	17.3	80.0
Condition				
Employed	0.5	1.8	29.6	68.1
Unemployed	1.6	2.5	39.4	56.5
Retired	4.5	34.2	10.4	50.9
Housewife	4.8	20.2	24.7	50.3
Student	0.2	0.2	17.3	82.3
Other	11.3	35.2	16.4	37.1
Type of household				
Single	6.6	22.7	35.3	35.4
Couple	2.4	21.6	19.9	56.1
Couple with children	0.9	5.5	22.6	71.0
Lone parent with children	3.0	12.1	27.4	57.5
Other	4.0	18.3	24.5	53.2
Housing tenure				
Tenant	3.4	8.3	48.2	41.1
Home owner	2.2	13.9	18.4	65.5
Other	3.1	12.4	31.1	53.4
Geographical area				
Northwest	1.5	13.2	20.4	64.9
Third Italy	1.6	14.7	19.9	63.8
South & Islands	3.7	11.4	29.9	55.0
All Individuals	*2.5*	*12.8*	*24.4*	*60.3*

Notes: (1) - (4) Landline availability means that the subject owns a landline phone at home; mobile phone availability refers to individuals using a mobile phone at least once a week regardless of whose the property is; (1) + (2) + (3) + (4) = 100 %.

Table 6.1b Personal phone arrangements (*continued*)

	Number not listed in white pages	*Total landline non-coverage*	*Have connection and use of the Internet* at home	on mobile
	(5)	(6)	(7)	(8)
Sex				
Male	18.1	41.0	41.4	4.0
Female	18.6	39.6	31.0	1.8
Age				
18-30 y.o.	22.0	47.9	64.6	6.5
31-40 y.o.	25.8	57.0	50.3	4.6
41-50 y.o.	21.3	42.9	45.3	2.8
51-60 y.o.	17.6	34.7	31.6	1.8
61-70 y.o.	14.1	28.5	12.9	0.5
older	10.1	24.7	2.0	0.0
Formal education				
Low level or none	11.5	33.1	2.0	0.1
Lower secondary	18.5	46.0	25.7	2.0
Upper secondary	21.2	41.3	56.0	4.5
University level	23.8	37.2	74.0	5.7
Condition				
Employed	22.6	45.9	53.2	4.4
Unemployed	19.6	52.6	39.8	4.5
Retired	11.9	25.0	9.4	0.3
Housewife	16.9	41.4	11.6	0.6
Student	18.8	33.0	83.5	5.9
Other	17.7	40.5	12.6	1.2
Type of household				
Single	20.4	53.8	21.3	2.0
Couple	16.1	34.8	22.4	1.8
Couple with children	18.6	37.7	47.3	3.7
Lone parent with children	20.1	44.4	38.1	2.9
Other	18.3	41.6	26.2	2.1
Housing tenure				
Tenant	27.9	65.1	27.6	3.0
Home owner	16.9	34.0	38.0	2.9
Other	19.1	46.8	35.4	2.4
Geographical area				
Northwest	21.9	39.0	39.5	3.0
Third Italy	14.9	33.2	38.6	3.0
South & Islands	18.7	46.0	32.1	2.6
All Individuals	*18.4*	*40.4*	*36.0*	*2.9*

Sources: own calculations from a 38,386 individuals sub-sample (aged 18 or over only) from the ISTAT survey Aspetti della vita quotidiana 2009.

(5) Percentages calculated on the ones with a landline phone available [(2)+(4)].
(6) is the total landline non-coverage (no phone or not in the white pages): (6) = (1) + (3) + [(2) + (4)] x (5) / 100. (7) - (8) The availability and use of an Internet connection at home refers to people having a connection to the Internet at home and have surfed the Internet at least once over the three months prior to interview; the availability and use of an Internet connection on mobile phones refers to people who have reported using an Internet connection from mobile phones or palmtops. "Third Italy" refers to the Northeast of the country and some of the regions in the centre of Italy (Emilia Romagna, Tuscany, Umbria and Marche).

schemes: 84.7% of Italian adults regularly use a mobile phone. This gives the opportunity – at least in principle – to use mobile phone surveys in order to speed up data collection and to reach those who are hard to find at home in an easier way.

A similar trend is evident when considering the recent popularization of smartphones that allow access to the Internet when moving. This will soon make self-administered mobile surveys possible. However, it is still to be considered that, according to the latest available estimates, only 2.9 % of adult individuals regularly use mobile connections (see Tables 6.1a and 6.1b).

Finally, the diffusion of the Internet gives the opportunity of self-administered surveys to be conducted via web. According to our estimates, 36 % of Italian adults has an Internet connection at home and regularly uses it. This may progressively lead to web surveys taking the place of the phone surveys anytime when there is no need for an interviewer.

Please consider that when estimating the use of mobile phones, of the Internet at home and on mobile we have relied on definitions stricter than simply owning such devices and connections (see notes to Tables 6.1a and 6.1b for details). Our interest is in their factual use, which is a more reliable precondition for the possibility to contact individuals for a survey. This may result in figures for the diffusion of mobile phones, the Internet at home and on mobile that are lower than the ones shown in other studies reporting ownership figures, often at the household level.

From the production side, phone surveys certainly contributed to the flourishing of the market research and polls industry in the past. They have allowed centralizing the field and the data management, to simplify data collection and data cleaning procedures. Overall, when compared to face-to-face interviews, phone surveys require less personnel, time and money. The liberalization of the telephone market – since the late 1990s – and the development of VOIP technologies have allowed further cuts on the costs of these activities.

On the other hand, the previously mentioned technological developments have also involved new methodological and logistics challenges to phone surveys.

Coverage error in landline phone surveys is in our opinion the most relevant one: 28% of the Italian households (ISTAT 2010: 40) own only mobile phones and no landline. This is a big issue in a country where surveys are generally carried out using phone directories as a sampling frame (see section 6.2).

Some other problems are also worth noting. Before 1998 a single state-owned company used to manage all landline telephone services and three companies the mobile ones. With the liberalization of the telephone market the number of actors in the field has increased. ISTAT (2008: 8) estimates that 52 different companies were managing mobile or landline phone services at the end of 2007. It is worth noticing that most of them were landline operators, while the mobile market is less fragmented (four operators only), given the higher costs to enter this market.

Furthermore, VOIP operators have recently started to use standard – geographically structured – telephone numbers. So, the latter are increasingly assuming virtual connotations and this may for example result in VOIP users living abroad to use a telephone number with Milan or other cities prefix. Because of both consumers' protection laws and marketing strategies, the telephone number portability between different operators has also been increasing.

Overall, these changes puzzle the setting of phone numbers, making it impossible to distinguish between landline and merely virtual phones, to use geographical information built in landline numbers and to discriminate by phone operator. This is clearly challenging when defining the survey strategy and logistics and in terms of eligibility assessment. This is especially true if one considers that it is a quite common practice to use more than one mobile working with different operators in order to take advantage of the different promotional schemes[4].

Furthermore, without a national authority, the multiplication of phone operators makes an effective cooperation between the operators and the organizations working with surveys to make the phone numbers structure understandable, practically impossible.

Several changes have also affected the broader social context and related legal provisions. When considering the use of the telephone, privacy has progressively become an issue, albeit if with some contradictory behavior. So while people commonly have very personal dialogues on their mobile phones in public spaces, they prefer that nobody calls them for phone surveys and especially telemarketing at home. Insisting marketing campaigns by phone operators, or by other companies, and polls for which people show no particular interest, have definitely become more and more common than in the past. Hence, receiving an unsolicited call is generally perceived as annoying. This might be one of the reasons why 19.1 % of the Italian households owning a landline phone are not listed in the white pages (own calculations, see Table 6.1b for data source).

The Italian Data Protection Authority and the legislators are still handling this issue at a normative level. In principle, Italian households wishing to be listed in the phone directories have been given the opportunity to give their consent to receiving unsolicited calls or not (market research surveys and telemarketing are considered to be the same under the Italian law)[5]. Companies in the field have eluded these regulations by using previous editions of the white pages, which still remain a useful public data source given the low mobility of Italians, and other similar trickeries. A “do not call” register has been established in February 2011 to counter such practices[6]. Citizens may ask to be recorded there at no cost while operators in the sector will have to check their potential contacts’ list against this register in order to get a list of phone numbers which can be contacted. Such list should be updated every 15 days.

To conclude, both social behavior and legal provisions regarding individual data protection have radically transformed the usability of phone directories as sampling lists.

---

[4] 90.2 million SIM cards were estimated active at the end of 2007 (ISTAT 2008: 2) vis-à-vis 56.3 resident individuals aged six or older.

[5] Italian Data Protection Authority, Act 15 July 2004.

[6] More details are available on the register web site: http://www.registrodelleopposizioni.it.

6.4 Coverage Error as the Big Issue in Phone Surveys

As we have already discussed in Callegaro and Poggio (2004) and anticipated in the previous section, coverage of sampling frames adopted in landline phone surveys is our main concern when thinking about the future of this type of inquiry. When dedicated lists – clients, alumni, for instance – are available as both a sampling frame and a source of contact information, there is clearly no problem. As such phone surveys organizations may fully take advantages of this, if their intention is to generalize to some specific sub-populations: customers, high school alumni.

The problem arises when surveys intend to provide estimates for the general population – or general sub-populations – and when phone directories are, directly or indirectly, used as a sampling frame. At the individual level, their non-coverage sums up to 40.4 % (see Table 6.1b) when considering "mobiles only" families, individuals with no phone at all and members of households with a landline phone that is not listed in the white pages. As already noted, this figure is likely to rise further.

Coverage error is a function of both non-coverage – whose magnitude has been assessed above – and the difference between covered and uncovered populations in the variable of interest (Groves 1989). Researchers can usually gain some control of non-coverage, while the variable of interest is typically an object of research.

However, we should consider that non-coverage is clearly structured along important social differentiation variables, as shown in Tables 6.1a and 6.1b consistently with previous research (Callegaro and Poggio 2004, Marbach 2010). Also investigated human behaviors, attitudes and opinions are usually related to such variables. This means that – with just some hypothetical exceptions – estimates from landline phone surveys are likely to be affected by a significant coverage error, because of both the high level of non-coverage and its association with typical variables of interest. Households' and individuals' characteristics are likely to influence both coverage and almost all potential variables of interest.

It is surprising to see how polls companies and even institutional and academic data collectors are underestimating the problem.

Web-based inquiries are emerging as an alternative to phone surveys, especially in the market research and in the polls industry. They may be used with a view of cutting costs further. They are also supposed to be representative of the Italian population. Indeed, according to the sources we have scrutinized, most of the web surveys are based on non-probabilistic opt-in samples. Moreover, they do not include the non-Internet population, estimated at about 64 % of Italian adults (see section 6.3)[7].

---

[7] Higher estimates are usually reported for having connection to the Internet. They usually refer to the household level. ISTAT for instance reports that 47.3 % of Italian households have an Internet connection at home (ISTAT 2010: 37). We have preferred to provide individual level figures, reflecting both within-household differences and the real use of the Internet. See notes to Tables 6.1a and 6.1b for details.

6.5 Conclusions: No Way to Skip a Low Cost–Low Quality Vicious Circle?

Is it possible to trace some of the most likely developments for phone surveys in Italy, given the ever-changing technological and social context and the coverage error implicit in the use of phone directories? Let us focus on general population surveys for which more quality problems exist.

Here we can trace different scenarios. The first one represents the worst hypothesis: low quality-low cost phone surveys will be carried out on the grounds of a substantial statistical illiteracy of both typical buyers and the general audience. Samples will continue to be defined on a non-probabilistic basis or, in the probabilistic case, to rely on increasingly unreliable sampling frames. No strict auditing practices – we here consider both legal provisions for polls and peer-judgment in the scientific community – will make an easy life for low quality-low cost solutions. Official statistics and academics will also continue to ignore coverage problems; if honest, they will limit their estimates' generalization to the households in the white pages only, regardless of the lack of substantiality of such a selection. As already mentioned, this kind of telephone surveys may be also replaced by low quality-low cost web surveys, based on self-recruited samples of some sort.

A second, more optimistic, scenario considers the attempt to conciliate the possibility of combining all the methodological and logistics advantages of phone (and web) surveys with the preservation of the capacity to generalize results to the whole population, on probabilistic grounds. The big issue here is the sampling frame to be used or, alternatively, the possible ways to control the coverage error deriving from public phone directories.

Different strategies may be adopted here, as we have already discussed in Callegaro and Poggio (2004). They can be grouped into three different options even if we have no pretention to be exhaustive.

One first option is to model the coverage probability using data from face-to-face surveys with a sample defined from population registers. Obtained parameters may be used in order to adjust the estimates of phone surveys derived from white pages sampling frames. However, this is not likely to be a very effective strategy due to the rapidity of the changes that have been discussed and the typical time needed to release face-to-face survey data. As an example, we can consider that non-coverage in the mentioned ISTAT's Birth Survey 2002 was estimated ex-ante at around 19 % of the population of interest, using data from a 2001 face-to-face survey (ISTAT 2006: 19). Non-coverage was actually twice this figure, without considering further problems deriving from wrong or non-existing telephone numbers (see section 6.2).

A second option is to consider alternative sampling frames. In principle, RDD may be considered for both landline and mobile phones. In practice, a traditional aversion to this sampling strategy remains. Further difficulties will arise from unknown eligibility problems deriving from what has been discussed in section 6.3. In any case, the implementation of the "do not call" register will only result in this effort becoming frustrating.

Finally, a third – and probably more reliable option – is to consider phone surveys within a broader mixed-mode survey strategy, that is something rather new in Italy. Reliable sampling frames exist: voting lists, population registers or geographical units, as in the US experience. They can be used instead of landline phone directories. Respondents may be contacted by post and be offered to participate in the survey in different ways, either through CATI and/or via a web questionnaire, to mention some of the options. Obviously, other methodological problems arise from mixed-mode surveys, and the time frame for their field periods are in any case longer than the one in phone surveys. Non-cooperation will remain an issue to deal with. But coverage error would not represent a big issue any more.

As an example, Bazzoli, Buzzi and Poggio have recently completed a small-scale mixed-mode survey, targeted at young households, on the demand for nursery schools in one single city in Northern Italy[8]. This small survey is also intended as a methodological experiment and seems to provide encouraging results for this survey strategy. The sample was drawn from the local population register. This allowed covering both national and foreign residents. All the households were contacted by post. The ones wishing to participate in the survey were given the possibility to complete a web questionnaire or to call a toll-free number to arrange a phone interview and give their landline or mobile phone number. Furthermore, all contacted households were informed that an interviewer would have contacted them for a face-to-face interview at home, as a default option.

All the ones who did not complete the web questionnaire were contacted for a CATI interview if their phone numbers were available, either because they provided themselves or because of successful linkage between the population register and the public phone directory records. An interviewer was sent to the address of the remaining sample persons for which no telephone number was available.

This methodological experience will be analyzed in more detail, but it is already possible to draw some rough conclusions about survey costs. Assuming CATI interviews as a cost benchmark, the minor costs of interviews collected through the web questionnaire to a large extent compensate for the greater costs for PAPI interviews conducted with less cooperative respondents and for posted materials[9].

This seems to mark the possibility of conciliating the use of a comprehensive sampling frame with the need to save time and money in the data collection

---

[8] http://www.crescereatrento.it.

[9] A mixed-mode approach (CAWI + CATI) is also adopted in the Almalaurea survey, the largest survey on graduates' professional outcomes, in order to control costs of interviews. The sample and contact information are derived from the universities' records. Alumni are contacted by email, whenever possible, and asked to complete a web questionnaire. Telephone interviews are arranged, as a default option, only for those whose email is unknown and for the nonrespondents to emails (Camillo et al. 2009).

process by using CATI or CAWI techniques whenever possible. This can be an effective option to break the low cost–low quality vicious circle without increasing the costs greatly.

Acknowledgements. We are grateful to Giancarlo Gasperoni, Michael Bosnjak, Enzo Loner and the book's editors for their useful comments. Figures presented are partially based on microdata from the ISTAT Multipurpose Survey on Household - Aspects of daily life 2009. Access to the dataset has been granted through the data archive of the Laboratorio di ricerca didattica at the Faculty of Sociology, University of Trento. The analysis presented is under the full and exclusive responsibility of the authors.

References

Billari, F., Rosina, A.: La guerra dei sondaggi (The war of the polls) (2006), `http://www.lavoce.info`

Callegaro, M., Gasperoni, G.: Accuracy of pre-election polls for the 2006 Italian parliamentary election: Too close to call. International Journal of Public Opinion Research 20, 148–170 (2008)

Callegaro, M., Poggio, T.: Espansione della telefonia mobile ed errore di copertura nelle inchieste telefoniche (Expansion of mobile telephony and coverage error in telephone surveys). Polis 18, 477–506 (2004)

Camillo, F., Conti, V., Ghiselli, S.: Integration of different data collection techniques using the propensity score. Paper Presented at the Wapor 62nd Annual Conference, Lausanne (September 11-13, 2009)

Di Gioia, R.: Nel nome del campione rappresentativo: Il Cati e il vitello d'oro (In the name of the representative sample: CATI and the golden calf.). Sociologia e Ricerca Sociale 90, 99–123 (2009)

Esomar (European Society of Opinion and Marketing Research). Global Market Research 2010. ESOMAR, Amsterdam (2010)

Gasperoni, G.: I sondaggi politici in italia. Tra arretratezza e diffidenza (Political polls in Italy. Between backwardness and distrust.). In: Corbetta, P. (ed.) I Sondaggi Politici Nelle Democrazie Contemporanee, pp. 19–48. Il Mulino, Bologna (2007)

Gasperoni, G., Callegaro, M.: "Non cantare vittoria". La capacità predittiva dei sondaggi pre elettorali pubblicati in occasione delle elezioni politiche italiane del 2001 e del 2006. "Do Not Sing Victory". The Predictive Power of Pre-election Polls Published Before the Italian Elections of 2001 and 2006, Polis 21, 463–487 (2007)

Gasperoni, G., Callegaro, M.: Un miglioramento immeritato? Le elezioni politiche del 2008 e la capacità predittiva dei sondaggi preelettorali (An improvement undeserved? The parliamentary elections of 2008 and the predictive ability of pre-election polls). Polis 22, 483–506 (2008)

Groves, R.M.: Survey errors and survey costs. Wiley, New York (1989)

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls II, W.L., Waksberg, J.: Telephone survey methodology. Wiley, New York (1988)

ISTAT (Italian National Institute of Statistics). L'indagine campionaria sulle nascite: Obiettivi metodologia e organizzazione (The sample survey on births: Objectives methodology and organisation). Metodi e norme. Istat, Rome (2006)

ISTAT (Italian National Institute of Statistics). Le imprese di telecomunicazioni (The telecommunications companies). Statistiche in breve. Istat, Rome (2008)

ISTAT (Italian National Institute of Statistics). La vita quotidiana nel 2009 (Daily life in 2009). Informazioni. Istat, Rome (2010)

Marbach, G.: Ricerche per il marketing (Marketing research), 2 edn. Utet, Torino (2010)

Piazza, T., Sniderman, P.M.: Incorporating experiments into computer assisted surveys. In: Couper, M.P., et al. (eds.) Computer Assisted Survey Information Collection, pp. 167–184. Wiley, New York (1998)

Pisati, M.: Sondaggi insondabili. Alcune note metodologiche sui sondaggi politici diffusi dai mezzi di comunicazione di massa (Polls unfathomable. Some methodological notes on political polls broadcast by means of mass communication). Sociologica (2008), http://www.sociologica.mulino.it

Chapter 7
Portugal

Paula Vicente and Elizabeth Reis

7.1 Introduction

The growing mobile phone penetration in developed countries is often accompanied by a downward trend in landline telephone coverage. An increasing number of households give up their landline telephones and rely merely on their mobile phones while others choose not to install a landline phone when moving to a new home. Portugal is no different from this trend.

Mobile phones are invading practically all aspects of modern life and sample surveys are no exception. Either used exclusively or in combination with a landline phone, the mobile phone has potential for survey work.

In this chapter we address the general issues of telephone surveys in Portugal, conducted using either with landline or mobile phones. In Section 7.2 we give an overview of the telephone market in Portugal presenting some figures regarding landline and mobile phone penetration. Section 7.3 presents a brief history of telephone surveys in Portugal. We then discuss a set of issues regarding telephone surveys in light of the current Portuguese market research practices and empirical research on the subject. In the final section we summarize the main features of telephone surveys in Portugal and discuss future trends for this mode of data collection.

7.2 Overview of the Telephone Market

Landline phone coverage achieved its highest rate, 80% of the Portuguese households, in 1997 but it has been in steady decline since 2004 (Statistics Portugal 2001, 2009). At present, 70% have a landline phone but the coverage rate is uneven across subgroups of the population: those with a landline phone tend to be older, more educated and living in bigger households than those without a landline connection at home (Vicente and Reis 2010a).

The telecommunications' market in Portugal was a monopolistic system until 1996 when Portugal Telecom (PT), then the only fixed phone service provider, became a 100% private company. Since then, other companies have entered the market although PT remains leader with a market share of nearly 63%[1] (Anacom 2010a).

Mobile communications service started in Portugal in 1991 at the hands of TMN, a company belonging to the PT Holding[2]. The market is currently shared with two other companies, Vodafone and Optimus, whose market shares are approximately 39% and 15%, respectively[3] (Anacom 2010b). Each operator has its own network and provides nationwide coverage. In 1995 TMN (together with the Italian TIM) launched the first prepaid mobile phone in the world, the Mimo, which contributed to democratizing mobile phone usage in Portugal and worldwide.

Portugal has one of the highest penetration rates of mobile phones in Europe, 148.9 per 100 inhabitants in the first quarter of 2010, making a total of 15.8 million active mobile stations[4] (Anacom 2010b). The ownership of multiple Subscriber Identification Mobile (SIM) cards explains this phenomenon. The percentage of Portuguese mobile phone subscribers with more than one number/SIM card activated is currently at nearly 10% (Anacom 2010b).

At the household level, there has been an extraordinary growth in mobile phone ownership. The percentage of Portuguese households owning a mobile phone reached 87% in 2008 in sharp contrast with 11% in 1997 (Statistics Portugal 2009). Portugal is also one of the EU27 countries with the highest percentage of mobile-only households, 44% in 2008, while the average for that period in the EU27 countries was 29% (EU Statistics 2008).

Mobile phone usage has grown particularly among young people and working people. The fact that the mobile phone is seen as a device that allows professional matters to be resolved at anytime and anywhere as well as the efficient management of personal and family issues has contributed greatly to the continuous growth of mobile phone penetration (Cardoso et al. 2007). Moreover, marketing campaigns by mobile operators are very aggressive and put great pressure on consumers to buy a mobile phone; in fact the range of both mobile phones and tariff plans on offer is so diverse that it is almost impossible not to find a combination of "mobile phone+tariff" that suits individual needs.

Portuguese mobile phone subscribers prefer prepaid plans which represent nearly 72% of all active mobile stations; approximately 15% of these were not used during the last month of the first quarter of 2010 (Anacom 2010b).

7.3 Overview of Telephone Surveys

Market research first began in Portugal approximately 30 years ago. Only one of the "top 5" companies conducting market studies in Portugal is 100% Portuguese,

---

[1] This figure represents the percentage of fixed phone accesses installed by PT.

[2] In fact the mobile phone service started to be offered in Portugal in 1989 by the consortium CTT + TLP. The company TMN resulted from this consortium in 1991.

[3] After 2007 two more networks entered the market, Phone-Ix and Zon Mobile but their market shares remain insignificant (0.4 % and 0.6 %, respectively) (Anacom 2010b).

[4] Active mobile stations are all those that can be used even if they are not in fact used.

Marktest; the other four are part of major international groups, IMS, AC Nielsen, Gfk and Kantar. Although telephone surveys in Portugal are almost as old as market research itself, the way in which they were first conducted was quite different. Telephone penetration was low in the early 1980s and as a result the telephone was only used to survey limited groups of clients or consumers that could be well represented by sampling telephone numbers. There were still no CATI facilities so the interviews were a mix of telephone and paper-and-pencil, the interviewers asked the questions on the telephone and registered the answers on paper.

Despite the uneven penetration of the landline phone across regions and subgroups of the population, telephone surveys grew in importance when the overall coverage rate of the landline phone reached nearly 80% in the mid 1990s (Statistics Portugal 2009) and became one of the major modes of data collection for market, social and opinion research of the general population. The introduction of CATI systems in the early 1990s led to a more efficient process of sample selection, respondents' contacts and interviewing. At present, 95% of market research companies conduct surveys via the telephone covering a wide variety of subjects: awareness and image studies, customer satisfaction studies, branding and segmentation studies, loyalty studies, election polls, and many others that intend to determine people's attitudes, beliefs, values and behaviors. Telephone surveys represent approximately 40% of all quantitative studies[5,6].

All companies conducting telephone surveys use CATI facilities which can be easily adapted for use with mobile phone numbers. However a number of obstacles, both methodological and administrative, hinder the wider implementation of telephone surveys with mobile phones. The next section discusses some of these aspects in light of recent empirical research on the subject.

7.4 Problems of Landline and Mobile Phone Surveys

Current telephone survey methods and practices were developed for landline phones. Using mobile phones to conduct surveys raises challenges for survey companies in terms of coverage, sampling frames, sample selection, cooperation rates and measurement as well as cost. An overview will be given of each of these issues with regard to landline surveys in light of current Portuguese market research practices; the differences arising from the use of mobile phones is also discussed.

[5] The share of telephone surveys is reported in terms of turnover.

[6] These figures relate to APODEMO member companies. APODEMO stands for Portuguese Association of Market and Opinion Research Companies. It was created in 1993 in order to "promote the confidence of market and opinion research studies among clients and population in general" and "contribute to establishing high quality and ethical standards by setting codes of conduct for the sector." (Apodemo 2010). APODEMO company-members represent 80 % of the total survey research market. The Association members follow the professional standards set out in the ICC/ESOMAR code.

7.4.1 Coverage

Landline phone coverage is presently 70% and is expected to continue decreasing. The population without a landline phone tends to be younger, less educated and living in smaller households (Vicente and Reis 2010a). The decrease in landline phone coverage means that samples of landline phone numbers are becoming less representative of the overall Portuguese population and a telephone survey that ignores households without a landline phone is likely to suffer from a serious bias in the estimation of some parameters As a result, landline telephone surveys have been steadily losing their leading position as a survey mode since 2007.

Contrary to the trend in landline phones, mobile phone penetration is increasing and is already higher than that of landlines. Despite the high coverage rate of mobile phones, a survey conducted exclusively with mobile phones will exclude people with distinct characteristics. People without a mobile phone tend to be older and less educated than mobile phone owners; they are more likely to live alone, in rural areas or villages and not to work. (Vicente and Reis 2010b).

In short, neither the landline nor the mobile phone can provide complete coverage of the overall Portuguese population when used exclusively. Using one mode or the other when targeting the overall population is likely to cause bias in survey estimates. However, good coverage can be accomplished in surveys of specific populations.

7.4.2 Sampling Frames

The PT Directory was the sampling frame for landline phone surveys for many years offering many practical advantages: coverage of the entire Portuguese territory, regular updating and availability. PT Directory coverage of the Portuguese with-landline-phone population is no longer at 100%, because since 1996 PT has not been the exclusive provider of landline communication services. However, PT is the only landline communications provider that publishes the numbers allocated along with the names and addresses of subscribers[7]. The sampling frame for residential phone numbers used by each telephone survey company is therefore a combination of the lists of landline phone subscribers and respective annual updates obtained from PT. Unlisted phone subscribers or subscribers of other operators are covered by Random Digit Dialing (RDD). ANACOM[8] provides information on the sets of landline numbers that have been allocated to each communications provider which allows survey companies to reduce the likelihood of generating non-attributed numbers when planning the RDD process.

---

[7] Many people do not want to be listed and ask for confidentiality. The percentage of confidential numbers is currently 20 % compared with just 10 % in 2005.

[8] ANACOM stands for National Authority of Communications; it is the legal entity regulating the communications service in Portugal.

There is no published list of mobile phone numbers[9]. Although many clients allow their mobile phone number and personal data to be published in a Mobile Phone Directory, mobile phone operators choose not to do so to prevent competitors “attacking” their clients with telemarketing aimed at seducing them to change mobile phone operator. Since no sampling frame is available for the general Portuguese population, samples of mobile phone numbers are created by randomly generating numbers. But this solution is not without difficulties. As mobile operators consider their numbering system confidential and provide no information on the attribution of numbers, there is a very high risk of generating many unattributed numbers. A comparison of a landline and a mobile phone survey in Portugal revealed that 59% of the mobile numbers dialed were of no use because they were not-attributed compared with just 26% for the landline phone survey (Vicente et al. 2009). RDD of mobile phone numbers is therefore economically unattractive to survey companies. The population surveyed with mobile phones is currently limited to specific populations such as the clients/employees of a company or subscribers of a specific product or service (e.g. a magazine or newspaper) for which a list of mobile phone numbers is available.

7.4.3 Sample Selection

Landline phone numbers are 9-digit numbers formed by a sequence of region code + area code + suffix. The geographic location of the person being called is known from the region code, while the area code delimits the area into neighborhoods. The geographic control of the sample in landline surveys is easily implemented by using the “region+area” information each landline number contains.

Mobile phone numbers are also 9-digit numbers which have a specific prefix of two digits + suffix. The prefix identifies the network to which the subscriber belongs but gives no information about where the person lives or works or the location at the time of the call. Therefore, mobile phone surveys based on samples of randomly generated numbers do not allow for a geographic control of the respondents and are very difficult to implement for local or regional surveys since a great deal of screening is required to guarantee that only respondents from the target area are surveyed.

7.4.4 Cooperation Rate

Nonresponse in landline and mobile phone surveys might differ either as a result of non-contacts or refusals. The advantage of mobile phones is that the person is always accessible, because the phone is a personal device carried at all times. The mobile phone makes respondents who were previously difficult to reach accessible. Nevertheless, there are situations in which a mobile call may not be answered, e.g. when driving a car or when the car is at a gas station (legislation

[9] Although lists of mobile phones for the three main operators are available for sale (e.g. on the internet), there is no information on how these lists are built and no survey company risks using them in a mobile phone survey intended to be representative of the general Portuguese population.

forbids mobile phone usage in these circumstances) or at the cinema, theater or any public exhibitions (when the silent mode or even turning off the phone is usually recommended). Whereas landline phones are considered a household/family item, a mobile phone is seen as a personal device and a call from strangers on the mobile phone may be considered an invasion of privacy. An unfamiliar number on the phone screen may be refused or the recipient may hang up without answering. However, research by Vicente and Reis (2009) revealed similar refusal rates for mobile and landline phone surveys (10.5% and 11.3%, respectively). This may be due to the fact that the mobile phone interviewers are obliged to begin the interview by asking respondents if it is a convenient time and place for the interview and only proceed if this is confirmed by the respondent (ICC/ESOMAR 2007, 2010). Results for non-contacts though were quite different: 68% of sampled persons in the mobile sample were contact compared with 56% in the landline sample.

7.4.5 Measurement

The interviewing process in a mobile survey is likely to be different from that of a landline survey. While in a landline survey the respondent is almost certainly at home[10], in a mobile survey the respondent can be anywhere. Respondents' attention may be distracted depending where they are and whom they are with or answers may be inhibited. Therefore mobile surveys may encourage satisficing, since people engage in more multitasking when speaking on mobile phones. Research by Vicente et al. (2009) showed some differences between a mobile and a landline survey regarding the percentage of respondents with omissions and the completion time of the interviews. The mobile survey had a higher percentage of respondents with fully completed questionnaires and interviews took longer to complete than landline phone interviews. The fact that the interview with mobile respondents only took place after confirming that it was a convenient moment may have contributed to these findings.

7.4.6 Costs

The charging system for mobile phone services in Portugal adopts the Calling Party Pays principle[11] which means that the caller pays for the entire cost of each telephone call (Littlechild 2006). The respondent in a mobile phone survey incurs no cost when a call is received soliciting participation[12] unlike what happens under the Receiving Party Pays principle[13]. However, for survey companies the adoption

[10] This may not be the case in every situation due to the possibility of diverting calls from the fixed telephone to a mobile phone.
[11] Calling Party Pays is used in most European countries, Australia and New Zealand (OECD 2000).
[12] The exception is when the recipient is abroad; in this case roaming costs are shared by the two parties i.e. the caller and the recipient.
[13] Receiving Party Pays is applied notably in the US and Canada, Hong Kong, Singapore and China (OECD 2000).

of mobile phones for survey research activities is likely to raise data collection costs because calls between two mobile phones or between a landline phone and a mobile phone are typically charged at considerably higher rates than calls between two landline phones. Moreover, calls between different networks are charged at slightly higher rates than within the same network.

7.5 Future Trends

The decreasing landline phone coverage means that there is a growing segment of the general Portuguese population that will systematically be excluded from landline phone surveys. The absence of this segment may bias market surveys because important subgroups of age, education or household size are inadequately represented by landline phone respondents. As this missing segment is expected to grow, its impact in the estimates of landline surveys will increase.

When telephone coverage was still above 80% companies conducting surveys of the general population were able to define the telephone population as the "adult population", as was recommended in the literature (Groves et al. 2001). However, in recent years survey companies have started to be cautious and in most landline surveys the target population is explicitly defined as the "adult population living in households with a landline phone". This indicates survey companies' acceptance of the fact that today the demographic, behavior and attitudinal characteristics of those without a landline phone are likely to be different from those of 15 years ago and they can no longer be represented by the with-landline-phone segment.

The coverage provided by mobile phones is now greater than that of landline phones, 87% of households have a mobile phone while 70% of households have a landline phone (Statistics Portugal 2009). This is surely a stimulus for survey companies to follow the path of "telephone surveys with mobile phones", taking mobile phones either as a supplement or as an alternative to landline phones. However, several aspects should be assessed prior to the decision by survey and marketing research companies to change from landline to mobile mode.

The major challenge for survey companies intending to use interviews on a mobile phone as a survey mode is building a sampling frame. Since mobile phone operators do not provide the necessary information, each survey company has to build its own sampling frame. Two alternative approaches are already being implemented by some companies:

a) Mobile numbers covering the three networks are randomly generated. Then, contacts are screened to identify the status of each number (attributed or non-attributed) and whether they are private numbers or numbers associated with companies/organizations. Finally, the previous list is used as a sampling frame (mother-sample) from which different random samples can be drawn at different times.

b) Mobile phone users are recruited face-to-face, via the Internet or landline phone calls, thus building a "panel" of volunteers. These panels need regular adjustments in order to reflect the general population structure regarding important aspects for market and social research.

Mobile phone surveys are already being implemented successfully when surveying specific populations with virtually complete mobile phone coverage and for whom a list of mobile phone numbers is available. However, the use of mobile phones for surveying the general population is still in its infancy.

Another issue to be taken in consideration when changing surveying modes is cost. The cost of calls between two mobile phones can be three times the cost of calls between landline phones. Survey companies and their clients will therefore use a cost-benefit approach by only adopting mobile phones for surveying large populations when they believe that the greater precision likely to be gained by using mobile phone calls outweighs the increased cost.

Telephone surveys achieved their dominant position when the landline phone service was provided exclusively by PT. Although new operators have entered the market, PT remains the leader with a market share of approximately 63%. In contrast, the mobile phone market is split between three operators that have important positions, 44% for TMN, 39% for Vodafone and 15% for Optimus (Anacom 2010b). The clients of the three networks are not significantly different in demographic terms but the penetration of each operator is different across regions. Additionally, differences have been found between the networks in terms of refusal and non-contact rates (Vicente and Reis, 2011). The pertinence of taking "the network" as a stratification variable in surveys with mobile phones should be monitored.

Finally, the foreseen evolution for telephone surveying using mobile phones cannot be separated from web surveying. In this paper, mobile surveys have been presented as a voice-based mode of data collection in which the interviewer calls the respondents and conducts the interview by asking the questions and registering the responses with the help of a CATI facility. In fact, this is currently the main use of mobile phones in the context of mobile surveys in Portugal. But mobile phones can also be the response device for Computer Self Administered Questionnaire (CASQ) surveys in which the respondents can be sent an SMS containing the link to an on-line questionnaire which they access on their mobile phone and enter the responses using their touchtone devices. Although CASQ surveys with mobile phones are still inadequate for studies on the general Portuguese population as less than 35% of the households have mobile phones with internet connection (Statistics Portugal 2009), they may be sufficient for surveying specific groups with access to this type of technology.

In sum, there is still no significant use of mobile phones as a survey mode to conduct surveys on the general population in Portugal. The non-existence of adequate sampling frames for mobile phone numbers is the major obstacle to this. However, some companies are already trying to overcome this obstacle and preparing their CATI facilities to deal effectively with the selection of mobile phone numbers. Although not yet upon us, a new era for telephone surveys certainly lies ahead.

Acknowledgments. The authors would like to thank APODEMO Director, Dr. António Gomes, for his assistance with this article.

References

Apodemo: Quem somos (about us) (2010), `http://www.apodemo.pt/.qsomos.html` (accessed July 5, 2010)

Anacom: Serviço telefónico fixo – 1º trimestre 2010 (fixed telephone service - 1st quarter 2010) (2010a), `http://www.anacom.pt/render.jsp?contentId=1031502` (accessed July 1, 2010)

Anacom: Serviços móveis - 1º trimestre 2010 (mobile service - 1st quarter 2010) (2010b), `http://www.anacom.pt/render.jsp?contentId=1028809` (accessed July 1, 2010)

Cardoso, G., Gomes, M.C., Espanha, R., Araújo, V.: Portugal móvel - utilização do telemóvel e transformação da vida social, Research Report. Observatório da Comunicação (Mobile Portugal - mobile phone use and transformation of social life, Research Report. Communication Observatory) (2007)

EU Statistics. Eurobarometer 68.2 wave November 2007-January 2008, European Commission, Brussels (2008)

Groves, M., Biemer, P., Lyberg, L., Massey, J., Nicolls, W., Waksberg, J.: Telephone survey methodology. John Wiley and Sons, New York (2001)

ICC/ESOMAR, International code on market and social research (2007), `http://www.esomar.org/index.php/codes-guidelines.html` (acessed July 5, 2010)

ICC/ESOMAR. Conduction survey research via mobile phone (2010), `http://www.esomar.org/index.php/codes-guidelines.html` (accessed September 29, 2010)

Littlechild, S.: Mobile terminations charges: calling party pays versus receiving party pays. Telecommunications Policy 30, 242–277 (2006)

OECD (Organisation for Economic Co-operation and Development). Cellular mobile pricing structures and trends (May 19, 2000), `http://www.oecd.org/dataoecd/54/42/2538118.pdf` (accessed March 5, 2009)

Statistics Portugal, Indicadores sociais de. 1999 Statistics Portugal (2001)

Statistics Portugal, Indicadores sociais de. 2008 Statistics Portugal (2009)

Vicente, P., Reis, E.: Telephone surveys using mobile phones: an analysis of response rates, survey procedures and respondents' characteristics. Australasian Journal of Market and Social Research 17(2), 49–56 (2009)

Vicente, P., Reis, E.: Telephone surveys in Portugal: the coverage error evaluation. In: Paper Presented at the XVIII Congress of the Portuguese Statistics Society, Termas de S. Pedro do Sul, September 29-October 2 (2010a)

Vicente, P., Reis, E.: Who is missing from mobile phone surveys? An analysis of European countries. International Journal of Mobile Marketing 4 (2010b)

Vicente, P., Reis, E.: Mobile phone surveys: a Portuguese case. Working Paper (2011)

Vicente, P., Reis, E., Santos, M.: Using mobile phones for survey research: a comparison with fixed phones. International Journal of Market Research 51(5), 613–633 (2009), doi:10.2501/S1470785309200852

Part II
Sampling for Telephone Surveys

Chapter 8
Who Can Be Contacted by Phone? Lessons from Switzerland

Dominique Joye, Alexandre Pollien, Marlène Sapin, and Michèle Ernst Stähli

8.1 Introduction

At one time, Switzerland was a country where most surveys were conducted by telephone. While it is still the dominant paradigm for many survey agencies, as well as for the Swiss Statistical Office itself, this choice is facing more and more challenges for many reasons. One reason is the high pressure on respondents[1], especially when we consider not only scientific and market surveys but also advertising operations trying to sell something under the label of a "survey"[2]. Another reason is the quality of responses. Many studies have shown that the content of responses differs in accordance with the mode of inquiry: respondents interviewed by phone underestimate alcohol consumption (Hill et al. 1997), exaggerate size, and minimize weight (Paccaud et al. 2001). Conrad and Schober (2000) have highlighted problems related to understanding questions, which is not as critical in the case of face-to-face surveys.

However another reason, the one that will be mainly investigated here, is the difficulty of reaching respondents by telephone. Of course, there is a technical change from landline telephony to other forms (VoIP, mobile phones, etc.), but there are also "human" changes: more and more people no longer want to be listed in a public directory. This is even more often the case with mobile phones, which are very rarely listed. Such an evolution is problematic, not only because many surveys were designed with CATI in mind and not only because the telephone directory is often used as a sampling frame but also because a telephone contact procedure can be a means of reducing non-contact, even for face-to-face surveys.

The coverage of a sampling frame based on a telephone registry is a longstanding question in survey methodology. An evident source of coverage error is households that do not have a telephone or only mobile phones (Vicente

[1] According to the Federation of Survey Agencies in Switzerland, more than four million interviews were conducted in Switzerland in 2008, which means that, on average, each household is interviewed at least once a year, and far more are contacted.

[2] Edith de Leeuw and Joop Hox (2004) published a famous paper, precisely titled "I Am Not Selling Anything: 29 Experiments in Telephone Introductions".

and Reis 2009). Keeter (1995) noted that non-telephone households are not necessarily a static population: subscription to telephone service can be episodic for many households for economic or spatial behavior reasons (Keeter et al. 2000)[3]. In the general literature about survey methodology the question of telephone coverage has been one of the most frequently discussed. It was already an important topic more than twenty years ago in Groves et al. (1988) and Collins and Sykes (1987), and it is still present in Lepkowski et al. (2007).

In this context, why offer a new chapter in this volume? There are at least two reasons: first, the very rapid evolution in the use of telephones and, in particular, the decline of classical connections, encourages us to examine the actual situation. Second, the important variations of telephone use between countries require a number of the studies on this topic. For example, in international comparisons, Switzerland has maintained more complete coverage with landline phones than other countries, despite the fact that mobile phones are also heavily used. In consequence, to examine the possibility of contacting and interviewing people by phone, it is important not only to examine the situation in one reference country but also to have information on local conditions in many different countries, as they can vary locally.

There is one more reason for taking an interest in telephone surveys: more and more attention is now dedicated to mixed-mode surveys, of which telephone surveys are a part. Availability by telephone is, consequently, important for this kind of strategy. This means, also, that, if the telephone survey will no longer be the preferred mode of survey, it will continue to play a role in the survey landscape over the next several years.

More precisely, the research question addressed here aims to determine potential bias caused by the accessibility problem through landline telephones, as well as the issue linked to inclusion in directories. This last point is not only important with regard to a sampling strategy based on registration, where correspondence with a telephone number has to be established, but also for the efficiency of contact procedure: in a pure face-to-face strategy, after five visits at home, it is difficult to have less than 15 % of non-contacts in a country like Switzerland. Even if all interviews are done face to face, a large number of telephone contact attempts is the best cost/benefit solution to lower the number of non-contacts, and doing that implies the ability to match a personal address and a phone number, a link that is possible to establish if a member of the household is registered in directories.

It is also important in this study to mobilize different datasets in the analysis for two reasons: the details of the information they can contain varies, and it is useful to establish the stability of the results in different methodological settings.

The Swiss implementation of high-quality surveys such as the ESS or the EVS allow us to investigate these issues to some extent: the last editions were based on a sampling frame with nearly 100 % coverage, independent of the telephone

[3] Once again, the field of health study seems to have been very active. This is not only the case in the original work of Blumberg et al. (Blumberg et al. 2006, Blumberg et al. 2008), but also of that of Dal and Taylor. (2010), Kempf and Remington (2007), Volker (2009), and Segri et al. (2010), to mention a few examples.

equipment and registration, and they included questions about landline availability and inclusion in the phone directory. Furthermore, they allow the testing of the impact of accessibility by phone, or the lack thereof, with respect to many different domains, including opinions, not just demographic variables. Finally, some elements of the design, such as, for example, a nonrespondent survey, allow us to check for a potential nonresponse bias. In the last section, we will examine the question of whether weighting with easy-to-obtain characteristics appears to be a practical solution to correct coverage or nonresponse bias linked to the use of telephone methodology in some part of the survey process.

8.2 Data

Two main datasets are used for this study: the fourth wave of the ESS and the last edition of the EVS, both conducted in 2008. The latter was completed with a nonrespondent survey. In Switzerland, these two surveys were based on addresses with a face-to-face contact strategy and, consequently, they were not linked to a telephonic contact procedure, except for reducing non-contacts at the end of the field procedure. Even if there is some repetition, let us explain some details of their execution in general and the sampling frame in particular.

The fourth wave of the ESS was conducted between September 2008 and April 2009. A total of 4,970 addresses were drawn from a file comprising all buildings in Switzerland, as well as the private letterboxes in each. Each address was visited. In the case of multiple letterboxes in one building, they were enumerated in a given order, from left to right and top to bottom, and the name on the letterbox, designated according its number, was transmitted back to the survey agency. From the valid addresses with an identified name returned to the survey agency, 3,801 were randomly drawn and used for the survey; 1,819 interviews were conducted, for a response rate of nearly 50 % (49.88 %, according to the ESS rules). This response rate was obtained through a strong investment in the field procedure: five face-to-face tentative contacts, a conversion refusal face to face with specialized interviewers, and, after that, tentative telephone recruitment for the people who had an identified telephone connection or a last letter.

The EVS survey was conducted in Switzerland between March and October 2008, with most of the interviews completed before the summer. The sample was also based on a file established by the Swiss post, containing all buildings as well as the private letterboxes in each. This allowed the calculation of the probability of sampling each letterbox in Switzerland. In contrast with the ESS using a national sample, a random selection of 120 sampling points, with a probability according to size, was made in this case. These sampling points were postal areas, used by the Swiss post for mail distribution[4]. In each of these sampling points, a first selection of 40 letterboxes was drawn, visited by a collaborator, and the names recorded. From there, a final random selection of 23 letterboxes in each sampling point was drawn. In other words, 2,970 addresses were used. A Kish grid was applied in

[4] The sampling points counting fewer than 300 letterboxes were aggregated to form a sufficiently large unit.

each household, and 1,270 valid interviews were conducted, for roughly a 45 % response rate, taking invalid addresses into account. The sequence of contacts, refusal conversion, and diminution of non-contacts was the same as in the ESS.

After this main stage of EVS, a nonrespondent survey was conducted by sending a short questionnaire and a ten Sfr unconditional incentive to all nonrespondents (1,554) as well as to 200 respondents to make a comparison and test for a possible mode effect. From these 1,754 addresses, 847 returned the paper questionnaire, and 74 more responses were obtained by a final CATI contact attempt. Globally, we obtained a 55.7 % response rate in this second stage. However, this also means that, in the end, we obtained information on 2,031 of 2,970 addresses, more than two-thirds of the potential respondents. The questions used were not only demographic variables but also substantive ones, ranging from associative participation to political interest, as well as trust.

We propose first to analyze the ESS as an example of a general social survey, and then to compare some of the results with the main and the nonrespondent surveys of the EVS to see whether the results obtained through the ESS suffer from a general nonresponse bias or are linked to a particular characteristic, such as the topic of the survey, the time of year, etc. From there, we will move to a multivariate perspective to learn more about the cumulative effects of social positions or opinions before testing how far a weighting process can overcome bias linked to the use of the telephone. However, first, let's go back to a more careful examination of a set of hypotheses.

8.3 Hypotheses

For many years, research has been conducted to compute the proportion of people without landline connections. In the U.S.A., the work of Blumberg et al. (2006) is currently a reference and mentions previous studies going back to the eighties. However, in fact, the coverage problem is probably as old as telephone surveys (Groves et al. 1988), and the question of mobile phones has already been discussed across countries for more than ten years (Kuusela 1999).

In most cases, age and education were identified as the most important predictors of availability of a landline, independently of potential registration in directories. Correlations between telephone availability and attitudes are less often studied, even though they are seen as important for the precision of polls, notably in the U.S.A. in the context of electoral studies or electoral prediction (Mokrzycki et al. 2008). The discussion of this point is important because, if age and education are the primary explaining factors, it will be easy to use post-adjustment to correct results, but if landline availability is directly linked to attitudes, a potential adjustment will require another kind of information.

In the case of Switzerland, we expect that the availability of a landline connection is not dependent only on socio-demographic characteristics (less for male young people living alone) but also on opinion and political variables, notably political preferences.

One reason explaining why there are fewer studies about registration in directories is that the rules involved vary far more between countries. In

Switzerland, Swisscom Directories, the organization in charge of maintaining directories, has stated that "various recent reports in the media have commented on the use of the phone book in the internet age. Despite these reports, we have registered only a few terminated subscriptions"[5]. However, this is probably only a partial argument, as the probability is high that new entries in directories are likely less frequent than before, which implies a bias toward the less mobile people and the oldest ones.

We can also expect than the people who are not listed in directories make privacy a high priority and will probably be less likely to answer surveys[6]. If this is true, this is one more indication that one should use not only variables linked to demographics but also variables related to openness to public spaces, participation, and similar indicators.

8.4 Analysis and Results

8.4.1 Landline Telephone Availability: The Lessons of the ESS08

The fourth wave of the ESS, conducted in 2008, contained a standard question about the presence of a fixed-line telephone in the accommodation. In Switzerland (weighted results)[7], 186 persons (10.2 %) mentioned not having this kind of connection[8]. The figures depend, of course, on sex and age, the greatest share of people without landline connections being the youngest males, nearly 20 % of them lacking landline connections. Education is also a factor, with people having vocational training indicating more often than the others that they do not have a landline telephone connection.

It is very important to underline that there is a relation not only to demographic but also attitudinal variables. For example, there is a link with political interest (see Table 8.1).

Table 8.1 Fixed-line according to interest in politics, base ESS08

	Yes	No	Total
Very interested	16.7	11.5	16.1
Quite interested	42.7	33.3	41.8
Hardly interested	31.3	32.8	31.5
Not at all interested	9.3	22.4	0.6

[5] http://www.directories.ch/base.aspx?do=content&pid=nudirfaqdirnav&language=en, Accessed 23 August 2010.

[6] The reverse hypothesis is also mentioned sometimes: as these people are less often contacted for surveys, they are more likely to answer.

[7] When not stated otherwise, all analyses have been conducted using design weights, but without post-adjustment weights.

[8] This is the figure observed in the autumn of 2008. According to informal discussion with people at the Swiss Statistical Office, a noticeable increase of this proportion is very likely.

Of course, the difference in the marginal distribution between people with fixed lines and the total is not large because the former make up the majority of the sample. The variations observed are, nevertheless, a signal that confirms the importance of access mode for the results, even for the simplest indicators.

The differences between people with and without fixed lines is even more impressive when one looks at political preferences, where “classic” parties are, by comparison, over-represented among the respondents with fixed lines, while alternative or populist parties are under-represented (see Table 8.2).

In this case, we also have the “true values” from the results of the last national election. Generally, the survey overestimates left and green parties and underestimates the right-wing, especially the populist rightwing movement (Swiss People Party).

Table 8.2 Political preferences according to landline availability, base ESS08

	Yes	No	Total	Election result 2007 in %
Liberal Radical	17.7	6.8	17.0	17.7
Christian Democrats and Christian-Social	13.1	5.1	12.6	14.9
Socialist and Swiss Labour party	27.0	28.8	27.1	20.2
Greens	13.8	20.3	14.2	12.0
Swiss People Party	23.3	39.0	24.3	30.0
Others	5.2	0.0	4.8	5.2

Source: http://www.politik-stat.ch/nrw2007CH_fr.html, Accessed 05 September 2010.

There are also differences if we look at attitude toward foreigners (see Table 8.3). This is a sensitive topic for the general population, as well as for surveys[9]. In this case, the people without landline connections seem to be more in favor of the foreign population, at least with respect to the question, “Is Switzerland made a worse or a better place to live by people coming to live here from other countries?”

Table 8.3 Immigration and landline connection, base ESS08

Position on immigration	Yes	No	Total
Made Switzerland a worse place	12.4	7.8	12.0
Neither	58.4	50.9	57.7
Made Switzerland a better place	29.1	41.3	30.3

This result, like the previous ones, is statistically significant. Many other significant results can be put forward based on the ESS, but we do not want to include too many tables. Can an identical tendency, showing not only differences

[9] The acceptance in 2009 of an initiative proposing the prohibition of minarets in Switzerland, while surveys conducted previously gave reason to expect widespread refusal of this project, has given space to a large debate on the validity of the surveys, as well as the methodological standards to consider before tackling such subjects.

in demographics but also in opinions, be found in other surveys, such as the EVS, for example? In other words, are these results stable?

8.4.2 Landline Availability and Directory Registration: The Lessons of the Main EVS 2008 Survey

The EVS is another data source available for such an analysis. Even if the response rate is a little bit lower than that of the ESS, we have already mentioned two reasons to mobilize another dataset in this context: testing the stability of the results across surveys and analyzing the effect on participation of accepting to be registered in the directory[10]. Are the respondents who have chosen not to be registered in directories different in values and opinions?

At first glance, the distribution of people without landlines does not differ too much from that in the ESS, even if the number with fixed lines seems to be a little bit lower: 82 % of the sample in the case. However, in addition, nearly 12 % mention not being listed in directories; the combination of these two percentages indicates a real coverage problem.

We can observe the replication of the tendency observed in the ESS for gender and age: more men and young people live nowadays without fixed connections. However, this is also true with regard to registration in directories, ending with nearly 50 % of the category of the youngest men only in a traditional sampling strategy. Education shows the same tendency as in the ESS, but without being statistically significant. Very significant results, well above the classic 0.05 % limit, are, nevertheless, found when considering relation to political behavior in general, beginning with political interest in particular (see Table 8.4).

Even if the answer categories were not the same as those retained for the ESS, each survey tends to preserve its own practices, the general tendency remains the same: a lower interest in politics for those without fixed telephone connections. However, in this case, being listed in directories or not seems to have no influence.

For political preferences (see Table 8.5), again, the difference goes in the same direction as in the previous survey considered: the more conventional parties are

Table 8.4 Interest in politics, landline connection, and directory registration, base EVS08

	Yes	Not listed	No	Total
Very interested	15.0	16.2	10.2	14.3
Somewhat interested	35.7	32.4	22.3	32.9
Not very interested	33.9	34.3	34.4	34.0
Not at all interested	15.5	17.1	33.1	18.9

[10] The accessibility is not only limited by having a fixed-line or not but also by entry in directories. The Swiss Statistical Office can have access to these “private” numbers but not academic surveys. This is also a limitation if we want to use the telephone as a way to maximize the number of recalls while keeping the survey at a more “reasonable” price compared to that associated with multiple face-to-face visits.

over-represented among people with a landline connection, while the leftist parties as well as the right-wing populists are overrepresented in the sample without fixed connections. In other words, there is probably a strong relationship between social integration, conventional political participation, and landline connection.

Table 8.5 Political preference, landline connection, and directory registration, base EVS08

	Landline listed	Landline not listed	No landline	Total	Election result 2007 in%
Liberal Radical	17.8	22.6	9.5	17.6	17.7
Christian Democrats and Christian-Social	15.2	9.4	2.4	13.1	14.9
Socialist and Swiss Labour party	26.5	17.0	40.5	26.7	20.2
Greens	12.9	18.9	11.9	13.6	12.0
Swiss People Party	19.4	22.6	31.0	21.0	30.0
Others	8.1	9.4	4.8	4.8	5.2

It is difficult to see a clear tendency for those not listed in directories. Even if the people situating themselves near the green parties tend not to figure in the directories, being perhaps more careful and less trustful regarding the police and government, this is not a significant relationship.

The analysis can be continued on many variables, showing a regular and statistically significant effect of having access to a landline connection. This is the case not only for the importance given, respectively, to family or work and politics or religion but also for participation in associations or opinions about immigration, as well as trust in others. However, there are nearly no examples where inclusion in directories induces a statistically significant difference.

8.4.3 Overcoming Potential Nonresponse Bias: The EVS, Main, and Non-respondent Survey

The results presented to date already show some important effects of the presence or absence of a landline telephone and, though less dramatically, registration in directories. However, a potential set of criticism can be addressed: even if the response rate is not so bad in a western European context, a value of 50 % means that the remaining half of the population may have different characteristics or opinions.

Adding a nonrespondent survey conducted in the framework of the EVS (EVS-NR) answers this precisely, having a nearly 70 % response rate for the main and nonresponse surveys (EVS-NR). In this context, we have to mention that EVS-NR is a mixed-mode survey, conducted in sequential modes: the main survey was CAPI, the recall in PAPI, and the last tentative in CATI. Even if the questionnaire was carefully designed, taking into account this mixed-mode context, such a design could be a challenge for comparability, and we have limited the choice of questions to those that can be asked without problems through these three different modes. All of the results presented in this subsection are based on a file putting together the answers to the main as well as the nonresponse surveys to minimize potential nonresponse bias[11]. In summary, we will answer two questions: 1) can we find the same types of results if we have a study minimising nonresponse? and 2) can we show a pattern of life conditions or opinions systematically linked to the availability or publicity of telephone connections?

This cumulated dataset shows, in the same line as the main ESS and EVS surveys, that only 75 % of the respondents are available to surveys with a sampling frame based on telephone directories; nearly 15 % were excluded by not having a landline connection, and nearly 10 % were excluded by not being registered in directories. This was the situation in 2008, and the proportion of people excluded is probably higher right now. However, can we find a more systematic explanation for the availability of a landline telephone in a given household, and what is the link with opinions and values?

Age, gender, professional activity, and education: a framework explaining telephone equipment?

Just in considering simple demographics, there are differences between people with or without landline connections, similar to those already shown for the main EVS survey.

- 13.1 % of men but only 9.8 % of women have no landline telephone, but the difference is not so important for inclusion in directories (8.1 % vs. 8.6 %).
- In the mean, education has no influence[12], while activity has some impact; the highest percentage of people without land-line connections is among fulltime active people (21.9 %), while the lowest one is among the part-time people (9.1 %).
- If people are living alone, the difference is great: 24.1 % of this group have no landline, while the rate is less than 9.4 % otherwise.
- Of course, age is another factor, as nearly 30 % of people under 30 years of age have no landline, while this is the case for less than 7 % for people aged over 50.

There is certainly a link between these variables, as they describe not only independent demographic characteristics but also a lifestyle incorporating many of these elements. A simple way to verify these interactions is to use an automatic interaction detection analysis (CHAID) to build the tree of the categories of the presence or absence of a landline connection (see Figure 8.1).

The differences between extreme categories is strong; in one case, people und thirty years old and living alone, more than two-thirds are without landlines, while this proportion is less than 1 % for the people living in households of more than one person, aged of 50 or more, and having had a general education rather than vocational training.

---

[11] In the next analyses, we will continue to use a very simple weighting schema, taking into account the size of the household.

[12] A broad categorization with a distinction between low, vocational, and general education also indicates an absence of relation.

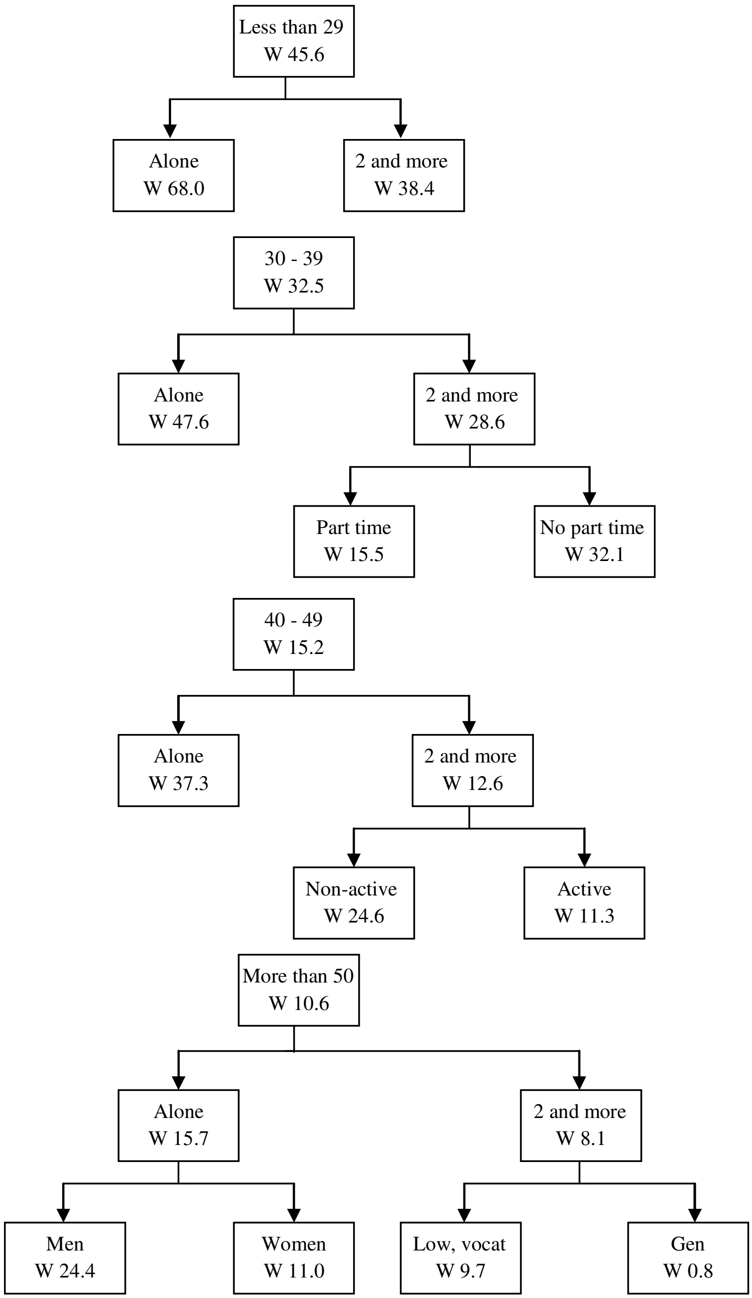


Fig. 8.1 Classification tree (CHAID)[13]

[13]The figures presented here, after the "W", are the proportion without landline connections.

More generally, this first multivariate analysis shows that, if age is the first and most important factor, the fact of living alone or in a multiperson household is also crucial. However, being active in a professional life or the type of education received also has significant effects in some cases. These results could indicate a link between the possession of a landline and a particular lifestyle. Such a relation would be based not only on age but also on social integration and education. If this is true, the hypothesis of a link between opinions and attitudes and the choice of a telephone connection is reinforced. We can try to look at this through another multivariate procedure using a discriminant analysis incorporating opinions and behaviors.

Opinions, behaviors, and telephone connection

We can perform the same kind of analysis on the cumulated EVS and EVS-NR with respect to opinion. The results show pronounced differences that are similar to those already mentioned in the analysis of the main EVS survey. A few of them are as follows:

- The importance given to different topics in life has an impact: the probability of having a landline is higher if religion or family is seen as important.
- The most discriminating variable is political interest: once again, the people without landlines are less interested.
- There are no significant variables among this set that explain registration or non-registration in directories.

All these results illustrate that there are differences not only in the socio-demographic characteristics but also the opinions of people with and without landline connections. That means that pure CATI surveys will probably present more or less biased results for opinion polls.

The significant differences are mainly the same in both the main EVS survey and the file cumulated with the non-respondent survey. This result can lead to, in this case, different interpretations, with the first two being reassuring with respect to the quality of the results:

1. Relative good quality of the face-to-face survey and a weak nonresponse bias in the EVS survey.
2. The absence of a mode effect on the variables chosen for the nonresponse survey[14].
3. Compensation between the mode effect and the nonresponse bias.

Even if the last explanation is theoretically imaginable, such a compensation effect seems very unlikely. This means that we are probably facing a low nonresponse bias in the case of the EVS survey because of its mainly face-to-face way of organizing the field[15].

[14] This is discussed here between the main and the nonrespondent survey. Some other studies, not presented here, have shown that there is not much difference between telephone and mail respondents in the nonrespondent survey.

[15] The only exception is a last tentative to contact, but never interview, the people by phone to minimize non-contacts.

To have the same relation between opinions and telephone equipment when considering nonrespondents is, of course, interesting but not the most important result in the context of such a chapter, dedicated to determining the differences in responses between people with or without landline connections and with or without registration in directories. We already know that there are some correlations between the variables examined so far. One way to draw a complete picture is to use a discriminant analysis, trying to explain membership in the three groups: 1) landline and registered (fix, reg); 2) landline but not registered (fix, none); and 3) without landline (no fix). Two axes are needed, of course, the first one, and the most important, showing the presence or absence of a landline, and the second indicating inclusion in directories. Figure 8.2 shows the position of the mean of these three groups in the discriminant space.

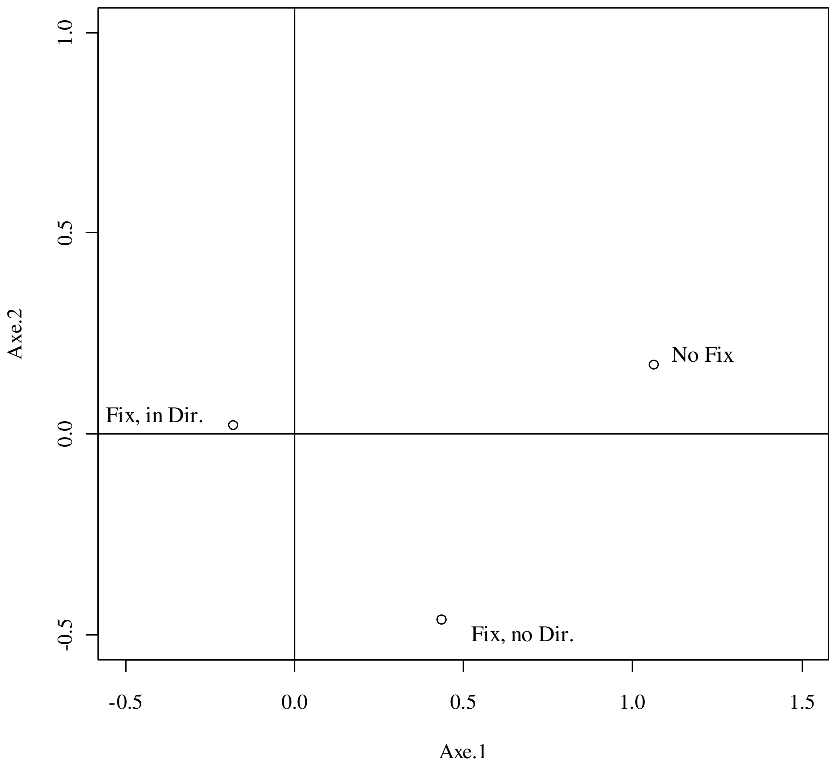


Fig. 8.2 Centres of discriminant classification functions

The interpretation needs not only to consider differences between groups but also to determine which variables are linked to these differentiations. If the discriminant scores are the best linear combination of variables to predict the group in which people are classified, we can try to identify the most important variables by looking at the correlation with the explicit indicators and the

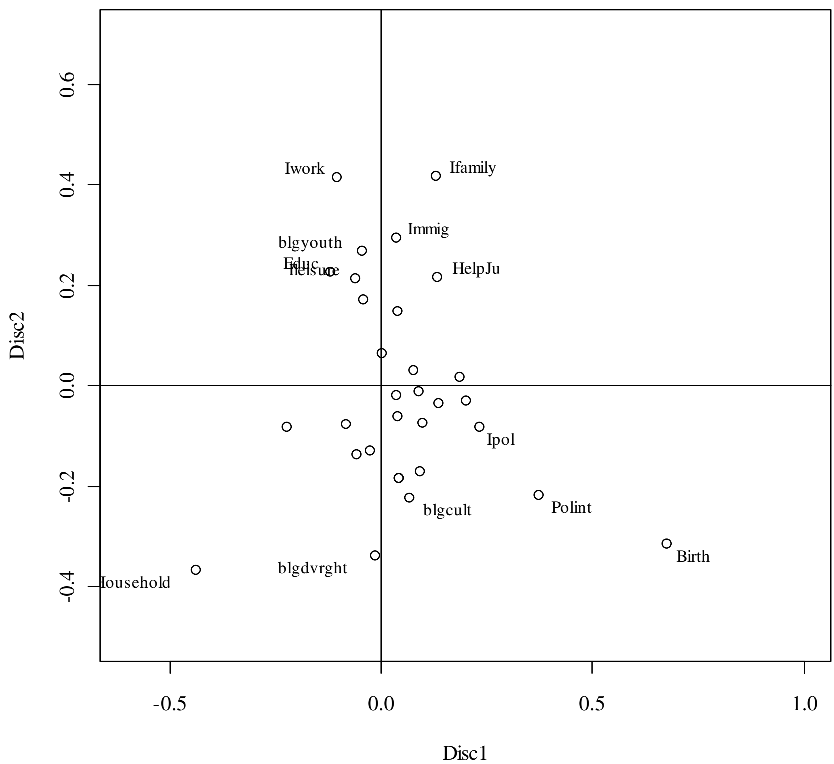


Fig. 8.3 Correlations of explanatory variables with discriminant functions

discriminant scores. The list of the variables used for this and the abbreviations used in Figure 8.3 are given in an appendix.

The picture obtained here (see Figure 8.3) is not much different from what we have obtained so far, where age, type of household, and political opinions, as well as the importance given to family or work, are the main variables. So far, we can see that there are significant differences between people with respect to their opinions. In other words, we can expect that surveys based on landline CATI will tend to overestimate political interest and conventional views about religion or family.

The coefficients presented in Figure 8.3 shows clearly that, if we can easily explain the effect of having or not having a landline, there is no such clear relation with registration in directories. In other words, the former is probably more important than the latter for potential coverage bias, at least in general social surveys such as ESS or EVS. If this result was already mentioned on the basis of the main EVS, it is important to mention that it is still valid when adding the results of the nonrespondent survey to the main dataset.

Figure 8.3 also shows, of course, the importance of the link with demographic variables such as the size of the household, on the left of the graph, or age, to the right of the first axis. However, opinion variables, as interest in politics or the importance given to work or family, are also very important which lower the probability of lacking a landline connection. Activity in voluntary organizations changes also a great deal the type of telephone connection, but the results are diverse depending on the type of organization: no fix is more probable in the case of belonging to a peace movement or welfare-related organization, while fix and registered is more probable in the case of activity in local movements. These results are interesting because they show a variety of behaviors and opinions linked differently to the type of telephone connection. This is one more reason to suppose that weighting only with demographic variables will not be enough to correct the difference in telephone accessibility. However, this will be explored more systematically in the next section.

8.5 Weighting?

The question here is whether weighting the people having only a landline connection[16] according to age and gender[17] is able to help obtain good estimates. Such a strategy is used here because this is the kind of information that surveys based on a telephone strategy will easily obtain. We used a very simple solution of post-stratification according to age and sex and applied it to the main EVS survey, more precisely to the people with landlines only. The idea here is to put ourselves in a situation where the available information is based on respondents having only a landline connection.

Table 8.6 Proportion of people by sex and age in the EVS08 (landline connection and registered people) and the Swiss population in 2008

	EVS08, men, fixed and registered phone	EVS08, women, fixed and registered phone	Men over 18 years old, Swiss resident population	Women over 18 years old, Swiss resident population
18 - 29	8.2	21.3	19.02	17.69
30 - 44	25.1	19.1	28.87	27.12
45 - 64	39.9	42.6	34.31	32.29
65 +	26.7	17.0	17.80	22.90
Total	100	100	100.00	100.00

The basis of comparison is presented in Table 8.6. From there, it is possible to compare the results with and without weighting in the case of political interest and political preference. Even if the difference (see Table 8.7) is relatively small, political interest is a little bit lower when considering the figures weighted according to sex and age in the direction of but still not completely in line with the results obtained when taking into account people without landline connections or directory registration.

[16] Such a strategy is, of course, relevant when considering the very fast expansion of mobile phones and mobile-only households. For a study related to political behavior and the problems of weighting, see Christian et al. 2010.

[17] Household structure is already a weighting criterion in the design weights.

Table 8.7 Interest in politics, influence of post-stratification by age and sex, EVS08

	Landline, design weights	Landline, post-stratification according to age and sex	Unlisted landline, design weights	No landline, design weights	Total
Very interested	15.0	14.3	16.2	10.2	14.3
Somewhat interested	35.7	34.3	32.4	22.3	32.9
Not very interested	33.9	35.3	34.3	34.4	34.0
Not at all interested	15.5	16.1	17.1	33.1	18.9

Table 8.8 Interest in politics, influence of post-stratification by age and sex, EVS08

	Yes, design weights	Yes, post-stratification according to age and sex	Not listed, design weights	No fixed connection, design weights	Total, design weights	Election result 2007 in %	Parliament 2007, % according to seats
PLR	17.8	17.0	22.6	9.5	17.6	17.7	17.5
PDC + Christian Democrats	15.2	14.7	9.4	2.4	13.1	14.9	16.5
PS+PdT	26.5	26.6	17.0	40.5	26.7	20.2	22.0
Greens	12.9	13.8	18.9	11.9	13.6	12.0	10.0
Swiss People Party	19.4	19.0	22.6	31.0	21.0	30.0	31.0
Others	8.1	8.9	9.4	4.8	4.8	5.2	3.0

By looking at the distribution by political family (see Table 8.8), post-stratification doesn't seem to change much, with a variation of less than 1% at most. The most extreme cases, notably the under-representation of the Swiss People Party, as well as the over-representation of the socialist party, are clearly not corrected by post-stratification, according to demographic criteria. This means that, if we want to really correct the results, we must not only consider demographic variables and attitudes such as political interest and social participation but also variables linked more to the process of interest (Voogt and Saris 2003, 2005) and use strategies of adjustment by propensity scores (Lee and Valliant 2009, Schonlau et al. 2004, 2009).

Of course, our data allow for a much more sophisticated post-stratification design than age/sex. However, without going in a direction too long to explore systematically in the context of this chapter, it is already important to show that a simple schema, very often used in the absence of additional information, corrects, perhaps a little bit, the results by comparison with the “true” value but not enough to obtain precise estimates. This is true, at least, for important indicators such as party preference, where a populist right-wing party such as the Swiss People Party is underestimated by surveys based only on people having landlines, being registered, and having connections.

8.6 Conclusion

The results presented so far show clearly that people without landline connections are really different in terms of age, education, and personal situations, measured through household composition. However, the differences are even more pronounced when looking at opinion variables, starting with political interest or party preference.

It is also important to emphasize that these results are stable across different surveys and seem not to be influenced by a non-response bias; a survey with a special module addressed to non-respondents, with a nearly 70 % response rate, shows the same results as a regular general social survey with a small response rate (50 %).

Even if we expect that registration in directories is linked to attitudes such as political opinions or trust in others, this is not really the case, at least not at a significant level. This means that, for telephone surveys, this coverage bias is probably less important here than the one induced by not having a landline at one's disposal. However, even if, at the time, the percentage of people having a landline connection but not being registered in directories is similar to that of those not having landlines, the latter group is growing far more rapidly than the former. This means that the potential bias of telephone surveys based on landlines will continue to gain importance.

The last results to mention are probably more reassuring: the stability of results in the nonrespondent survey, which is a mixed-mode survey, indicates that it is probably possible to use such a strategy to obtain estimates of substantive variables that can later be used to compute propensity scores. However, adjustment strategies need information on substantive variables, not just demographic ones. In other words, CATI surveys can have a place in the landscape of the surveys in a country like Switzerland even if they will probably be less dominant than they were some years ago. However, in any case, the use of a CATI strategy will imply complementing the fieldwork with a specific study of the impact of coverage or nonresponse biases. This will be a huge challenge in conducting high-quality surveys with such a methodology.

References

Blumberg, S.J., Luke, J.V., Cynamon, M.L.: Telephone Coverage and Health Survey Estimates: Evaluating the Need for Concern About Wireless Substitution. American Journal of Public Health 96, 926–931 (2006)

Blumberg, S.J., Luke, J.V.: Coverage Bias in Traditional Telephone Surveys of Low-Income and Young Adults. Public Opinion Quarterly 71, 734–749 (2007)

Blumberg, S.J., Luke, J.V., Cynamon, M.L., Frankel, M.R.: Recent Trends in Household Telephone Coverage in the United States. In: Hoboken, N.J., et al. (eds.) Advances in Telephone Survey Methodology (2008)

Christian, L.M., Keeter, S., Purcell, K., Smith, A.: Assessing the cell phone challenge to survey research in 2010, Chicago (2010)

Collins, M., Sykes, W.: The Problems of Non-Coverage and Unlisted Numbers in Telephone Surveys in Britain. Journal of the Royal Statistical Society A 150, 241–253 (1987)

Conrad, F.G., Schober, M.F.: Clarifying Question Meaning in a Household Telephone Survey. Public Opinion Quarterly 64, 1–28 (2000)

Dal Grande, E., Taylor, A.: Sampling and coverage issues of telephone surveys used for collecting health information in Australia: results from a face-to-face survey from 1999 to 2008. BMC Medical Research Methodology 10, 77 (2010)

De Leeuw, E.D., Hox, J.: I am not selling anything: 29 experiments in telephone introductions. International Journal of Public Opinion Research 16, 464–473 (2004)

Groves, R.M., Biemer, P.P., Lyberg, L., Massey, J.T., Nichols II, W.L., Waksberg, J.: Telephone Survey Methodology. John Wiley & Sons Inc. (1988)

Hill, A., Roberts, J., Ewings, P., Gunnell, D.: Non-response bias in a lifestyle survey. Journal of Public Health Medicine 19, 203–207 (1997)

Keeter, S.: Estimating Telephone Noncoverage Bias With a Telephone Survey. Public Opinion Quarterly 59, 196–217 (1995)

Keeter, S., Miller, C., Kohut, A., Groves, R.M., Presser, S.: Consequences of Reducing Nonresponse in a National Telephone Survey. Public Opinion Quarterly 64, 125–148 (2000)

Kempf, A.M., Remongton, P.L.: New Challenges for Telephone Survey Research in the Twenty-First Century. Annual Review of Public Health 28, 113–126 (2007)

Kuusela, V., Vikki, K.: Change of Telephone Coverage due to Mobile Phones. In: International Conference on Survey Nonresponse, Portland (1999)

Lee, S., Valliant, R.: Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment. Sociological Methods & Research 37, 319–343 (2009)

Lepkowski, J.M., Tucker, C., Brick, J.M., De Leeuw, E.D., Japec, L., Lavrakas, P.J., Link, M.W., Sangster, R.L.: Advances in Telephone Survey Methodology. John Wiley & Sons Inc. (2007)

Mokrzycki, M., Keeter, S., Kennedy, C.: Cell-Phone-Only Voters in the 2008 Exit Poll and Implications for Future Noncoverage Bias. Public Opinion Quarterly 73(5), 845–865 (2008)

Paccaud, F., Wietlisbach, V., Rickenbach, M.: Body mass index: comparing mean values and prevalence rates from telephone and examination surveys. Revue D'épidémiologie Et De Santé Publique 49, 33–40 (2001)

Schonlau, M., Zapert, K., Payne, S.L., Sanstad, K., Marcus, S., Adams, J., Spranca, M., Kan, H.-J., Turner, R., Berry, S.: A Comparison Between Responses From a Propensity-Weighted Web Survey and an Identical RDD Survey. Social Science Computer Review 22, 128–138 (2004)

Schonlau, M., van Soest, A., Kapteyn, A., Couper, M.: Selection Bias in Web Surveys and the Use of Propensity Scores. Sociological Methods & Research 37, 291–318 (2009)

Segri, N.J., Cesar, C., Alves, M., Carandina, L., Barros, M., Goldbaum, M.: Health survey: comparison of interviewees according to ownership of a residential telephone line. Revista de Saúde Pública 44 (2010)

Vicente, P., Reis, E.: The Mobile-only Population in Portugal and Its Impact in a Dual Frame Telephone Survey. Survey Research Methods 3, 105–111 (2009)

Volker, H.: Non-coverage and Health Telephone Survey Estimates: Evaluating the need for a dual frame (2009) (accessed September 08, 2010)

Voogt, R.J.J., Saris, W.E.: To Participate or Not to Participate: The Link Between Survey Participation, Electoral Participation, and Political Interest. Political Analysis 11, 164–179 (2003)

Voogt, R.J.J., Saris, W.E.: Mixed Mode Designs: Finding the Balance Between Nonresponse Bias and Mode Effects. Journal of Official Statistics 23, 367–387 (2005)

Appendix

Abbreviation	Description
Educ	Years of education
Actif	Economic activity
Sex	Sex of respondent
Birth	Year of birth
Polint	How interested are you in politics
Ptrust	Most people try to take advantage of you or try to be fair
Immig	Immigrants living in your country: there are too many
Iwork	How important in your life: work
Ifamily	How important in your life: family
Ifriends	How important in your life: friends and acquaintances
Ileisure	How important in your life: leisure time
Ipol	How important in your life: politics
Irel	How important in your life: religion
HelpJu	Give information to help justice
Private	People should stick to own affairs
Household	Size of household
blgwlfr	Do you belong to: welfare organisation
blgrel	Do you belong to: religious organisation
blgcult	Do you belong to: cultural activities
blgun	Do you belong to: trade unions
blgpol	Do you belong to: political parties/groups
blgdvrght	Do you belong to: development/human rights
blgloccm	Do you belong to: local community action
blgenv	Do you belong to: environment
blgprof	Do you belong to: professional associations
blgyouth	Do you belong to: youth work
blgsprt	Do you belong to: sports/recreation
blgwmn	Do you belong to: womens groups
blgpeace	Do you belong to: peace movement
blghlth	Do you belong to: voluntary health organisations

Chapter 9
Sampling Frames for Telephone Surveys in Europe

Christiane Heckel and Kathrin Wiese

9.1 Introduction

When drawing samples, transnational market research surveys face the key problem of how to apply equivalent standards using the available sampling frames. In contrast to scientific research, which is often allowed to draw samples from the population registers (Häder et al. 2009: 184), commercial research has to rely on sources that are available to the general public. From the point of view of market research, the fundamental considerations concerning "workable and equivalent sampling strategies" need to be extended by the dimensions of their being "accessible and available in the first place".

Against this background, the "Häder-Gabler" model offered an excellent solution for telephone-based market research in Germany, allowing representative national samples to be drawn with a manageable amount of effort and without having to rely only on numbers listed in the phone directories, as had been common practice until then. (von der Heyde 2002: 32ff.) It can be demonstrated that this fundamental idea may also be applied to other European countries.

Samples based on online panels do not present a genuine alternative to representative face-to-face or telephone surveys. Due to the recruitment process and the repeated participation in surveys, in online panels statements about "the general population" are only possible to a limited extent based on the samples drawn according to specific quota criteria. Mobile phone samples drawn from online panels are even less suitable as a means of arriving at generalizable conclusions.

The liberalization of the telecommunication markets along with the high penetration of mobile phones among the European population, have however made it necessary to extend the model.

When constructing sampling frames for telephone surveys, two basic sets always need to be considered nowadays: the set of all landline numbers and the set of all mobile phone numbers. These two sets only partially intersect, and the

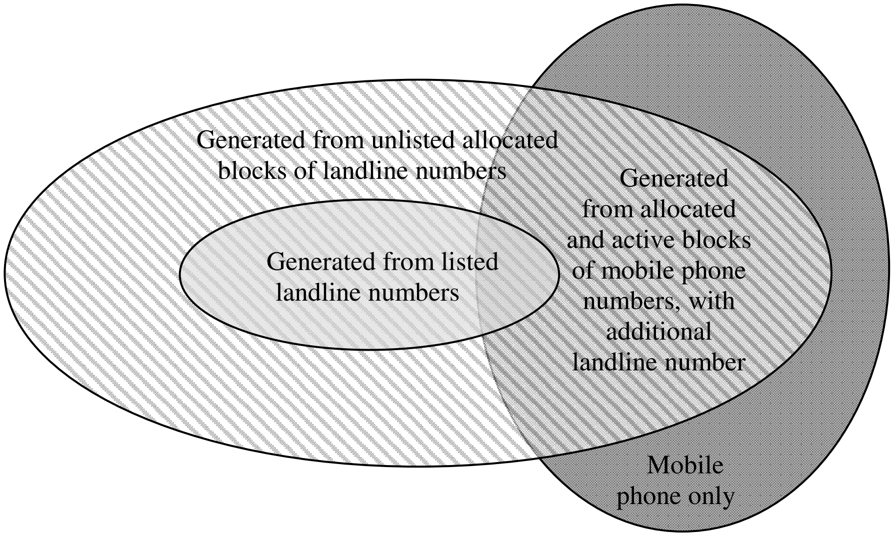


Fig. 9.1 Connection between landline and mobile phone sampling frames

nature of the base used for generating the phone numbers is crucial. Since only a subset of all potential subscribers with landline or mobile phone connections are listed in a public telephone directory in the countries considered here, two types of error can be made if such numbers are confined to data obtained from phone directories:

- In the landline sector, all subscribers who use a landline phone that is not listed in the directory are systematically ignored, and
- The sampling frame does not include people who can only be reached by mobile phone and who are not listed in the phone directory (so-called "mobile-only" subscribers).

The first part of this article will therefore outline the revised method of drawing landline samples, based on the five largest European countries; the second part will then demonstrate the need for "dual frame" samples in order to ensure that "mobile-only" respondents are found in the sample too.

9.2 Constructing European Landline Sampling Frames

9.2.1 Market Penetration in the Landline Sector

The ITU (International Telecommunication Union) reports a negative trend in the number of landline connections per capita in Europe for the year 2009, down 1.7 % for the period between 2003 and 2008, while mobile phone subscriptions grew by 10 % over the same period. Does this mean that we should in future only use mobile phone sampling frames?

Objections are in order here. While mobile phones tend to be personal devices with a separate phone number, a landline phone tends to be associated with an entire household. This means that the per-capita rate for landlines will automatically be lower than that for mobile phones. In addition, a subscriber can be reached at anything from one to five or ten numbers, depending on the provider. The telephone "channels" that are counted by the ITU do not reflect the actual number of phone numbers either.

Neither may we ignore the fact that the share of households that can still be reached via landline phones only (see Table 9.4) is no less than 20 % in Germany, and 11 % in Europe as a whole.

And finally, the regional focus and stratification of samples is only possible in the landline sector, not in the mobile phone sector, since mobile phone dialing codes are not subdivided according to local exchanges.

Important reasons therefore remain for continuing to use landline samples in 2010, though these ought to be combined with mobile phone samples.

9.2.2 Drawing up Landline Sampling Frames

The literature that deals with drawing up CATI sampling frames in the landline sector requires us to use either "a known system for the assignment of phone patterns via activated blocks" or "a complete list of listed phone numbers in which the last digit is replaced by a random digit" (Schnell 2006: 9). However in most European countries, phone books have been used due to the absence of other sources of numbers.

Yet even constructing the sampling frame based only on listed numbers and the number space defined by them, does not go far enough.

Table 9.1 Relationship between listed private phone numbers of private households and number of private phone numbers allocated*

Country	Households in 2009 (million)	Listed private numbers (million)	Allocated private phone numbers (million)	Private households listed (percent)	Allocated phone numbers listed (percent)
DE	39.331	22.658	536.221	57.6	4.2
IT	24.610	17.052	382.013	69.3	4.5
UK	26.753	13.153	257.542	49.2	5.1
FR	27.393	14.622	214.090	53.4	6.8
ES	17.076	10.917	127.758	63.9	8.5

Sources: Eurostat, DE: RegTP, FR: ARCEP, UK:OFCOM, IT: AGCOM, ES: CMT.

* The number of "Allocated private phone numbers (million)" includes all phone numbers used in business, if the telephone provider is classified as "private + business". In the case of Germany, a large number is hold by Deutsche Telekom.

As seen in Table 9.1, restricting the pool of phone numbers to those that are listed in the public phone directory in the five most populous European countries will only produce a fraction (4 to 9 %) of all the phone numbers that have been assigned.

Since the landline connections are at a fixed location, by nature of the technology used (with the exception of hybrid phone numbers which permit both a mobile phone and a landline number to be used within a certain radius of a home), it makes sense to describe features of the sample in terms of household numbers. Thus one in two households still use a listed landline. This does not take into account VOIP numbers, which get by without a location-specific area code. Since samples in the landline sector are normally stratified by region, such phone numbers are missed out as is pure Internet telephony, which uses its own dialing code.

Following the liberalization of landline telephony in Europe, generating sampling frames based only on listed phone numbers (Gabler and Schürle 2002: 59ff.) simply does not go far enough. The private phone numbers assigned by competitors of the former monopolists are only rarely listed (see Table 9.2), meaning that the current model substantially underestimates this subset of phone numbers operated by competitors. In all the countries investigated here, the price of porting a number from one provider to another within the landline network and within a local exchange costs at least 6 € (EU 2010: 49f.). Therefore once the blocks of numbers have been assigned, major shifts cannot be assumed to take place.

Table 9.2 Market shares of the former monopolists and share of listed phone numbers

Country	Name of former monopolist	Market share of former* monopolist	Former monopolist's share of allocated phone numbers**	Former monopolist's share of listed phone numbers**
DE	Deutsche Telekom	51	78	97
IT	Telecom Italia	62	64	99
UK	British Telcom	48	65	94
FR	Orange	56	67	99
ES	Telefonica de Espana	68	42	99

Sources: Eurostat, DE: RegTP, FR: ARCEP, UK: OFCOM, IT: AGCOM, ES: CMT; * national calls by volume 2008 in percentage; ** in percentage.

The less liberalized the landline market, the higher the market share of the former monopolist, and the more numbers are still listed in the public phone directory (Spain and Italy), as seen in Table 9.2. At the same time, the share of listed numbers of private households is higher. This is linked to the publication practice when switching providers. Only when a household switches providers is it asked whether it still wants its number to be listed. This also explains why virtually no numbers allocated by competitors are listed. Only in the UK, where competition has been permitted for the longest amongst all the countries considered here, are significant numbers of entries found for competitors' numbers.

For this reason, the regulatory authorities in charge of telecommunications in France (ARCEP) have started requiring all market participants to report allocated numbers to the authorities. (ARCEP July 2010) However directory enquiries are now only possible online in France, being provided by four agencies which are obliged to obtain their data from the ARCEP. A DVD with the corresponding data is no longer available, making it considerably more difficult to update a sampling frame in France.

The only way of including the phone numbers of competitors in a sampling frame is to expand it by additional blocks of phone numbers, which have been assigned according to the respective national telephone numbering plan. For this reason, the approach proposed by Schnell (2006) and cited at the beginning of this chapter, as an “either-or procedure” cannot be recommended, because important information is lost. If phone books are no longer used, it becomes difficult to assign probabilities to regional phone numbers (this is particularly important in France, since the area codes here only follow a very rough regional pattern). By relying on the phone book only, subscribers who have switched from the former monopolist to a competitor are not adequately reached. Furthermore, difficulties arise with phone systems that use phone numbers of different lengths. The large number of phone numbers allocated in Germany and Italy are also due to the fact that numbers of three or four different lengths are used here, depending on how old those numbers are.

Since the numbers generated on the basis of allocated blocks of numbers without any directory listing are not always actually active blocks, a preliminary technical test of these sectors is necessary.

The procedure described here has already been used in the UK since 2002. The main reason given by Nicolaas and Lynn (2002) for generating phone numbers on the basis of the list of allocated blocks of numbers provided by OFCOM is the marked decrease in the density of listings in the UK. By 1998, ex-directory households are claimed to have accounted for about a third of all households. The authors also describe an uneven regional distribution of ex-directory households, being located particularly in larger cities. Regional stratification was difficult because regional probabilities based on listings and valid combinations of postal codes, area codes and community codes were dispensed with.

9.2.3 Connection between the Sampling Frame and the Ability to Reach Special Target Groups

Apart from the blocks of numbers obtained using entries in the phone directory, the list of allocated blocks of number is also expanded by those for which no entries are recorded in the directory, using the lists published by the respective network agency. In doing so, however, only those providers are taken into account that do not confine themselves solely to commercial customers. In addition, samples per one hundred numbers in a block of a thousand can be used to identify those regions which only return codes for inactive numbers in a technical test. A uniform basis for generating phone numbers can then be created by merging the tested blocks without directory entries and the blocks with entries.

The great importance of expanding the sampling frame to include blocks of numbers with ex-directory numbers will be shown once again using Germany as an example. The media analyses data was kindly supplied to us together with an additional criterion (interview based on a listed or a generated phone number).

Table 9.3 Interviews based on sample sources of landline phone numbers in percent

	ma radio 2010 Wave 1	*ma radio* 2010 Wave 2	*AGOF 2010* Wave 1
Telekom generated	34.0	35.8	34.5
Telekom listed	46.3	43.5	44.4
Competitors generated	18.8	19.8	20.3
Competitors listed	0.9	1.0	0.8
n =	*32,373*	*32,891*	*10,577*

Note: Each wave is weighted, and all the fieldwork was carried out between September 2009 and April 2010.

As the three waves of the survey demonstrate, some 20 % of all interviews are conducted now on phone numbers that are allocated by Telekom competitors, the vast majority of which were not listed in the directory (see Table 9.3). If a sampling frame were constructed solely based on the phone directory, these 20 % of respondents would be missing from it.

Furthermore, the structure of respondents is markedly different from those from the Telekom source. All data were weighted using the official 2008 microcensus.

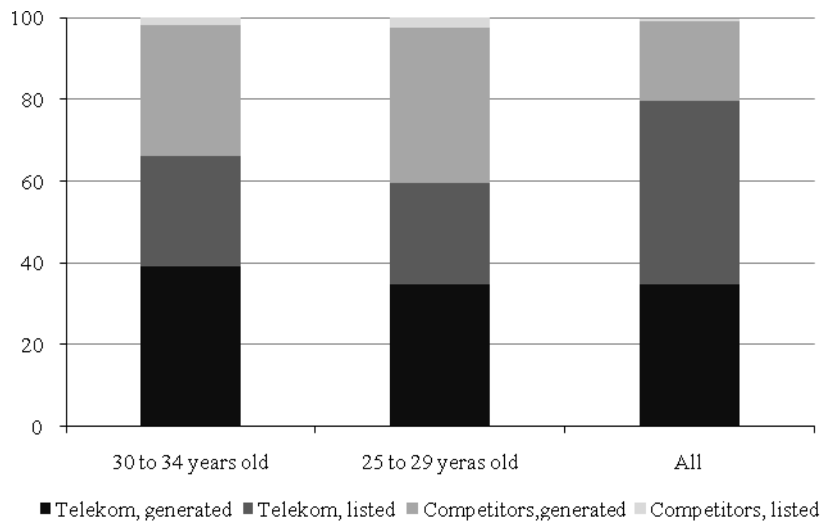


Fig. 9.2 Interviews based on selected age groups and sample sources

Some 38 % of all interviews conducted with people between the ages of 25 and 29 were carried out on the basis of generated phone numbers allocated by competitors. In the overall sample of respondents, they only made up 19 % (see Figure 9.2).

In terms of the nationality of respondents, too, it emerges that 28 % of foreign citizens living in Germany have ex-directory phone numbers allocated by Telekom competitors.

The discrepancy between urban and rural regions already noted for the UK, was also displayed by the German data. 31 % of interviews in the cities of Hamburg, Berlin and Bremen were based on phone numbers that belong to Telekom competitors and are not listed in the phone directory.

In all the countries presented here, the blocks of numbers based on the public phone book only account for one part of the total universe. Whereas the competitors' market share of domestic calls fluctuates at around 50 %, their share of the phone numbers listed in phone directories is only 4 % at most.

However, the empirical results presented here show clearly how important it is to take these blocks of numbers into account when constructing a sampling frame, also in Germany. The composition of the interviews from different sample sources shows that certain sub-populations can scarcely be reached any more only using the phone directory and numbers generated from it.

9.3 Constructing European Sampling Frames for the Mobile Phone Network

9.3.1 Market Penetration of Mobile Phones

A key measure usually used to determine the market penetration in the mobile phone sector is the mobile phone penetration rate. For estimating this rate, the total population of a country is compared with the number of mobile phone subscribers. Mobile phones are assumed to be used by individual persons and in varying locations, in contrast to landline phones which are used in a fixed location and in the context of a household. In Europe, the average mobile phone penetration was 122 % in 2009 (see Table 9.4) – four percentage points higher than the year before. As early as in 2000, the penetration rate for mobile phones in Europe exceeded that for landline telephones (ITU 2009: Yearbook of Statistics 1999 - 2008).

At a mere 90 %, France has by far the lowest penetration rate compared with other European countries, while countries with a similar level of development and population size (e.g. Germany, Italy, the UK and Spain) have a much greater market penetration, with rates between 118 and 146 %. The European Commission attributes this peculiarity primarily to the limited promotion of prepaid offers by French mobile service providers (European Commission 2009).

Table 9.4 Number of mobile subscribers and penetration rate (in percent), October 2009

Country	Number of mobile subscribers	Penetration rate in percentage
DE	108,215,000	132
IT	87,660,560	146
UK	77,797,554	126
FR	58,045,197	90
ES	53,914,549	118
PL	41,201,470	108
RO	24,784,087	115
NL	21,182,000	128
PT	15,535,758	146
EL	14,090,504	125
CZ	14,030,802	134
SE	11,190,300	121
AT	11,150,597	133
BE	11,064,960	103
HU	10,654,339	106
BG	10,567,831	139
CH	8,746,418	112
FI	7,280,000	137
DK	6,944,899	126
SK	5,411,510	100
IE	5,328,813	119
LT	4,933,397	147
LV	2,316,257	102
SI	2,076,751	102
EE	1,556,834	116
CY	1,077,060	136
LU	701,087	142
MT	417,699	101
Average 2009		122
Average 2008		118

Source: European Commission, E-Communications Household Survey – Special Eurobarometer 335, 2010.

At the top end of the scale, countries such as Lithuania, Portugal, Italy or Luxembourg have a particularly high market penetration, with rates of more than 140 %. To some extent these high rates can be explained by the comparatively low landline penetration (usually for historical reasons). Under such circumstances, mobile phones serve as a substitute for a landline phone more often than in other EU countries. High penetration rates are, however, also attributable to a high percentage of prepaid customers, some of whom have several SIM cards ("double SIM effect").

Apart from such effects, which result from the structure of the market, the way in which active subscribers are charged (e.g. deactivation periods that are not taken into account) also affects the mobile phone penetration rate, as does the number of commercially used SIM cards (e.g. "multi-SIM registration" using a single account/contract).

Contrary to what the above numbers might lead one to think, it is not in fact the case that every European owns a mobile phone. According to the E-Communications Household Survey 2010, approx. 13 % of all households in the 27 EU member states still cannot be reached by mobile phone; and in Germany the percentage is as high as 21 %. In the UK 11 % of households, in France 12 % and in Greece and Bulgaria approx. 15 % and 21 %, respectively, have a landline phone but not a mobile phone (see Table 9.1). Most households that only have a landline phone are single-person households whose members are over the age of 60, living in rural settings or small to medium-sized towns (E-Communications Household Survey 2010).

Table 9.5 Telecommunications equipment in Europe in 2010 in percent

Country	*Landline & mobile phone*		*Landline but no mobile phone*		*Mobile phone but no landline*		*Neither*	
	2010	Change from 2006	2010	Change from 2006	2010	Change from 2006	2010	Change from 2006
DE	69	+6	20	-4	11	0	1	-1
IT	62	+2	5	-7	32	+7	1	-2
UK	69	-3	11	-2	20	+7	1	-1
FR	76	+12	12	-9	11	-3	1	0
ES	60	+5	10	-9	29	+6	2	-1
PL	41	-5	11	-12	44	+24	4	-6
RO	33	+0	13	-7	39	+15	15	-8
NL	83	-4	5	-4	11	+7	0	0
PT	46	+7	8	-6	41	+3	5	-4
EL	66	-3	14	-3	19	+7	1	-1
CZ	21	-20	4	-8	73	+31	2	-2
SE	94	+1	5	-2	1	+1	0	0
AT	43	-8	11	-8	45	+17	1	0
BE	62	+6	11	-6	25	+1	2	-1
HU	40	-4	11	-8	44	+14	6	0
BG	48	+12	14	-17	30	+13	7	-9
FI	24	-22	4	-3	71	+24	1	1
DK	62	-11	6	-7	32	+18	0	0
SK	30	-8	7	-10	59	+21	4	-3
IE	66	0	6	-8	28	+10	1	-1
LT	37	+7	7	-6	52	+4	4	-5
LV	42	+1	5	-7	51	+11	2	-5
SI	75	+2	7	-5	18	+5	0	-2
EE	44	-1	8	-6	45	+9	3	-2
CY	77	+2	7	-8	16	+6	0	0
LU	85	+7	6	-7	9	+1	0	0
MT	80	0	15	-2	5	+2	0	0
EU27	*62*	*1*	*11*	*-7*	*25*	*7*	*2*	*-1*

Source: European Commission, E-Communications Household Survey – Special Eurobarometer 335, 2010.

However, the percentage of European households that only have a landline is declining further and further. Since the winter of 2006 (E-Communications Household Survey 2006) their share has dropped by approximately seven percentage points from 18 % to 11%.

9.3.2 Mobile-Only

The percentage of households that can only be reached by landline is dropping; conversely the number of European households that only uses the mobile phone network is steadily increasing. Whereas on average 18 % of households interviewed in Europe could only be reached by mobile phone in 2006, their share has now risen to a quarter of all households (see Table 9.5). At more than 70 %, most households that only have a mobile phone are found in the Czech Republic and Finland, while Sweden and Malta have the smallest proportion of mobile-only households, at 5 % or less. People in the latter two countries are also more likely to have both types of telephone; 94 % of Swedes and 80 % of Maltese can be reached by both landline and mobile phone.

Comparing the 15 EU member states (EU15 before EU expansion in May 2004) and the twelve new members (NMS12), different socio-demographic trends may be noted concerning the exclusive use of mobile phones. In the NMS12, the people who only have a mobile phone most often live in households with several members, in small to medium-sized towns. By contrast, in the EU15 those who can only be reached by mobile phone most often live in single-person households –particularly the under-29-year-olds. In principle, this trend holds throughout Europe, but it is more pronounced in the EU15 member states. In Italy, for example, as many as 70 %, in the UK and Greece more than half, and in Spain 43 % of all single-person households with members under the age of 29 only have a mobile phone (E-Communications Household Survey 2010).

The percentage of German households found in the E-Communications Household Survey 2010 reachable only by mobile phone is 11 %, a relatively low figure compared with other European countries and one that has been stagnating since 2006 (E-Communications Household Survey 2006). Nevertheless, the percentage of mobile-only households is substantially higher than that found in major national surveys, such as media analyse (ma). Apart from the small number of cases (the E-Communications Household Survey is based on 1,522 unweighted cases for Germany) this is undoubtedly also due to the wording of the question, which asks respondents about their devices rather than active phone numbers. Using a face-to-face methodology, the ma2010 Tageszeitungsdatensatz (TZD, a large national readership study that uses strict representative selection procedures and large sample sizes to represent even small target groups as well as the usership of local and target group media with very high quality), whose universe is all private households in Germany at their main place of residence, finds that 6.9 % of all private households can only be reached by mobile phone. The proportion of such households varies from region to region. Thus, in the East German states distinctly more private households, 11.1 %, can be reached only by mobile phone than in the West German states (5.9 %). Furthermore, as seen for the European average, the distribution is found to differ between rural and urban communities. According to ma2010 TZD, central areas of major cities with more than 100,000 inhabitants have distinctly more mobile-only households, at 7.7 %, than are found in the corresponding surrounding communities (5.2 %) or rural communities of less than 5,000 inhabitants (5.5 %).

Furthermore those households in Germany that can only be reached by mobile phone are found to have a lower income. Thus, according to ma2010 TZD more than 33 % of households with a net household income of up to 500 € can only be reached by mobile phone. By contrast, up to about 95 % of private households with higher incomes (above 1,750 € net) have at least one landline phone.

In surveys relying solely on landline samples, experience shows that the "under-coverage error", i.e. the proportion of respondents who cannot be reached by landline, is particularly high in the age group of 20-29-year-olds. For this age group, the percentage of "mobile-onlys" in Germany is already 13 % according to the ma2010 TZD.

In view of the facts outlined here, in most European countries it is becoming increasingly important to use a combination of landline and mobile phone sampling frames in order to achieve as broad a coverage of the population as possible.

Pure mobile phone samples, or the use of mobile phone access panels, are not feasible in Germany, France or the UK, for example, due to the still high proportion of "landline-only" households (see Table 9.5). However using pure landline samples is also becoming more and more problematic in most EU member states, due to the high proportion of mobile-only households. The regionally and demographically inhomogeneous distribution of private households that can only be reached by mobile phone further aggravates the problems associated with using pure landline sampling frames. Certain target groups – in Germany, for example, low-income households with a young principal earner – can no longer be accurately depicted using pure landline samples. These developments make it more and more important to switch to "dual-frame" sampling methods.

9.3.3 Creating the Mobile Phone Sampling Frame

BIK □ Aschpurwis + Behrens GmbH creates mobile phone sampling frames for a range of European countries as part of its BIK □ RSSE (Random Sampling System EUROPE) product family. In order to create adequate mobile phone sampling systems it is crucial – as for landline systems – not to exclude any potentially valid numbers from the sampling frame. In a first step, all theoretically allocated mobile phone numbers are established based on blocks of numbers, for each mobile phone network or dialing code, depending on the system adopted by the national regulatory authorities. This allows the theoretically assigned range of numbers to be generated from random numbers (RDD).

In Germany, with 25 allocated mobile phone dialing codes and 7-8-digit numbers, this leads to a universe of some 340 million numbers, 3.1 times the total number of mobile phone subscribers in Germany. In Switzerland, Spain or France, for example, the allocation of phone numbers by the national regulatory authorities is more detailed; in contrast to Germany, they do not allocate entire dialing codes to a network operator, but instead assign blocks of 10,000 to 1,000,000 phone numbers. Furthermore, a block list is published containing details of the allocation to network operators and in some cases mobile phone service providers. In Switzerland, which has some 20 million allocated phone numbers, 2.3 times as many have been allocated as there are mobile phone subscribers. Spain's 54 million mobile phone subscribers are only overestimated by a factor of 1.5 using the 79.6 million phone numbers allocated in blocks of 10,000. In France, too, there are approximately 1.5 allocated phone numbers for each subscriber.

The fact that the construction of mobile phone sampling frames draws on the list of blocks of phone numbers allocated by the respective national regulatory authorities right from the start, is due to the miniscule proportion of listed mobile phone numbers. In the countries considered here, these account for less than 5 % of all listed numbers.

In the interest of the effectiveness of the sampling frame, care must now be taken not to include number ranges which have not yet been activated. In order to avoid overestimations, only blocks of phone numbers should be generated which are actually used by mobile phone subscribers. For this purpose, a three-step test is conducted on each mobile phone network or range of phone numbers:

In the first step, all blocks of phone numbers are identified from which numbers are listed in the phone directory. In Germany, only about 2.2 million mobile phone numbers attributable to private households are listed in the phone directory (Klicktel January 2010). Hence 98 % of subscribers have unlisted numbers. In Switzerland the density of mobile phone listings (3.7 % of all market participants) is similarly low, with approximately 327,000 listed mobile phone numbers (directories 11/09). The situation is even more extreme in France. Here the coverage is only 0.1 % of all subscribers, with just about 70,000 mobile phone numbers listed in the last published phone book DVD (infobel 2007). Hence the listed numbers are only of limited use in identifying blocks of mobile phone numbers actually used.

In the next step, all blocks of numbers which are not relevant – for research surveys – are identified and excluded from the sampling frame. This, for example, includes blocks which are used by operators for internal technical testing purposes, mailboxes, and blocks reserved for special services. The remaining blocks of phone numbers which are not listed in any directory are then verified by doing manual research with the help of Internet search engines, phone number exchanges and supplier hotlines, to find ranges of numbers that have already been allocated. If this manual search provides no evidence to the contrary, the block of numbers is also switched from "active" to "blocked".

Another possibility, which is however associated with additional costs, is a so-called "look-up" in the Home Location Register (HLR). Various agencies offer to test the activity of a list of mobile phone numbers. However in view of the overall cost associated with creating a sampling frame, we have not yet made use of this option.

Based on such research, the theoretical number of some 340 million allocated numbers in Germany can be reduced to approx. 227 million – 2.1 times the total number of mobile phone subscribers. In Switzerland 17.6 million mobile phone numbers are generated from the used blocks of numbers. Hence only twice as many mobile phone numbers are included in the sampling frame as there are market participants.

In addition, all blocks of mobile phone numbers are coded – if the details are available – to indicate which provider that range of numbers is assigned to and what type of service it corresponds to (prepaid, monthly payment, home zone/home option etc.). This allows these blocks to be boosted or filtered out of a sample if necessary.

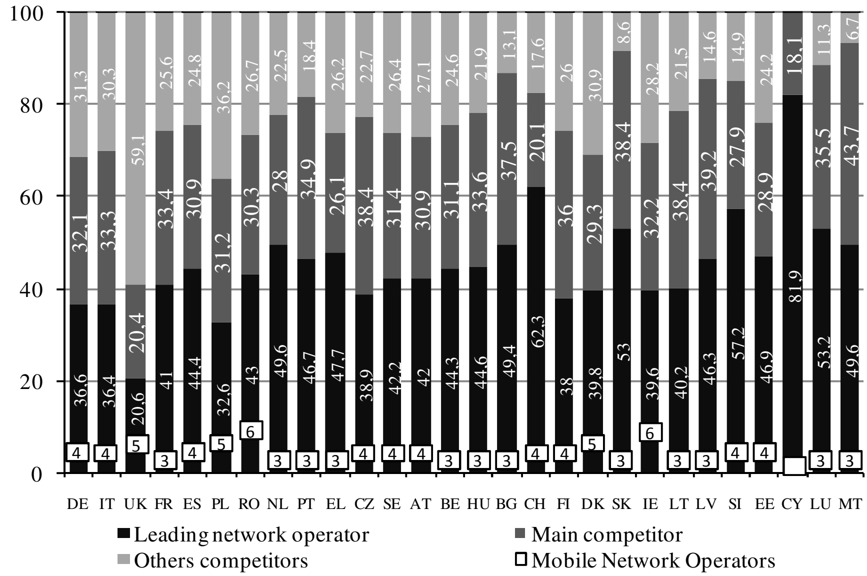
However it is important to watch out for number porting in mobile phone networks (Mobile Number Portability, MNP)), i.e. the option of keeping one's entire phone number, including the dialing code, when switching providers. When numbers are ported it is no longer possible to clearly identify the network they operate in from the dialing code. As of 2010 porting mobile phone numbers is possible in all EU member countries: in the Netherlands and the UK, it has been possible to take one's mobile phone number when switching to a different provider ever since 1999. In Finland, almost 5 million mobile phone numbers have already been ported since 2003, corresponding to 67 % of market participants in 2009 (see Table 9.4). In Denmark too, at 41 %, in Spain at 34 % or in Sweden at 32 %, the proportion of subscribers with ported mobile phone numbers was very high in 2009. By contrast, in Germany the mobile phone network can be ascertained with a fairly high accuracy from the dialing code or the allocated block of numbers, since only 2.6 % of numbers have been ported, and the situation is similar in France, at 4.7 % and Austria, at 5.7 % (E-Communications Household Survey 2010).

9.3.4 Sample Structure

Throughout Europe, mobile phone numbers display no regional characteristics, which means that stratification by administrative divisions (e.g. NUTS) or community size is not possible – as it is in most representative landline surveys. No regional samples can be drawn, nor can individual regions be boosted in the sample, since this characteristic first needs to be established in the course of the interview.

The patterns of competition in the mobile phone market are not as varied as those in landline telecommunications. As a rule, there are three or four mobile network operators (MNOs), of which Vodafone (28 %), Telefónica/O2 (19 %), T-Mobile (17 %) and Orange (14 %) are the most common in Europe (European Commission 2009). The mobile phone industry, therefore, is highly concentrated. Almost 80 % of Europeans (about 460 million mobile phone customers) make calls through these operators. However, since 2005 the market shares of the leading operators and their main competitors have been dropping slightly, which the European Commission attributes to an increase in Mobile Virtual Network Operators (MVNOs).

On a national level the market share of the leading operator is almost always less than 50 % and the most important competitor has on average a market share of about 30 %. The market share of the leading operator is lowest in the UK (21 %) and Poland (33 %), but in Germany and Italy, too, strong competition means that the leading network operator has very little of a lead. In Cyprus, Switzerland and Slovenia, on the other hand, the traditional operators hold a considerable share of the market. For example Swisscom has a share of 62.3 % in Switzerland, which is remarkably high compared with the European average of 37.8 % (see Figure 9.3).



Source: European Commission, 15th Progress Report on the Single European Electronic Communications Market (2009); for CH: 15th European Union implementation report extended to include Switzerland (2009)

Fig. 9.3 Mobile market share based on subscribers, October 2009

Studies investigating the subscriber structure in Germany and Switzerland, for instance, show that mobile phone operators within individual countries have very different subscriber structures and regional distributions.

In Germany, for example, whereas T-Mobile is used more by people in West Germany, Vodafone's main emphasis lies in East Germany. In rural communities, the German population is more likely to use T-Mobile and Vodafone, in cities on the other hand E-Plus and O2 have large market shares as well. Mobile phone users under the age of 40 are more likely to use the E-Plus or O2 network than the older generation, most of whom have their most recent mobile phone contract with T-Mobile. Differences between network operators are also noticeable in terms of household incomes. For example the population with a net household income above 3,000 € is much more likely to use T-Mobile than subscribers with a lower income (VerbraucherAnalyse 2010: Klassik I Märkte).

In Switzerland, although Swisscom is the market leader in almost all sectors, with a market share of 62.3 %, clear differences can be indentified in the use of mobile networks depending on demographic characteristics such as age and linguistic region. Thus the Sunrise network is used especially by Italian-speaking Swiss citizens between the ages of 14 and 39, whereas young French-speaking citizens are found more often within the network of the French company Orange. As a successor to the former state-owned PTT (Post, Telegraph, Telephone), Swisscom AG has a much higher share among older mobile phone subscribers and in rural communities – much like T-Mobile in Germany (MACH Consumer 02/2009).

Based on these different demographic market structures, as seen in the example of Germany and Switzerland, it is important to take all network providers into appropriate consideration when putting together a mobile phone sample. In order not to risk ignoring certain customer segments, proportional random sampling must be performed based on network operators, i.e. the market shares of the network operators must be used as a stratification criterion.

9.4 Conclusions

The sampling frames and parameters presented here for Europe's "Top 5" show how important it has become to combine both sampling frames. In order for a sample to reach all the different groups within a population, it is increasingly not enough to use only a single sampling frame. If only 70 % of the population can be reached by landline phones, as in Italy or Spain, the sample must be expanded using the mobile phone networks. However, this happens at the expense of regional stratification, since the region to which the respondent is to be assigned can only be determined correctly once the interviewer has made contact with him or her.

When weighing up whether it is "sufficient" to work only with a landline sampling frame based on phone directory entries, or whether to expand the landline sampling frame as described here and to include the mobile phone network, the matter of cost plays an important role in commercial market research. The effort involved in removing from the sample numbers that do not lead to an actual contact, and the cost of phoning people in the mobile phone network who may live in a region outside the desired population, should not be underrated. On the other hand, particularly the younger target group, highly mobile people and people with below-average household incomes can increasingly no longer be reached via the above-mentioned landline samples. The costs and benefits must be weighed up with the client for every CATI survey.

References

Arbeitsgemeinschaft media-Analyse e.V., Media-Micro-Census GmbH, ma 2010 Tageszeitungsdatensatz, Universe: German-speaking population above the age of 14 years in private households at main place of residence

European Commission. eCommunications household survey: The results of a special Eurobarometer survey - Spezial-Eurobarometer 249 (July 2006), http://ec.europa.eu/public_opinion/archives/eb_special_260_240_en.htm

European Commission. 15th Progress Report on the Single European Electronic Communications Market, Staff Working document (2009), http://ec.europa.eu/information_society/policy/ecomm/library/communications_reports/annualreports/15th/index_en.htm

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Progress Report on The Single European Electronic Communications Market (15th Report), (COM(2010) 253) (2010)

European Commission. eCommunications household survey: The results of a special Eurobarometer survey - Spezial-Eurobarometer 335 (October 2010), http://ec.europa.eu/public_opinion/archives/eb_special_339_320_en.htm

Gabler, S., Schürle, J.: Zur Stabilität des Gabler-Häder-Auswahlrahmens (On the stability of the Gabler-Häder selection frame). In: Telefonstichproben, Methodische Innovationen und Anwendungen in Deutschland, pp. 59–67. Waxmann, Münster (2002)

Häder, S., Ganninger, M., Gabler, S.: Die Stichprobenziehung für den European Social Survey: Prinzipien und Ergebnisse (The sampling for the European Social Survey: Principles and results). In: Weichbold, M., et al. (eds.) Umfrageforschung 2009, pp. 181–193 (2009)

International Telecommunication Union. Yearbook of statistics, Telecommunication/ICT Indicators, Chronological Time Series 1999-2008. International Telecommunication Unioun, Geneva (2009)

International Telecommunication Union. In: Information Society Statistical Profiles 2009, Europe, v1.0. International Telecommunication Union, Geneva (2009)

MACH Consumer. Analysis by DemoSCOPE - WEMF AG (2009), MACH Consumer 2009-2 - Potentialanalyse, Universe: Residential population of Switzerland above the age of 14 in private households, adequately assimilated to the official language(s) in the district of residence, mobile phone ownership (February 2009)

Nicolaas, G., Lynn, P.: Random-digit dialling in the UK: viability revisited. Journal of Royal Statistical Society 165(2), 297–316 (2002)

Schnell, R.: Die Umstellung des LFS (Labour Force Survey) auf RDD-CATI (Random Digit Dialing – Computer Assisted Telephone Interviewing) (The conversion of the LFS to CATI RDD). Économie et statistiques – Working papers du Statec 14 (2006)

VerbraucherAnalyse, Klassik I Märkte, Axel Springer AG, Bauer Media Group, VerbraucherAnalyse, Klassik I Märkte - Strukturanalyse, Universe: German-speaking population above the age of 14 in private households at main place of residence, private mobile phone ownership or private and business mobile phone usage (2010), http://online.mds-mediaplanung.de/vakm/1c9224ca8c08f24ff91eb5e9e9547eda/client#start

Von der Heyde, C.: Das ADM-Telefonstichproben-Modell (Model of the ADM-Sample). In: Gabler, S., Häder, S. (eds.) Telefonstichproben, Methodische Innovationen und Anwendungen in Deutschland, pp. 59–67. Waxmann, Münster (2002)

Appendix

Country: Regulatory authorities and telephone directories, including details of allocation of landline numbers according to providers

Germany: Bundesnetzagentur, Verzeichnis der zugeteilten Festnetznummern, März 2010, KLICKTEL Jan 2010

France: ARCEP (Autorité de Régulation des Communications électroniques et des Postes) Le plan de Numérotation Observatoire semestriel de l'annuaire universel en France, Situation au 30 juin 2010 Les annuaires téléphoniques pour la France, http://www.lesannuaires.com/annuaire-telephonique.html INFOBEL FRANCE OFFICE 2008

United Kingdom: OFCOM (The Office of Communications), The national numbering plan, Q1 2010 UK White Pages 2010

Italy: AGCOM (Autorità per le garanzie nelle communicazioni) Piano di numerazione nel settore delle telecomunicazioni e disciplina attuativa Allegato A alla delibera n. 26/08/CIR ISCTI, Instituto Superiore delle Communicazioni e Technologie dell'Informazione – Ufficio 5, Reporto 4 Data Family Italia 2009

Spain: CMT – Comisión del Mercado de las Telecommunicaciones, Bases de Datos de Numeración 2010 INFOBEL Espana Office, Directario Telefónico 2009

Chapter 10
Mobile- and Landline-Onlys in Dual-Frame-Approaches: Effects on Sample Quality

Götz Schneiderat and Tino Schlinzig

10.1 Introduction

Mass media, politics, and commercial enterprises are increasingly using social survey data to substantiate their reasoning and interventions. Although the number of telephone surveys is decreasing more than one third of all studies of market and social research in Germany is conducted on the telephone (ADM 2011). This raises questions about the quality of the gathered data. One of the central variables to look at is the quality of the realised sample. Reflecting the number of sample units interviewed as a percentage of the number of eligible sample units, response rates are considered reliable indicators of data quality. However, irrespective of the mode declining response rates question this assumption (Schnell 1997, de Heer 1999, Däubler 2002, de Leeuw and de Heer 2002, Curtin et al. 2005, Stoop 2005). In addition to the number of ineligible cases (including people not part of the target population or unable to participate) the quality of surveys of the general population suffers from nonrespondents. A growing number of contacted persons refuses participation in market and scientific surveys. Numerous studies have examined different aspects of this phenomenon in detail and offered solutions to the problem (Blasius and Reuband 1995, Hüfken 2000, de Leeuw and Hox 2004, Meier et al. 2005, de Leeuw et al. 2007, Schnauber and Daschmann 2008). It is highly probable that the nonrespondents in the sample differ from the respondents in ways that are associated with the objectives of the survey. This is one good reason for looking at the realised sample to measure survey data quality. Another one concerns the question of which people can be surveyed by which mode. Coverage errors result if some members of a population under study do not have a known nonzero chance of being included in the sample. The latter turns out to be highly relevant especially for telephone surveys. In the past, the potential for coverage errors has been comparatively low in RDD landline telephone surveys in Germany. At the end of the 1990`s almost every household had landline telephones. But from the early 2000`s onward mobile phones have become increasingly popular and widespread and have started to replace traditional landline phones. Persons as well as households have temporarily or permanently switched to mobile phones as an exclusive means of telecommunication. Today this group of people so-called 'mobile-onlys' cannot be ignored anymore by survey research. In Germany their numbers have increased from about 5 % in 2006 to approximately 10 %[1] at the moment. In consequence of the declining number of people with landline phones the problem of under-coverage of certain populations has become virulent. Consequently, surveys conducted via landline telephones only are not an appropriate means for estimates of a general population under study (Ehlen and Ehlen 2007, Carley-Baxter et al. 2010, Peytchev et al. 2010, Busse and Fuchs 2011, Peytchev et al. 2011). On the other hand, about 20 % of the German population does not own mobile phones. Together with Bulgaria and Romania Germany leads the list of countries within Europe. Hence, exclusive mobile phone surveys would cause significant bias in certain measures as well. Data from our study CELLA 2 (Cell Phone and Landline) suggest that conclusions about the German population can still only be drawn from surveys applying a dual frame approach. Moreover, the data confirm that the quality of general population surveys conducted by telephone should not only be assessed by response rates but rather by looking at the realised sample and thereby at the respondents included.

Based on a discussion on the importance of response rates as a quality indicator for telephone surveys this chapter examines advantages and disadvantages of a dual frame approach with respect to possible coverage biases. Differences in demographic characteristics and personal traits across respondents of certain subpopulations using cell phones and/or landline telephones will be considered and conclusions will be drawn[2].

Our data stem from CELLA 2 'General Social Telephone Surveys via Mobile Phone', a project funded by the German Research Foundation (DFG, HA 1898/3-2) . 1,507 landline and 1,500 mobile phone interviews were conducted by the forsa institute[3] in Germany. Both RDD samples – landline and mobile – were drawn by GESIS – Leibniz Institute for the Social Sciences, Mannheim/Germany. Field time started on July 10th 2010 and was finished on August 30th 2011.

---

[1] While in Germany the number of mobile-onlys rose only very little in the last years, a rapid development can be observed in the replacement of landline telephones by mobile phones in the European Union. Frontrunners are Czech Republic and Finland where more than 70 % of the population is only reachable on the mobile phone (Europäische Kommision 2010: 58). Similar observations have been made in the United States. By 2005 the prevalence of mobile-only households had increased to approximately 8 - 10 % and the proportion of U.S. adults that exclusively use mobile phones to 8 - 9 % (Ehlen and Ehlen 2007, Blumberg et al. 2008). Numbers for the EU however should be considered with care. In particular for Germany it is well-known that a large number of persons that according to self-reports only have a mobile phone can be reached via so-called virtual landline numbers. However, in contrast to the general notions market and social research define 'mobile-onlys' exclusively by the criterion of availability on any possible landline number. According to Eurobarometer data (Europäische Kommision 2010: 57ff.) two trends can be observed in Europe. On the one hand, there are a number of countries in which dual availability (landline and mobile phone) is predominant. On the other hand, there are countries in which only one mode dominates.

[2] The authors wish to thank Matthias Lehmann and Stefan Fehser for their assistance and insightful comments.

[3] FORSA. Gesellschaft für Sozialforschung und statistische Analysen GmbH (http://www.forsa.com/, accessed June 30, 2011)

10.2 Response Rates

Schneekloth and Leven (2003: 18) claim that the scientific discussion pays only little attention to the quality of survey data in general. Moreover, they argue that there are no practical concepts defining quality requirements (criteria) and measures that constantly evaluate data quality on an empirical basis (monitoring). This argument partially also holds for the discussion on response rates.

However, despite the fact that there are many ways of calculating response rates, they are still supposed to be a resilient indicator for survey quality (Groves 2004, Stoop 2005). Response rates are considered a criterion for a good application of the population under study.

In a review of the research literature Porst (1996) compared statements on minimum response rates and concluded that the stated recommendations are rather arbitrary instead of systematically acquired. Finally, there is neither a theoretical nor an empirical way of defining what constitute a sufficiently high response rate.

Both laypeople and researchers consider high response rates as a quality indicator, since they at least reduce the risk of bias (Groves and Peytcheva 2008, Aust and Schröder 2009). Whether nonresponse at a given response rate induces bias or not is an additional concern. Previously lacking nonresponse coding standards within survey research the American Association for Public Opinion Research (AAPOR) first published their standard definitions for final dispositions of case codes and outcome rates for surveys in 1998. Unfortunately, no comparable standards are available in Germany. This makes a comparison of survey outcomes across different surveys more difficult. Particularly, the way in which households that were not reached are treated (as non-contacts or with unknown eligibility) can be interpreted differently and thus affects the published response rates (Fuchs 2010: 236). In addition, the basis for comparison is missing if detailed documentation of nonresponding cases are missing. AAPOR standard definitions provide a precise catalogue of disposition codes differentiated by mode and help categorizing different types of response and nonresponse of sampled cases. Beginning with the 5th edition published in 2008 these definitions contain important modifications for mobile phone surveys as well as an improved classification of contact attempts with cases of unknown eligibility.

Additionally, AAPOR provides a „Response rate Calculator“ to calculate four different types of response rates from the final disposition codes entered. Response rate 1 (RR1) represents the minimum response rate, which would apply under the improbable case that all cases with unknown eligibility can be classified as ineligible. In contrast, response rate 3 (RR3) – a conservative estimate – takes the level of insecurity about the cases of unknown eligibility “e“ into account[4]. It

4 ‘e’ is the estimated proportion of cases of unknown eligibility that are eligible. Enter a different value or accept the estimate in this line as a default. This estimate is based on the proportion of eligible units among all units in the sample for which a definitive determination of status was obtained (a conservative estimate). This will be used if you do not enter a different estimate. For guidance about how to compute other estimates of ‘e’, see AAPOR's (2009) Eligibility Estimates.
http://www.aapor.org/Standard_Definitions/3049.htm, accessed November 27, 2011

estimates what proportion of cases of unknown eligibility is actually eligible. Response rates 2 (RR2) and 4 (RR4) are considered equivalent to RR1 and RR3, respectively, but in addition include partial interviews as respondents. Since incomplete interviews are unusable for our purpose, our analysis exclusively uses complete interviews. For this reason Table 10.1 only presents response rates 1 and 3. Empty fields were omitted for better comprehension.

Table 10.1 CELLA 2 Response Rates

	landline	mobile phone	total
Interview (Category 1)			
Complete	1,500	1,507	3,007
Partial	41	90	131
Refusal	5,100	4,538	9,638
Break off/ implicit refusal (internet surveys)	121	197	318
Telephone answering device (confirming HH)	215	1,155	1,370
Miscellaneous	1,196	2,144	3,340
Always busy	2,388	506	2,894
No answer	2,992	3,186	6,178
Answering machine - don't know if household	447	8,167	8,614
Call blocking	1	54	55
Out of sample - other strata than originally coded	1,858	1,117	2,975
Fax/data line	1,059	208	1,267
Non-working number	13,004	21,053	34,057
Disconnected number	43	14	57
Temporarily out of service	393	3,029	3,422
Special technological circumstances	0	1	1
Number changed	2	0	2
Total phone numbers used	*30,360*	*46,966*	*77,326*
I = Complete Interviews (1.1)	1,500	1,507	3,007
P = Partial Interviews (1.2)	41	90	131
R = Refusal and break off (2.1)	5,221	4,735	9,956
NC = Non contact (2.2)	215	1,155	1,370
O = Other (2.0. 2.3)	1,196	2,144	3,340
UH = Unknown household (3.1)	5,828	11,913	17,741
Calculating e	0.333	0.275	0.299
Response rate 1			
I / (I + P) + (R + NC + O) + (UH + UO)	0.107	0.070	0.085
Response rate 3			
I / ((I + P) + (R + NC + O) + e (UH + UO))	0.148	0.117	0.130
Refusal rate 1			
R / ((I + P) + (R + NC + O) + (UH + UO))	0.373	0.220	0.280
Refusal rate 2			
R / ((I + P) + (R + NC + O) + e (UH + UO))	0.516	0.367	0.431

Higher response rates can be gained by allocating cases of known eligibility and including not completed cases or partial interviews as respondents as suggested by AAPOR standard definitions RR4 (AAPOR 2011). Independently, several studies reported declining response rates (Schnell 1997, de Heer 1999, Däubler 2002, Curtin et al. 2005, Holbrook et al. 2008, Lepkowski et al. 2008).

Using AAPOR definitions of RR3 for our study the response rate is 13 %. The RR3 is 14.8 % for the landline sample and 11.7 % for the cell sample. The difference of 3.1 percentage points can be explained primarily by the fact that for mobile phones a high portion of contact attempts resulted in contact with mailboxes, which were handled as cases of unknown eligibility.[5]

15.8 % of all contacted mobile phone numbers were identified as mailboxes with preset announcements on behalf of mobile service providers. Instances where personalized mailboxes were reached (two percent of all cases) the sampled phone number is likely to be a SIM card in use.

Response rates are comparatively low, however, are close to the rates of telephone surveys today (Diekmann 2010). As is described in the following, the study still succeeds in estimating socio-demographic variables of the target population under study.

In comparison to an earlier study CELLA 1[6] which was adapted to AAPOR standards following the suggestion of Meier (2009) response rates of CELLA 2 were lower. An earlier publication "Telefonbefragungen über das Mobilfunknetz" (Häder and Häder 2009) describes this study, in which a RR3 of 21.9 % was reached, in detail. The difference of 8.9 percentage points between CELLA 1 and 2 can be explained by such things as the use of a different methodology. In CELLA 1 approximately 70 % of the sampled mobile units received a prior notice via text message announcing the upcoming interview (Schlinzig and Schneiderat 2009). Positive effects of warm contacts were also discussed in the context of several other studies (Dillman et al. 1976, Dillman 1978, Goldstein and Jennings 2002). We assume that compared to CELLA 1 response rates are clearly lower due to the lack of pre-announcements in our current study. Response rates dropped by approximately 14.1 percentage points. However, for practical considerations the method was not applied to CELLA 2. Nevertheless, the CELLA 2 response rate for the mobile sample is 9.2 percentage points lower than the response rate in CELLA 1 without prior notice. It could be argued that the higher response rates in CELLA 1 can be explained by the fact that mobile phone surveys were less

---

[5] In this context a new procedure seems promising. The Home Location Register (HLR) Lookup verifies the status of generated mobile phone numbers without calling. By means of data inquiries a commercial provider (e.g. http://mobimex.com/, accessed May 03, 2011) examines whether and how numbers are used. Struminskaya et al. (2011) showed that this procedure can reduce the amount of cases with unclear status in mobile surveys and moreover provides means of calculating response rates more accurately and consequently allows higher response rates. However, further research is needed to examine the reliability of this procedure.

[6] CELLA 1 was an earlier study financed by the German Research Foundation (DFG) and conducted by the Department for Social Empirical Research at Dresden University of Technology in cooperation with GESIS – Leibniz Institute for the Social Sciences, Mannheim from 2006 to 2008. 1,009 landline and 1,162 mobile phone interviews were conducted in our CATI-studio at Dresden University of Technology. Field time started on November 2007 and was finished on April 2008 (Häder and Häder 2009). Together with two pretest studies about 3,500 telephone interviews were realized both via landline and mobile phones.

widespread at the time, whereas traditional landline surveys had achieved certain saturation on part of the respondents. Obviously, saturation continued to progress in both modes even if legal regulations in Germany started to prohibit unwanted telephone advertisements by direct marketing calls. Further reasons for lower response rates in CELLA 2 might be due to the shorter field period of seven weeks compare to seven month in CELLA 1. To what extent the university background has contributed to the increase – calls were conducted by student interviewers in the department's own CATI lab in CELLA 1 – can be only speculated, since in CELLA 2 no university was mentioned by the Forsa interviewers. This would be in line with the findings of Meier et al. (2005). He argues that mentioning a university leads to higher response rates. Finally, also the response rate of the landline sample declines with 5.7 percentage points compared to CELLA 1.

This result is also reflected in the refusal rates. Refusal rates are computed similar to response rates with or without the estimated proportion of cases of unknown eligibility that are eligible. With 51.6 % the refusal rate is 3.3 times higher than the response rate in CELLA 2. The mobile sample did 3.14 to 3.54 times better than the traditional landline sample. With respect to nonrespondents we distinguished between physically or mentally unable and respondent refusal. However, these categories were not particularly selective. Less than one percent mentioned "too ill" and "too old" as the reason for their refusal.

However, in contrast to the response and refusal rates, the rate of non-contacted sample units is higher for the mobile sample; this is mainly due to the cases of unknown eligibility. Among these cases of unknown eligibility might also be cases where the sample unit did not answer the phone, because the phone number of the call centre was transmitted and thus displayed on the phone. This is more likely to have been the case for mobile phones than for landline phones.

10.3 Comparisons with Data from the German Microcensus

A further indicator whether the achieved sample reflects the population sampled from a correspondence between sample characteristics and population characteristics know from other data sources. In Germany for example, the data from the Microcensus, which are official and representative data of the German population collected by Statistics Germany, are often used for data comparisons. These data are collected annually and, due to legally mandatory participation, response rates reach 95 %. The data are collected by interviewers who visit the households and gather data by means of computer-assisted face-to-face interviews (CAPI). As such, the Microcensus data serve well as reference data for sociological studies. Those characteristics that are collected in both the Microcensus and the survey data can be compared to each other. However, to make this possible the Microcensus questions have to be adopted in the survey as well. Standard demographic questions, which form the basis for comparable data, exist for self-completion and telephone surveys in Germany. However, the comparison is limited to certain variables such as age, sex, education, family status, and income.

A number of studies examine survey results with the aggregate Microcensus statistics and discuss response rates and socio-demographic characteristics. Hartmann and Schimpl-Neimanns (1992: 11ff.) compared German General Social Survey[7] data (ALLBUS 1986, 1988, 1990) with Microcensus data (1985, 1987, 1989) and proved differences across datasets in the proportion of small households and in the proportion of persons of low socio-economic status (middle class bias). The authors showed that by person-representative weighting the bias of small household size – in particular of respondents aged 70 and older – was reduced. By means of multivariate analyses (Hartmann and Schimpl-Neimanns 1992: 20ff.) they further proved that the middle class bias was mainly due to an education bias, which is according to expectations, since nonresponse is often found associated with education and employment status.

Koch (1998: 81) also presents results showing a tendency in this direction and proving that response rates do not affect data quality. Similarly, Schneekloth and Leven (2003) argue on the basis of "best practice" ALLBUS data from 1994, 1996, 2000 that high response rates are not realizable by tough documentation of survey processes.

The comparison with Microcensus data can be considered as a relatively simple method for detecting nonresponse bias. However, this comparison makes two assumptions. First, it assumes that the data from official statistics are correct and unbiased. Second, it assumes that all population units had an equal non-zero inclusion probability, i.e. that the bias found in the comparison of the survey data with the Microcensus data was solely due to nonresponse bias. Some survey methodologists, such as Groves (1989: 201) question these assumptions.

In Table 10.2 below the socio-demographic variables of our mobile and landline sample are compared with the ALLBUS and Microcensus data. The data are design-weighted. The design weight adjusts the unequal selection probabilities of households of the landline sample and compensates for differential selection probabilities by the number of telephone numbers. This also affects the effective sample size.

Table 10.2 Sex

	CELLA 2			*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
Male	45.5	55.8	50.4	49.4	48.8
Female	54.5	44.2	49.6	50.6	51.2
(n)	*(1,580)*	*(1,427)*	*(3,007)*	*(3,469)*	*(69,501)*

Design-weighted data.

[7] The German General Social Survey (ALLBUS) is a representative cross-sectional survey using face-to-face interviews. It collects data on attitudes, behavior, and social structure in Germany. Moreover, ALLBUS allows international comparative analysis. It includes questions also asked in the American General Social Survey (GSS). The sample size is about n = 3,500. For more information please consult the GESIS website: http://www.gesis.org/en/allbus/general-information/, accessed April 12, 2011.

Table 10.2 shows that the mixing of modes in our study had a corrective function, particularly regarding the distribution of sex. Men are underrepresented in the landline sample but overrepresented in the cell phone sample, and the opposite is of course true for women. These differences disappear when total sampled population of CELLA 2 is compared to the reference statistics.

Table 10.3 Age Groups

		CELLA 2		*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
16-19	4.7	8.6	6.5	2.7*	5.2
20-29	13.0	25.0	18.7	12.5	14.2
30-39	14.7	18.8	16.7	14.2	14.5
40-49	21.1	20.0	20.6	20.4	19.8
50-59	19.1	16.8	18.0	17.2	16.3
60-69	13.3	7.3	10.5	16.2	13.6
70+	14.0	3.5	9.0	16.7	16.3
(n)	*(1,556)*	*(1,409)*	*(2,966)*	*(3,457)*	*(69,501)*

Design-weighted data, * ALLBUS interviews people aged 18 and older. Therefore, the age group of 16 - 19 year-olds is not comparable.

Differences across subsample and reference data sets can also be found with respect to age groups as Table 10.3 depicts. Separated according to mode and subsamples it is remarkable that landline data quality is superior in particular when considering younger respondents (up to 39 years) and persons in the older age groups (60 years and older). Middle age groups (40 to 59 years) are better represented by our mobile sample. However it should also be noted that in general younger persons are slightly overrepresented and the oldest age group (70 +) is underrepresented.

Table 10.4 Nationality

		CELLA 2		*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
German	95.4	94.3	94.0	95.5	91.0
Not German, EU citizen	2.2	2.5	2.3	1.6	3.3
Citizens of Non-EU countries	2.4	3.2	2.8	2.9	5.7
(n)	*(1,573)*	*(1,407)*	*(2,980)*	*(3,400)*	*(69,501)*

Design-weighted data.

Conducting telephone surveys via mobile phones favors traditionally underrepresented subpopulations such as migrants. Our mobile phone sample mirrors the official statistics when looking at eligible Non-European respondents, as Table 10.4 shows.

Table 10.5 Marital Status

	CELLA 2			*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
Single	28.0	46.8	37.1	23.8	29.7
Married living together	53.9	37.7	46.4	58.1	52.8
Married but separated	1.8	2.7	2.2	1.6	2.0
Widowed	7.0	2.8	5.0	8.3	8.0
Divorced	8.9	9.3	9.1	8.3	7.5
(n)	*(1,580)*	*(1,427)*	*(2,989)*	*(3,464)*	*(69,501)*

Design-weighted data.

As presented in Table 10.5, implementing a mobile sample decreases data quality with regards to the distribution of marital status. While our landline sample provides an optimal estimation of the distribution of different marital status groups, the mobile phone data bias the overall representativeness. Only the group "married but separated" is adequately represented. Traditional landline surveys observed that this group is difficult to contact (among others Blasius and Reuband 1995: 76). We assume that surveys conducted exclusively by mobile phone would only insufficiently estimate marital status.

Table 10.6 Education

	CELLA 2			*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
Lower secondary level	21.9	19.0	20.5	36.3	38.8
Secondary level	25.5	27.8	26.6	31.4*	21.8
Secondary level former Democratic Republic of Germany	5.2	7.2	6.2	--	6.7
Highschool degree allowing access to polytechnical universities	7.2	8.2	7.7	6.1	6.0
A-level	29.3	27.8	28.6	22.2	19.7
Other	10.9	10.0	10.4	4.1	7.0
(n)	*(1,580)*	*(1,427)*	*(2,988)*	*(3,460)*	*(69,501)*

Design-weighted data, * ALLBUS does not distinguish between secondary levels of education in the former Democratic Republic of Germany and in current German school system (POS and Realschule).

Telephone surveys typically suffer from an education bias. As described above, predominantly higher educated persons participate in surveys. In our CELLA 2 data (Table 10.6) we find a similar bias; higher educated respondents are severely overrepresented while people of lower education are severely underrepresented. Implementing a mixed-mode survey obviously only corrects for educational biases to a very limited extent. More respondents with lower secondary education were successfully interviewed in the landline sample and fewer respondents with a high school degree were included.

Table 10.7 Households by Household Size

	CELLA 2			*ALLBUS 2008*	*Microcensus 2009*
	landline	mobile phone	total		
Number of Persons					
1	20.3	22.0	21.1	20.7	22.3
2	37.0	31.6	34.4	39.8	38.2
3	17.9	20.4	19.1	17.4	18.0
4	16.2	14.7	15.5	14.7	15.1
5 +	8.7	11.3	9.9	7.4	6.5
(n)	*(1,570)*	*(1,409)*	*(2,979)*	*(3,451)*	*(69,501)*

Design-weighted data.

Table 10.7 presents data on household size. It is worth noting that single households are better represented by our mobile sample than by the traditional landline sample. The landline sample results are more promising with respect to households with more than one person. However, altogether CELLA 2 data are quite close to the reference statistics with respect to household size.

Table 10.8 Comparison of income groups with Microcensus data from 2009

	CELLA 2			*Microcensus 2009*
	landline	mobile phone	total	
< 500 €	1.2	0.8	1.1	1.5
500 € - 1500 €	12.6	17.3	14.8	24.4
1500 € - 3500 €	55.3	55.1	55.2	49.0
3500 € - 5000 €	20.8	16.9	19.0	11.4
> 5000 €	10.1	9.9	10.0	9.8
(n)	*(912)*	*(805)*	*(1,716)*	*(63,614)*

Design-weighted data.

A comparison of the weighted CELLA 2 data with Microcensus data from 2009 regarding grouped household net incomes results in the following picture: the lowest income groups (up to 500 €) are well estimated by our CELLA 2 data with altogether (1.1 % in CELLA compared to 1.5 % in the Microcensus). However, respondents with low income (500 € to 1,500 €) are clearly underrepresented with 14.8 % compared to 24.4 % in the Microcensus data. However, as Table 10.8 shows the two subsamples deliver different data. Again our mobile sample has a corrective effect in the direction of the reference statistic. Nevertheless and independent of the survey mode middle incomes (1,500 € to < 3,500 €) are slightly overrepresented in CELLA 2. With respect to the higher income groups the mobile sample performed somewhat better; with 16.9 % the mobile sample provided a better estimate than the landline sample (20.8 %). Very high incomes are well represented by both subsamples.

In summary, it can be stated that despite low response rates, comparisons of the CELLA 2 data to official reference statistics showed that our samples performed in representing the characteristics of individual subgroups of the survey's target population. The integration of a mobile sample by applying a dual frame approach nearly always leads to better sample quality. However, conducting a mobile phone survey only would cause biases with respect to the age distribution in favor of younger respondents. Regarding the education bias and biased estimates of the oldest age group the dual frame approach did not improve the sample quality. However, these biases are only partly due to low response rates, since even when response rates were still high (such as in the 1980s when ALLBUS still achieved 65 % response) such biases were already present. Thus, the low response rates achieved in CELLA must be associated with a new and different social process, which explains a general reluctance in the population to participate in survey. However, our interim results illustrate that the additional use of a mobile phone sample can improve the contact rate of the targeted population but not the willingness of sample units to participate. Persons that are more likely to refuse in other survey modes are also difficult to motivate to participate in surveys via mobile phones.

10.4 Landline- and Mobile-Onlys

So far we have presented findings for the question of which respondents are more likely to be interviewed by which calling mode. However, thus far we have not distinguished between whether the respective respondent was more likely to be a respondent to one mode or the other or whether they would not have been sampled in one mode and, consequently, need to be considered as mobile- or landline-only. Methodologically this question is concerned with two different types of error: unit nonresponse and under-coverage. Flexible and mobile persons who also have a private landline can be sampled into the mobile and landline samples even though they might need more contact attempts on landline phone than on the mobile phone. The situation is different for respondents who only own telephone mode.

In the following we present findings indicating which group can be reached by which mode and describe the characteristics of the group members.

Table 10.9 Socio-Demographic Characteristics of Mobile- and Landline-Onlys

	Landline-Onlys	both	Mobile-Onlys (+ virtual landline)
Sex	more likely female	balanced	more likely male
Age	more likely 40+	mostly 20 - 60	more likely 20 - 49
Average Age	57	42	37
Education			
Low	38 %	18 %	23 %
Medium	24 %	32 %	38 %
High	30 %	45 %	32 %
Other	8 %	5 %	7 %
Average Income	1,472 €	1,784 €	1,353 €
Marital Status	• more likely married living together • more likely widowed	• more likely married • more likely single	• more likely single
Partnership			
Without partner	74 %	46 %	54 %
Living apart	6 %	26 %	23 %
Politics	• more likely Green Party • CDU° and SPD°° balanced	• more likely SPD than CDU • FDP and Green Party balanced	• more likely SPD than CDU • FDP and Green Party balanced
*Migration Background**	21 %	14 %	16 %
n = 2,221	305	1,672	244

Design-weighted data, * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$, ° Conservatives °° Labor.

Apart from the fact that mobile-onlys are more likely to be male, single[8] and aged 30 - 39 (see Table 10.9), these respondents also have the lowest average income and a medium level of education. Although mobile-onlys are primarily single, they are also more likely to be living in a partnership compared to landline-onlys. The landline-only in contrast has the highest average age, is likely to be divorced and lives alone. These people predominantly have lower education which is likely to be an age effect.

Regarding incomes it is remarkable that mobile-onlys have the lower average income. At the beginning of the mobile phone development and distribution as consumer product, mobile phones represented a status symbol. The notion of mobile calls being more expensive compared to landline calls originates from this time. But by various tariff amendments (including discount offers) expenditure patterns for telephone communications have shifted. It is highly plausible that persons with low income can only afford prepaid contracts, since they result in no monthly fees.

[8] Single is to be defined as partner-less living arrangement by choice.

Prior to our study we assumed that persons with a migration background would be more likely to belong to the group of mobile-onlys. The opposite is the case. Migrants are more frequently landline-onlys. This particularly applies to the higher educated, married and older migrants aged 50 and older in West-Germany. Due to low response rates our data can only be generalized with a certain degree of reservation, since we have no insights into whether responses are evenly distributed across migrant groups. In addition, migrants are likely to be heterogeneous with respect to socio demographic variables and responses.

We computed a logistic regression to specify the differences between mobile and landline-onlys. Table 10.10 examines mobile-onlys compared to landline-onlys, while Table 10.11 shows mobile-onlys and the intermediate type (landline and mobile phone available). In addition to the socio-demographic characteristics of the respondents we included 'Big Five' personality variables in the analyses. The term 'Big Five' describes a psychological model which claims to capture the main personality characteristics based on five valid and reliable dimensions. These characteristics are considered temporally stable and culturally independent. The dimensions are neuroticism, extroversion, openness, agreeableness, and conscientiousness.

In CELLA 2 personality characteristics were operationalized with three items per factor. Often complex phenomena are conceptualized with many more questions. Examples include NEO-FFI and NEO-PI-R, which contain 60 and 240 items per dimension (Ostendorf and Angleitner 2001, 2004, Muck 2006). A set of publications considers the measurement of personal traits based on different versions of the 'Big Five' (Lang et al. 2001, Rammstedt et al. 2004, Rammstedt and John 2005, 2007, Rammstedt 2007, Häder and Häder 2011). The German Socio-Economic Panel Study (SOEP) for example uses 15 items (Gerlitz and Schupp 2005, TNS Infratest Social Research 2010: 21) similar to our CELLA 2 study. First analyses suggest that such, reduced personality measures' can be just as successful[9]. A Varimax rotation showed that the included items are distributed across the five factors the way it was expected. The internal reliability of the respective factors varies with a Cronbach's α between 0.52 and 0.64. With rising number of items per dimension Cronbach's α would increase accordingly, however at expense of the reliability (Field 2005). Results are satisfactory considering that in our study only three items were used for each index. The 'Big Five' variables were generated by computing average values for each dimension.

[9] Häder and Häder (2011) show that differences between CELLA and SOEP data are negligible.

Table 10.10 Factors Influencing Type of Telephone (Mobile-Only without Landline Number)

	Telephone User Type (R = Landline-Only)
Age	0.94***
Sex (R = male)	0.42**
East-West-Germany (R = west)	2.58**
Marital Status/ Living Arrangement, (R = living together)	3.11***
Big-5 „neuroticism”	0.68*
Big-5 „extroversion“	1.42**
Big-5 „openness“	0.68**
Big-5 „conscientiousness“	1.45*
Migration Background (R = „yes“)	0.47*
Household Net Income	0.99**
Pseudo – R²	0.41
n	271

Logistic Regression (Odds Ratios) of Factors Influencing Telephone Type (“Mobile-Only” without landline number), * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

The used regression model shows good variance explanation with a pseudo-R^2 of 41%. Table 10.10 illustrates that persons using only mobile phones are predominantly younger, male and do not live together with their spouses or partners. Furthermore, there are significant influences of certain personality characteristics (Big Five). Mobile-onlys are less “neurotic” (fearful and nervous) and less open to new experiences (rather pragmatic[10] and conservative). Furthermore, they are rather extroverted (sociable and active) and conscientious (reliable and careful). In addition, a migration background is significantly less probable for mobile-onlys compared to landline-onlys. The household income shows only a very small effect. However, mobile-onlys tend to have a lower household net income available than landline-onlys.

Table 10.11 Telephone Type Influencing Factors (Mobile-Only without Landline Number)

	Telephone User Type (R = Intermediate user)
Age	0.99*
Sex (R = female)	1.69***
East-West-Germany (R = West)	0.60***
Marital Status/ Living Arrangement ,(R = living together)	0.38***
Big-5 „agreeableness“	0.79*
Big-5 „neuroticism“	0.76***
Big-5 „conscientiousness“	1.28*
Pseudo – R²	0.08
n	2,694

Logistic Regression (Odds Ratios) of Factors Influencing Telephone User Type (“Mobile-Only” without Landline Number), * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

[10] For mobile-onlys this result appears to be counterintuitive and instead seems to describe the age group of the mobile-onlys. The fact that young people are rather pragmatic also corresponds to the findings of the Shell Jugendstudie 2010 (Shell Youth Study). According to data of this study personal success is crucial for young people in a performance and consumer society. For a summary see http://www.shell.de/home/content/deu/aboutshell/our_commitment/shell_youth_study/2010/values/, accessed June 30, 2011.

When considering mobile-onlys and comparing them to those using both mobile phones and landline telephones, some of the influences disappear. ‚Age' is no longer significant: mobile-onlys are only slightly younger than the so-called intermediate telephone users. However, the effects of both sex and partnership and of the 'Big Five' dimensions "neuroticism" and "conscientiousness" remain significant (see Table 10.11).

The personality trait of "agreeableness" (understanding and helpful to others) is significant in this model; Mobile-only users are significantly more agreeable than the intermediate user type. However, the explained variance of the model with the reference category 'intermediate user type' dropped substantially when compare to the model with the reference category 'landline-only user'.

10.5 Number of Contact Attempts and Effects on Sample Structure

Sociological surveys are based on the voluntary participation of persons from the target population. If in Germany a sampled person refuses their cooperation, they cannot be contacted again for legal reasons[11]. With household selection procedures the contacted person can also refuse access to the sample person. Thus an increase of contact attempts can only influence the number of persons who are hard to contact and the number of person who are not reached at all. The final dispositions "unable to answer" and "refusal" cannot be influenced by increased contacting efforts. Since nonresponse is unequally distributed across subgroups, an increase in the number of contact attempts can cause bias within these groups. Many studies showed that older respondents with low education and physical impairments are easily to contact, but at the same time they more often refuse participation in surveys.

In contrast to studies conducted face to face frequent contact attempts are comparatively easy to realize in telephone surveys. But increasing contact attempts lower chances of successfully contacting persons in the target group. Therefore, it makes sense to consider the literature on call attempts for the administration of surveys and their efficiency. Purdon et al. (1999) demonstrated the importance of the number of contact attempts and used a logistic regression to model days and calling times, number of calls and temporal distance between them. Their results showed a decrease in contact probability for each additional contact attempt. Blasius and Reuband (1995) examined possible response rates for a telephone survey based on a register sample in Cologne (n = 362). They furthermore compared respondents easy to contact with those hard to contact with respect to their demographic characteristics, attitudes, and opinions. The authors concluded that the optimum number of contact attempts is between two and five. According to Blasius and Reuband further attempts would cause an overrepresentation of better educated and younger respondents. Also Neller (2005) showed that an increase in calls does not improve data quality. An increase in response rate will not always be beneficial since the positive effects on a

[11] Guideline on Telephone Surveys: http://www.adm-ev.de/fileadmin/user_upload/PDFS/R04_E_08.pdf, accessed June 30, 2011.

correction of sample biases are very limited. Although counterintuitive, according to Neller additional conversion attempts even seem to be counterproductive since they work best for those groups that cooperate and that tend to be already overrepresented in the sample. The results of Neller however are contrary to the findings of Philippens et al. (2004) who concluded that persons who initially refused but could be converted into respondents are more likely to be older, lower educated and with only little interest in politics. In their study conversion attempts thus led to a reduction of sample biases. Also Knesebeck and Lüschen (1999) achieved an improvement of sample quality with response-enhancing measures and reduced nonresponse among elder people and hence sample bias.

In the following we examine to what extent the number of contact attempts affects estimates of the target population under study. Therewith we are trying to show whether a high number of contact attempts leads to higher data quality and whether differences can be shown with respect to the applied mode. A distinguishing trait of mobile phones is that they allow owners to be reachable nearly everywhere. However, in contrast to landline telephones mobile phones can be switched off or muted, such that effectively persons are not always reachable.

Table 10.12 Number of Contact Attempts (CELLA 2 and Microcensus data)

	easy to reach	medium difficult to reach	difficult to reach	Microcensus data 2009
Sex (n)	*(1,866)*	*(873)*	*(268)*	*(69,501)*
Male	49.7	56.0	37.3	48.8
Female	50.3	44.0	62.7	51.2
Age Group (n)	*(1,844)*	*(860)*	*(261)*	*(69,501)*
16-19	7.0	6.1	4.9	5.2
20-29	19.2	18.1	16.7	14.2
30-39	15.2	18.6	20.9	14.5
40-49	18.9	23.5	23.3	19.8
50-59	17.4	17.9	22.2	16.3
60-69	11.8	8.7	6.8	13.6
70+	10.4	7.1	5.3	16.3
Marital Status (n)	*(1,866)*	*(873)*	*(268)*	*(69,501)*
Single	37.0	38.5	31.6	29.7
Married living together	45.3	45.8	54.0	52.8
Married separately living	2.0	2.4	2.6	2.0
Widowed	5.3	5.0	3.1	8.0
Divorced	10.0	7.2	8.5	7.5
Education (n)	*(1,851)*	*(873)*	*(266)*	*(69,501)*
Lower Secondary Level	23.3	19.6	19.0	38.8
Secondary Level	31.7	31.0	36.1	28.5
Highschool degree allowing access to polytechnical universities	--	--	--	6.0
A-level	40.0	46.2	38.5	19.7
Other	5.0	3.2	6.3	7.0

Design-weighted data, easy to reach = 1 - 2 contacts, medium difficult to reach = 3 - 4 contacts, difficult to reach = 5 + contacts.

Our CELLA 2 data show that especially women tend to be more difficult to reach than male respondents (see Table 10.12). This finding is difficult to explain, but according to our findings from CELLA 1 we assume that the places in which mobile phones are kept are important in this respect. While about 90 percent of male respondents keep their phones in the pockets of their jacket or trousers, almost three out of four women keep them in handbags, cases or bags (Schneiderat and Schlinzig 2009: 108). This makes it more difficult for women to access their mobile phone and receive calls. An unequal distribution of mobile phones across men and women is also a reasonable assumption, however so far one without empirical evidence. Something similar applies to the assumption that landline telephones are equally distributed or equivalently spread across the sexes of a certain population.

However, no significant differences can be observed with respect to age. As already described above only the elder respondents (age 60 +) are obviously easier to reach than other age groups which is due to the fact that this group is more often found at home. Also somewhat surprising is the fact that married respondents are more difficult to reach than others, since we know that especially single persons are significantly more mobile and likely to be in employment, take part in leisure activities and outdoor pursuits. Respondents with a lower formal education are easier to contact which might be explained by them being at home more often due to higher unemployment rates and exclusion from expensive leisure activities. However, this interpretation is not supported by the fact that our data showed no significant differences with respect to income.

Table 10.13 Results of a Linear Regression on Contact Attempts

Variables	
Regression Constant	3.37
Mode (R = Landline)	$-.34^{***}$
Age of Respondents	$-.01^{*}$
R^2	0.01
n	2,966

Design-weighted data, * $p < 0.05$. * $p < 0.01$. *** $p < 0.001$, R = landline.

Even if the explanatory power of the model is comparatively small, it shows that contact attempts to a mobile phone are more efficient than those to a landline (see Table 10.13), even though 77 % of the mobile phone respondents in our CELLA 2 study were reached at home. Blasius and Reuband (1995) showed that single and divorced respondents were difficult to recruit for an interview only after frequent contact attempts. Our mobile sample seems to compensate for this deficit. Respondents difficult to reach on a landline can be contacted more easily on a mobile phone. The opposite is true for the group of elder persons. They tend to be available only via a landline telephone and remain difficult to convince, as it was also found in previous telephone surveys. The inclusion of a mobile sample has somewhat increased the number of younger respondents who were reached after few contact attempts. In addition an increase in numbers of contact attempts leads to a further increase in the number of younger respondents while older persons may not be contacted once they have refused. Our CELLA 2 data also showed that the problem of high nonresponse rates among elder people cannot be solved by applying a dual frame approach. Oversampling in advance and additional weighting adjustments afterwards cannot provide a satisfactory solution for a biased sample. To research into telephone survey nonresponse the studies mentioned above conducted personal face-to-face interviews instead. Therefore, another complementary mode in addition to two telephone modes seems to achieve better response rates (Tailored Design Method).

10.6 Concluding Remarks

Despite an increase in online surveys[12] telephone surveys are still the most widespread kind of survey conducted to gather representative data on a population. For this reason it remains important to examine carefully the quality criteria of this kind of survey. As shown above, response rates and the number of interviews cannot be considered the only criteria for measuring survey quality, particularly since no generally accepted standards for the calculation of response rates exist in Germany. However, the use of equal standards for all kinds of telephone nonresponse outcomes is a precondition for comparing response rates and for describing trends in survey outcomes. Based on common standards, statements on declining response rates would have an empirical basis which as Schnell (1997) criticizes is currently lacking.

Following this concern the presentation of response rates for our CELLA studies were computed according to AAPOR standard definitions. The result of 21.9 % in CELLA 1 and about 13 % in CELLA 2 using the estimate ‚e' seems to be insufficient for scientific analysis at first sight. The comparison with data from the Microcensus studies as reference statistics showed however that certain subpopulations are well described by the CELLA data and that nonresponse is quite equally spread across the subpopulations under study.

To compare telephone survey data with reference data sets is a frequently used method to examine data quality. What is new is the application of this strategy in the combination of telephone survey data based on a dual frame approach and on the AAPOR standards.

The comparatively low response rates of our CELLA studies are both the result of a large number of cases with unknown eligibility (due to voicemail messages from the provider or mailboxes with different types of messages) and of high refusal rates. The scientifically common classification of nonresponse into reachability, ability to participate and willingness to participate cannot accurately represented by our CATI-survey. On the one hand, the distinction between ability to participate and willingness to participate cannot be selectively shown and, on the other hand, the reachability ratio cannot be measured directly. Furthermore, reachability can only be estimated by other parameters because at present there is

[12] In 2011 members of ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V) for the first time conducted more online than telephone interview (for numbers see http://www.adm-ev.de/, accessed May 03, 2011).

no reliable criterion to determine which generated telephone number of a RRD sample is connected to an eligible sample unit.

We showed that telephone surveys applying a dual frame approach can improve the reachability of respondents, however, this not the case with ability to participate and the willingness of the respondents to participate. We suspect that the novelty of contact attempts on mobile phones increased the response rates in our CELLA 1 study. Possible negative experiences of respondents with unexpected mobile phone calls from (commercial) call centers might have diminished the acceptance of telephone surveys.

In our view dual frame approaches have the following advantages: persons with both mobile and landline telephones can be contacted by integrating a mobile frame more easily and at the same time mobile-only respondents do not get excluded (under-coverage error). In contrast to other European countries, it will be necessary to conduct telephone surveys via both modes in Germany for a long time, since it is highly probable that both landline-onlys and mobile-onlys remain to exist for the years to come. The analysis of the two user types also showed that landline- and mobile-onlys strongly differ with respect to their demographic characteristics and their personality operationalized by the ‘Big Five’ dimensions.

Since reachability of respondents has improved by applying a dual frame approach, fewer contact attempts were necessary to interview respondents able and willing to do be interviewed. Too many contact attempts however might reduce the number of the potential respondents usually underrepresented. Particularly elderly people can only be contacted via a landline. In turn, this group is very hard to convince to participate in a telephone survey.

The disadvantage of the dual frame approach is that two separate samples have to be drawn; they have to be separately administered and combined with by means of weighting. In order to be able to accomplish a correct weighting according to reachability criteria some important items need to be integrated into the questionnaire, which tend to be annoying for respondents.

Nonetheless, we argue that telephone surveys should apply a dual frame approach, since it provides the best conditions for high heterogeneity and a lower risk of coverage bias. In addition, this approach needs fewer contact attempts when aiming at good estimates of the target population. As Blasius and Reuband (1995), Schneekloth and Leven (2003), Neller (2005), and our CELLA 2 showed a higher number of contact attempts might decrease the representativeness by including a lower proportion of lower educated respondents and elder people, and by consequently increasing the risk of coverage bias.

References

ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.). Figures on the Market for Market Research, vol. 2) (2011), `http://www.adm-ev.de/index.php?id=zahlen&type=1&L=1&type=1` (accessed April 13, 2010)

AAPOR (The American Association for Public Opinion Research). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Vol. 7 (2011), `http://www.aapor.org/Standard_Definitions/3049.htm` (accessed June 30, 2011)

Aust, F., Schröder, H.: Sinkende Stichprobenausschöpfung in der Umfrageforschung – ein Bericht aus der Praxis (Declining Sampling Coverage in Survey Research – a Report from Practical Experience). In: Weichbold, M., et al. (eds.) Umfrageforschung – Herausforderungen und Grenzen. Österreichische Zeitschrift für Soziologie, vol. 9, pp. 195–212 (2009)

Blasius, J., Reuband, K.-H.: Telefoninterviews in der empirischen Sozialforschung: Ausschöpfungsquoten und Antwortqualität (Telephone Interviews in Empirical Social Research: Response Rates and Response Quality). ZA-Information 37, 64–87 (1995)

Blumberg, S.J., Luke, J.V., Cynamon, M.L., Frankel, M.R.: Recent Trends in Household Telephone Coverage in the United States. In: Lepkowski, J., et al. (eds.) Telephone Survey Methodology, pp. 56–86. John Wiley & Sons, Hoboken (2008)

Busse, B., Fuchs, M.: The components of landline telephone survey coverage bias. The relative importance of nophone and mobile-only populations. Quality & Quantity 45, 1–17 (2011)

Carley-Baxter, L.R., Peytchev, A., Black, M.C.: Comparison of Cell Phone and Landline Surveys: A Design Perspective. Field Methods 22(1), 3–15 (2010)

Curtin, R., Presser, S., Singer, E.: Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly 69(1), 87–98 (2005)

Däubler, T.: Nonresponseanalysen der Stichprobe F des SOEP (Nonresponse Analyses SOEP Sample F). DIW Materialien 15 (2002), http://www.diw-berlin.de/documents/publikationen/73/38830/diw_rn02-05-15.pdf (accessed September 21, 2008)

de Heer, W.: International Response Trends: Results of an International Survey. Journal of Official Statistics 15(2), 129–142 (1999)

de Leeuw, E.D., Hox, J.J.: I am not selling anything: 29 experiments in telephone introductions. International Journal of Public Opinion Research 16(4), 464–473 (2004)

de Leeuw, E.D., Callegaro, M., Hox, J.J., Korendijk, E., Lensvelt-Mulders, G.: The influence of advance letters on response in telephone surveys: a metaanalysis. Public Opinion Quarterly 71(3), 1–31 (2007)

de Leeuw, E., de Heer, W.: Trends in household survey nonresponse. A longitudinal and international comparison. In: Groves, R.M., et al. (eds.) Survey Nonresponse, pp. 41–54. John Wiley & Sons, New York (2002)

Diekmann, A.: Die Demoskopie verletzt Informationspflichten. Zur Bedeutung von Interview-Verweigerung und misslungenen Kontaktversuchen (Polls Violate Reporting Requirements. On the Importance of Refusals and Unsuccessful Contact Attempts). Neue Züricher Zeitung 276, 14 (2010)

Dillman, D.A.: Mail and telephone surveys. In: The total design method. John Wiley & Sons, New York (1978)

Dillman, D.A., Gallegos, J.G., Frey, J.H.: Reducing Refusal Rates for Telephone Interviews. Public Opinion Quarterly 40, 66–78 (1976)

Ehlen, J., Ehlen, P.: Celluar-Only substituions in the United States as lifestyle adoption. Public Opinion Quarterly 71(4), 717–733 (2007)

Europäische Kommission. E-Communications Haushaltsumfrage – Bericht. Spezial-Eurobarometer 335 (Household survey) (2010), http://ec.europa.eu/public_opinion/archives/ebs/ebs_335_de.pdf (accessed October 10, 2010)

Field, A.: Discovering statistics using SPSS (and sex, drugs, and rock ’n’ roll). SAGE Publications, London (2005)

Fuchs, M.: Zukünftige Herausforderungen der Umfrageforschung (Future Challenges of Survey Research). In: Faulbaum, F., Wolf, C. (eds.) Gesellschaftliche Entwicklungen im Spiegel der Empirischen Sozialforschung, pp. 227–252. VS Verlag für Sozialwissenschaften, Wiesbaden (2010)

Gerlitz, J.-Y., Schupp, J.: Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP (Research on Big5 Personality Traits with SOEP Data). DIW Berlin Research Notes 4/2005 (April 2005), http://www.diw.de/sixcms/detail.php?id=diw_02.c.231308.de (accessed April 28, 2011)

Goldstein, K.M., Jennings, K.M.: The Effect of Advance Letters on Cooperation in a List Sample telephone Survey. Public Opinion Quarterly 66, 608–617 (2002)

Groves, R.M.: Survey Errors and Survey Costs. John Wiley & Sons, New York (1989)

Groves, R.M.: Survey Errors and Survey Costs. John Wiley & Sons, Hoboken (2004)

Groves, R.M., Peytcheva, E.: The Impact of Nonresponse Rates on Nonresponse Bias. A Meta-Analysis. Public Opinion Quarterly 72, 167–189 (2008)

Häder, M., Häder, S.: Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung. In: Telephone Interviews over Mobile Phone. Concept, Design and Implementation of a Data Collection Strategy. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Häder, M., Häder, S.: Der Einfluss der Persönlichkeit auf die Antwortqualität – empirische Befunde aus der Umfrageforschung über das Festnetz und den Mobilfunk (The Impact of Personality on Answer Quality – Empirical Findings from Research on Landline and Mobile Phone Surveys). In: KZfSS (2011) (in print)

Hartmann, P.H., Schimpl-Neimanns, B.: Zur Repräsentativität soziodemographischer Merkmals des ALLBUS – Multivariate Analysen zum Mittelschichtbias der Umfrageforschung (Representativity of Socio-Demographic Characteristics of ALLBUS data – Multivariate Analysis on Middle Class Bias of Survey Research). ZUMA Arbeitsbericht (January 1992)

Holbrook, A., Krosnick, J.A., Pfent, A.: The Causes and Consequences of Response Rates in Surveys by the News Media and Government Contractor Survey Research Firms. In: Lepkowski, J.M., et al. (eds.) Telephone Survey Methodology, pp. 499–528. John Wiley & Sons, Hoboken (2008)

Hüfken, V.: Kontaktierung bei Telefonumfragen (Contact Modes of Telephone Surveys). In: Hüfken, V. (ed.) Methoden in Telefonumfragen, pp. 11–31. VS Verlag für Sozialwissenschaften, Wiesbaden (2000)

von dem Knesebeck, O., Lüschen, G.: Die Bedeutung ausschöpfungssteigernder Maßnahmen für die medizinsoziologische Forschung: Das Beispiel einer telefonischen Befragung alter Menschen (The Importance of Response Increasing Measures for Medical Sociological Research: The Example of a Telephone Survey with Older People). Sozial- und Präventivmedizin/Social and Preventive Medicine 44, 233–241 (1999)

Koch, A.: “mehr” nicht gleichbedeutend mit "besser" ist: Ausschöpfungsquoten und Stichprobenverzerrungen in Allgemeinen Bevölkerungsumfragen (If„more“ is not synonymous with„better“: response rates and sample bias in general population survey). ZUMA-Nachrichten 42(22), 66–90 (1998)

Lang, F.R., Lüdkte, O., Asendorpf, J.B.: Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen (Test Quality and Psychometric Equivalence of the German Version of the Big Five Inventory with Young, Middle and Old Adults). Diagnostica 47(3), 111–121 (2001)

Lepkowski, J., Tucker, C., Brick, J.M., de Leeuw, E.D., Japec, L., Lavrakas, P.J., Link, M.W., Sangster, R.L. (eds.): Telephone Survey Methodology. John Wiley & Sons, Hoboken (2008)

Meier, G.: Telefonbefragungen über das Mobilfunknetz: Konzept, Design und Umsetzung einer Strategie zur Datenerhebung' ('Telephone Interviews over Mobile Phone. Concept, Design and Implementation of a Data Collection Strategy') (2009); Zu , R., Häder, Häder: In GESIS – Leibniz-Institut für Sozialwissenschaften . Medien, Daten Analysen 3(1), 116-119 (2009)

Meier, G., Schneid, M., Stegmann, Y., Stiegler, A.: Steigerung der Ausschöpfungsquote von Telefonumfragen durch geschickte Einleitungstexte (Increasing Response Rates of Telephone Surveys by Using Suitable Introductions). ZUMA-Nachrichten 57(29), 37–56 (2005)

Muck, P.: Persönlichkeit und berufsbezogenes Sozialverhalten (Personality and Job-related Social Behaviour). In: Schuler, H. (ed.) Lehrbuch der Personalpsychologie, vol. 2, pp. 527–578. Hogrefe Verlag, Göttingen (2006)

Neller, K.: Kooperation und Verweigerung: eine Non-Response-Studie (Cooperation and Refusal: A Nonresponse Study). ZUMA Nachrichten 57(29), 9–36 (2005)

Ostendorf, F., Angleitner, A.: NEO-PI-R: NEO-Persönlichkeits-Inventar – revidierte Form (Personality Inventory – Revised Version). In: Sarges, W., Wottawa, H. (eds.) Handbuch Wirtschaftspsychologischer Testverfahren, pp. 409–415. Pabst Science Publishers, Lengerich (2001)

Ostendorf, F., Angleitner, A.: NEO-PI-R – NEO Persönlichkeitsinventar nach Costa und McCrae - Revidierte Fassung. NEO-PI-R – NEO Personality Traits Inventory by Costa and McCrae – Revised Version. Hogrefe, Göttingen (2004)

Peytchev, A., Carley-Baxter, L.R., Black, M.C.: Coverage Bias in Variances, Associations, and Total Error From Exclusion of the Cell Phone-Only Population in the United States. Social Science Computer Review 28(3), 287–302 (2010)

Peytchev, A., Carley-Baxter, L.R., Black, M.C.: Multiple Sources of Nonobservation Error in Telephone Surveys: Coverage and Nonresponse. Social Methods & Research 40(1), 138–168 (2011)

Philippens, M., Billiet, J., von Loosveldt, G., Stoop, I., Koch, A.: Non-Response and Fieldwork Efforts in the ESS: Results from the Analysis of Call Record Data. Work Package 7, Data-Bases Quality Assessment in the ESS, (2004) `http://naticent02.uuhost.uk.uu.net/methodology/nonresponse_fieldwork_efforts.pdf` (accessed September 22, 2009)

Porst, R.: Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Die Sicht der Institute (Response Rates in Social Surveys. The Perspective of the Institutes). ZUMA-Arbeitsbericht (July 1996)

Purdon, S., Campanelli, P., Sturgis, P.: Interviewers' Calling Strategies on Face-to-Face Interview Surveys. Journal of Official Statistics 15, 199–216 (1999)

Rammstedt, B., Koch, K., Borg, I., Reitz, T.: Entwicklung und Validierung einer Kurzskala für die Messung der Big Five Persönlichkeitsdimensionen in Umfragen (Development and Validation of a Short Scale for the Measurement of Big5 Personality Dimensions in Surveys). ZUMA Nachrichten 55, 5–28 (2004)

Rammstedt, B., John, O.P.: Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit (Short Version of Big5 Inventory (BFI-K): Development and Validation of an Economic Inventory for the Assessment of Five Personality Dimensions). Diagnostica 51(4), 195–206 (2005)

Rammstedt, B., John, O.P.: Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41, 203–212 (2007)
Rammstedt, B.: The 10-Item Big Five Inventory: Norm Values and Investigation of Sociodemographic Effects Based on a German Population Representative Sample. European Journal of Psychological Assessment 23(3), 193–201 (2007)
Schlinzig, T., Schneiderat, G.: Möglichkeiten zur Erhöhung der Teilnahmebereitschaft bei Telefonumfragen über Festnetz und Mobilfunk. Zum Potenzial von Warmkontakt und randomisierter Geburtstagsauswahl (Possibilities to Improve Response Rates in Landline and Mobile Phone Surveys. On Effects of Prior Notices and an Alternative Within Household Respondent Selection Technique). In: Weichbold, M., et al. (eds.) Umfrageforschung. Grenzen und Herausforderung, Sonderheft der Österreichischen Zeitschrift für Soziologie, vol. 9, pp. 21–43. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)
Schnauber, A., Daschmann, G.: Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews ? (Factors Influencing the Willingness to Participate in Telephone Surveys). Methoden, Daten, Analysen 2(2), 97–123 (2008)
Schneekloth, U., Leven, I.: Woran bemisst sich eine gute allgemeine Bevölkerungsumfrage? (How to measure a good general population survey?). ZUMA-Nachrichten 53(27), 16–57 (2003)
Schneiderat, G., Schlinzig, T.: Das Mobilfunktelefonverhalten in der Allgemeinbevölkerung (Mobile Phone Habits of the General German Population). In: Häder, M., Häder, S. (eds.) Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung, pp. 99–119. VS-Verlag für Sozialwissenschaften, Wiesbaden (2009)
Schnell, R.: Non-Response in Bevölkerungsumfragen (Non-response in Population Surveys). Leske + Budrich, Opladen (1997)
Stoop, I.A.L.: The Hunt for the Last Respondent. Nonresponse in Sample Surveys. Social and Cultural Planning Office of the Netherlands, The Hague (2005)
Struminskaya, B., Schaurer, I., Bandilla, W., Gabler, S., Häder, S., Kaczmirek, L.: Improved cost-effectiveness in mobile surveys using HLR-Lookup. Paper presented at the GOR 11 Conference, Düsseldorf, March 14-16 (2011), http://conftool.gor.de/conftool11/index.php?page=downloadPaper&filename=Struminskaya-Improved_cost-effectiveness_in_mobile_surveys_using_HLR-Lookup-173.pdf&form_id=173&form_index=&form_version=final (accessed April 28, 2011)
TNS Infratest Sozialforschung, Methodenbericht zum Befragungsjahr 2009 (Welle 26) des Sozio-ökonomischen Panels. (Methodological Survey Report 2009), München (2010), http://www.tns-infratest-sofo.com/downloads/2010MB_%20SOEP2009.pdf (accessed April 04, 2011)

Part III
Weighting and Nonresponse

Chapter 11

Weighting for Unequal Inclusion Probabilities and Nonresponse in Dual Frame Telephone Surveys

Siegfried Gabler, Sabine Häder, Iris Lehnhoff, and Elisabeth Mardian

11.1 Introduction

There are two main reasons for using weights in sample surveys. The first is the need to account for unequal inclusion probabilities which may result from various processes such as disproportional selection of units within strata. In telephone surveys, inclusion probabilities vary for example because of differential amounts of telephone numbers people have or differential numbers of persons in the household belonging to the target population. Therefore, design weights have to be calculated and applied during data analysis. In section 11.2 of this chapter, we describe how this may be handled in combined landline and mobile phone surveys.

In addition to, the efforts involved in finding suitable frames without over- or under-coverage and in drawing gross samples accurately, there is a second reason why the data might need to be weighted. If nonresponse to a survey is systematic, the estimates will be biased. The chapter by Schneiderat and Schlinzig in this book illustrates this at the example of the CELLA 2 samples. They describe how (design weighted) distributions of socio-demographic variables differed from the corresponding distributions in the Microcensus, a 1 % sample of about 800,000 individuals which provides official representative statistics for the population of Germany.

In the following chapter we show how samples from a landline phone survey and a mobile phone survey can be combined. The appropriate weighting procedure and the necessary parameters are described in detail.

However, in our combined – i.e. design weighted – sample there still seems to be a notable nonresponse error. We show this in an excursus for the variable procrastination where we compare survey participants and drop-outs (as proxies for nonrespondents). To account for this, we finally describe an adjustment weighting procedure in order to reduce this error.

11.2 Design Weighting: The Dual Frame Approach

11.2.1 Theoretical Background

Often a single sampling frame does not cover all the units of a target population, as it is the case with the frame of landline phones. If a second frame is available, for example a frame of mobile phone numbers, and both frames together cover (nearly) the whole population, then samples should be drawn independently from the two frames. The difficulty is that the two frames may have an overlap. Thus, the task of the researcher is to handle this overlap to get approximately unbiased estimates. One possibility is to compute the inclusion probabilities for the units of the whole population and to use the Horvitz-Thompson (HT) estimator for estimating the total of a variable of interest. The advantage of this procedure is its simplicity and the possibility to extend this estimator using the GREG (generalized regression) estimator to take additional biasing error sources, such as nonresponse, into account (see section 11.3 of this chapter).

For a general formula of the inclusion probabilities of persons we need to define the relevant parameters (see Table 11.1).

Table 11.1 Parameters needed for the Dual Frame Model

Landline phone		Mobile phone	
M^F	frame size of numbers	M^C	frame size of numbers
m^F	sample size of numbers	m^C	sample size of numbers
k_i^F	size of landline numbers allowing access to target person *i*	k_i^C	amount of mobile phone numbers allowing access to target person *i*
z_i	size of household to which target person *i* belongs		

Furthermore, to simplify the formula we make the following fundamental assumption:

The probability that two (not necessarily distinct) members of the same household are selected into the sample through different channels is negligible.

Obviously, this assumption might be problematic for relatively small populations. But in our case, where the frames contain 139 million numbers (landline phones) and 197 million numbers (mobile phones), we can neglect the chance of selecting an individual twice from different frames.

According to Gabler and Ayhan (2007) the inclusion probability of a target person via landline frame is

$$\pi_i^F = \sum_{j=0}^{k_i^F} \left(1-\left(1-\frac{1}{z_i}\right)^j\right) \frac{\binom{k_i^F}{j}\binom{M^F-k_i^F}{m^F-j}}{\binom{M^F}{m^F}}$$

$$\approx \frac{1}{z_i}\sum_{j=0}^{k_i^F} j \frac{\binom{k_i^F}{j}\binom{M^F-k_i^F}{m^F-j}}{\binom{M^F}{m^F}}$$

$$= k_i^F \frac{m^F}{M^F}\cdot\frac{1}{z_i} \ .$$

Quite analogously, the inclusion probability of a target person via mobile phone frame is

$$\pi_i^C = \sum_{j=1}^{k_i^C} \frac{\binom{k_i^C}{j}\binom{M^C-k_i^C}{m^C-j}}{\binom{M^C}{m^C}} = 1-\frac{\binom{M^C-k_i^C}{m^C}}{\binom{M^C}{m^C}}$$

$$= 1-\prod_{j=0}^{m^C-1}\left(1-\frac{k_i^C}{M^C-j}\right) \approx 1-\left(1-\frac{k_i^C}{M^C}\right)^{m^C}$$

$$\approx k_i^C \frac{m^C}{M^C} \ .$$

Applying the fundamental assumption above we deduce

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

and for the inclusion probability of the target person i we get

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

11.2.2 Fixing the Model Parameters

In the following section we discuss how the parameters of the Dual Frame Approach that we introduced above may be estimated.

Contactability

To include mobile phone samples in telephone surveys we are not only interested in a considerable part of the population owning such a phone. In addition, this phone needs to be turned on (at least from time to time) so that the potential respondent may be reached by the survey agency. In our survey CELLA 2 we asked respondents how often they usually have their mobile phone switched on. 58.5 % answered that it is always ready-to-receive. 1,064 respondents told us that their mobile phone is switched on only from time to time. Of these 70.1 % turn it on daily and 17.4 % several times a week. Thus, only 133 out of 2,567 (5.2 %) have their mobile phone ready-to-receive only several times a month or even less often. Fortunately, there seems to be no gender difference in this behavior in the sample: Male respondents reported having their mobile phone switched on about 13.69 hours a day, while female respondents reported 13.56 hours a day.

During longer fieldwork periods, when sample persons are connected on different days and at different times of a day, it should be possible to reach nearly all of them sooner or later if enough contact attempts are undertaken. Therefore, we have no parameter to adjust for differential contactability in our model.

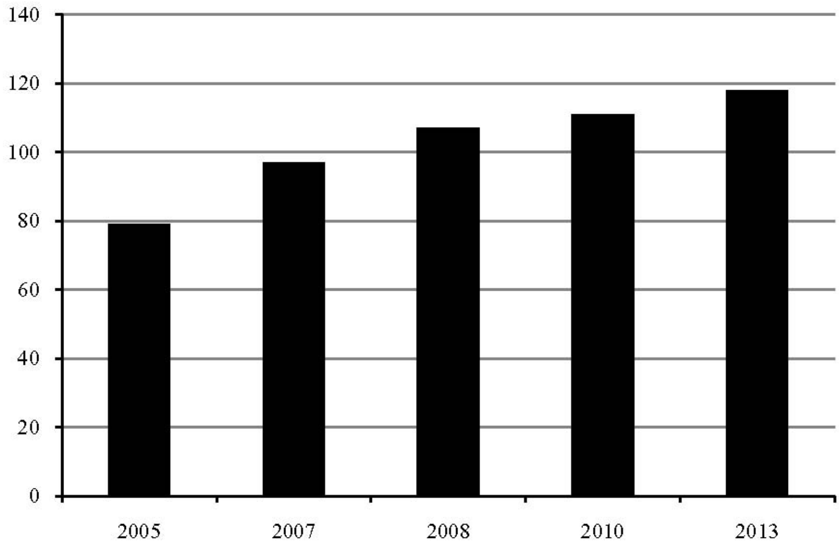
Number of Mobile Phones, Sharing Mobile Phones

A crucial problem in telephone sampling for landline phones is the selection of the target person within the household, since landline phones are typically used by the whole household. Usually, a method derived from the Kish grid or the last/next birthday selection method is applied to ensure a random selection with known inclusion probabilities. In the subsequent design weighting the number of persons belonging to the target population in the household has to be considered – information that has to be collected during the interview. In some cases this is a sensitive question since especially older women living in 1-person-households do not like giving such information to a stranger.

However, for mobile phone samples there is wide agreement amongst survey statisticians that this procedure can be dropped, because a “mobile telephone is a personal appliance, like a wrist watch.” (Kuusela, spotlight on Finland in this book). In the following we show that this is a reasonable assumption. We illustrate this with the example of Germany.

In 2009 approximately 77.7 million individuals aged 6 years and older lived in Germany[1] – we assume that children start using a mobile phone from the age of 6 onward. Compared to this the number of mobile phones seems to be very high. In 2010 about 111 million customers owned a mobile phone in Germany (see Figure 11.1). Obviously, this is an indication of people using multiple SIM cards concurrently. In fact, in particular younger people sometimes use different cards of different providers for different purposes.

[1] See http://www.destatis.de.



Source: DigiWorld Yearbook 2009 p. 82 and 2010 p. 82

Fig. 11.1 Number of mobile phone users in Germany 2005 – 2013 (in millions)

Table 11.2 shows the amount of mobile phone numbers on which a respondent can be reached. It seems to support our hypothesis of multiple SIM card or mobile phone usage since about 10 % of the respondents are available on more than one mobile phone number.

Table 11.2 Quantity of mobile phone numbers on which a respondent is available

Numbers	Frequency	Percent
One	2,282	89.3
Two	230	9.0
Three	30	1.2
Four	5	0.2
Five	2	0.1
Six and more	6	0.2
	2,555	100

Source: CELLA 2

In CELLA 2 we also asked the respondents how many people usually use their mobile phone (passive sharing, Busse and Fuchs 2011). The answers are shown in Table 11.3. One might assume that those who have more than one mobile phone number are those who have an own phone and share another phone with others. But this is not the case. Only 38 out of 273 respondents who have more than one mobile phone number belong to those people who share their mobile phone.

Table 11.3 Who uses the respondent's mobile phone?

	Frequency	Percent
Exclusive use by myself	2,323	90.4
Other people use it from time to time	186	7.2
Share it with others	51	2.0
Other	10	0.4
	2,570	100

Source: CELLA 2

At the same time, only 1.9 % of those who share their mobile phone with others are available on only one number. For this part of the population a within-household respondent selection would be necessary.

To also investigate if respondents use mobile phones of other household members or other persons close to them we asked in CELLA 2 whether they answer calls on the mobiles of these persons. The answers are shown in Table 11.4.

Table 11.4 Do you answer calls on mobiles of other household members or other persons close to you?

	Frequency	Percent
Yes, always	170	6.6
Yes, mostly	162	6.3
Yes, seldom	519	20.2
No, never	1,258	49.0
Only if requested by the owner	438	17.0
Other household members do not have a mobile phone	22	0.9
	2,569	100

Source: CELLA 2

Obviously, the majority of the respondents would never come into a situation where they might answer calls on mobile phones of other household members. This becomes even clearer when analyzing the answers to the question "How likely or unlikely are you to participate in a social survey on someone else's mobile phone?" (see Table 11.5). This question was asked to all respondents who answered "Yes" to the question in Table 11.4.

Table 11.5 How likely or unlikely are you to participate in a social survey on someone else's mobile phone?

	Frequency	Percent
Very likely	28	2.1
Likely	36	2.8
Partly likely/partly unlikely	57	4.4
Unlikely	441	34.3
Very unlikely	725	56.4
	1,285	100

Source: CELLA 2

Consciously, we have not restricted our question to possible “sharers” within the household but included also other persons close to the respondent. This restriction was criticized by Busse and Fuchs (2011). However, the amount of individuals having only one mobile phone number and sharing this phone with others is so small that we do not worry about unequal inclusion probabilities and instead treat mobile phones as personal devices – as was suggested by the spotlights of various countries in this book. In addition, the small advantage of including a within-household selection procedure in both the landline and the mobile phone survey would be erased by an increase in nonresponse during the contact phase.

Amount of Landline Phone Numbers

To construct design weights for unequal inclusion probabilities we need to know the amount of landline phone numbers in a household. The reason for this is that the more numbers are included in the sampling frame the higher is the probability of an individual to be drawn. However, what initially seemed to be a simple issue turned out to be a difficult question during the course of our experiments. In CELLA 1 we chose the following solution. In households where an analogue modem was used, we set the amount of landline numbers to 1 for technical reasons. For households with ISDN connections (Integrated Services Digital Network) we calculated a mean of 2.5 numbers per household from a previous survey. We applied this rule of thumb – parameter $k_i^F = 1$ for analogue modems and $k_i^F = 2.5$ for ISDN-connections – for the adjustment in the dual frame model (Häder and Häder 2009, Callegaro et al. 2011). However, this rule turned out to be too imprecise and constantly changing. Thus, in CELLA 2 we asked respondents about the amount of their landline phone numbers – although we have learnt that the answers are often wrong because people simply do not know the exact answer.

In the following Table 11.6 we show the results for analogue and ISDN modems.

Table 11.6 Amount of landline phone numbers in households with analogue and ISDN modems

Amount of numbers	*Analogue modem*		*ISDN modem*	
	Frequency	Percent	Frequency	Percent
1	1,227	76.6	615	58.1
2	242	15.1	251	23.7
3	95	5.9	137	12.9
4	17	1.1	31	2.9
5	9	0.6	10	1.0
6	12	0.8	16	1.5
	1,602	100	1,060	100

Source: CELLA 2

Respondents with an analogue modem in their household reported a mean of 1.36 landline numbers, those in ISDN-households reported a mean of 1.69 numbers. Obviously, our above assumption does not hold any longer and we have to revise it. For the computation of the design weights we used the numbers reported by the respondents. 10 respondents stated having 10 or more landline phone numbers – these were cut to 6.

To ultimately determine the exact amount of telephone numbers on which a household may be reached further research is necessary. Experiments have shown (Meier 2007) that the amount of landline telephone numbers reported by respondents and the amount of phone numbers specified in contracts do not match.

Household Size

In the course of the interview we also asked how many people aged 16 years and older are living in the respondent's household. The household size of people who reported living in households with more than 6 members aged 16 years or older was cut to 6 since the target population was restricted to people living in private – not institutional – households. The sample distribution is shown in Table 11.7.

Table 11.7 Target population household size

Household size	Frequency	Percent
1	86	3.7
2	1,480	63.0
3	482	20.5
4	205	8.7
5	69	3.0
6	26	1.1
	2,349	100

Source: CELLA 2

Furthermore, we need the following parameters for fixing our model.

M^F = 139,366,300 (Number of landline phone numbers in the universe)

m^F = 31,358 (Number of landline phone numbers in the sample)

M^C = 197,490,000 (Number of mobile phone numbers in the universe)

m^C = 44,330 (Number of mobile phone numbers in the sample)

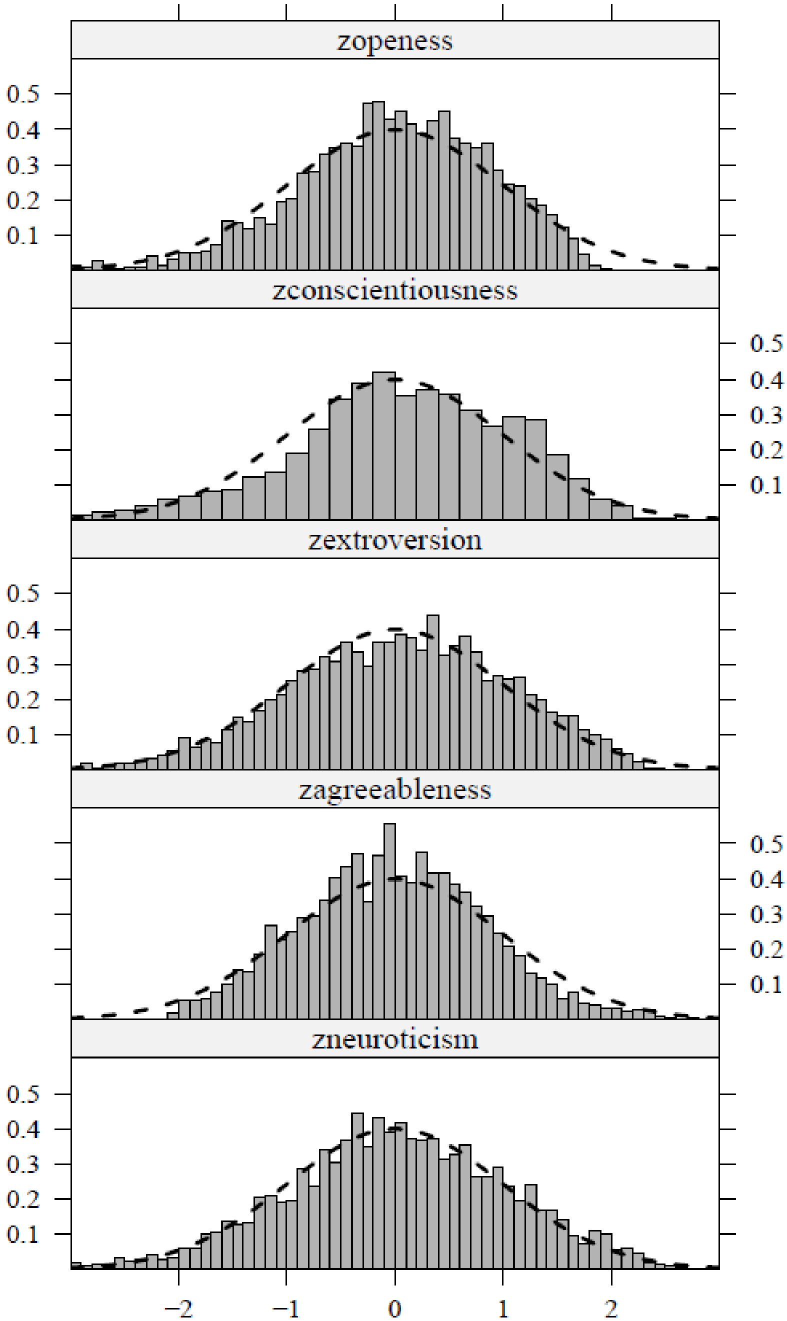
The concluding inclusion probabilities $\pi_i \approx k_i^F \frac{m^F}{M^F} \cdot \frac{1}{z_i} + k_i^C \frac{m^C}{M^C}$ are used to construct both the HT-estimator and the GREG-estimator. The results of applying these estimators will be shown in section 11.4.

11.2.3 Practical Application

In the following application we demonstrate the impact of design weighting on a set of variables for which we do not have the population distributions. One part of the CELLA 2 questionnaire included the 15 item inventory of the Big Five (see John et al. 2008, Dehne and Schupp 2007). This concept provides a score for each of the Big Five personality traits (Conscientiousness, Agreeableness, Neuroticism, Extroversion, and Openness). Scores on these traits can often explain issues of daily life such as participation in surveys (see Häder and Häder forthcoming).

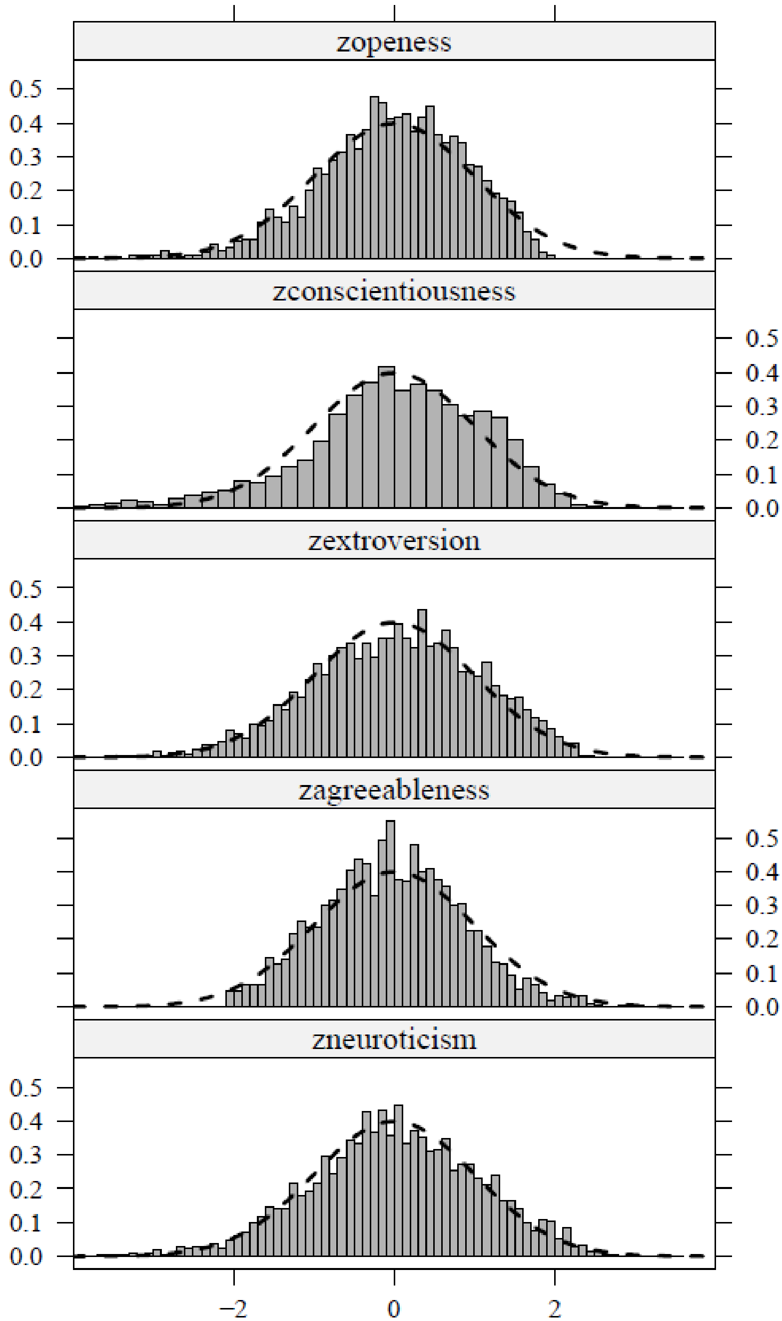
In the literature we found indications that the factor indices resulting from a factor analysis of the Big Five items are nearly normally distributed: "Scores on dimensions will fall along a normal distribution (bell curve)."[2] Thus, we have a criterion for the assessment of the data quality. We will check how closely the un-weighted and weighted data follow a normal distribution. For this, we first plot the standardized Big Five factor indices together with the standard normal distribution (see Figures 11.2a and 11.2b).

[2] See http://www.centacs.com/research-development/the-big-five.



Source: CELLA 2

Fig. 11.2a Histograms of the unweighted standardized factor indices of the Big Five and normal distribution



Source: CELLA 2

Fig. 11.2b Histograms of the weighted standardized factor indices of the Big Five and normal distribution

Applying the adjusted Pearson or Likelihood Ratio statistic test for testing the homogeneity (see Rao-Scott 1984) we first have to define the number of intervals (bins). For the number of bins we use the smallest integer greater or equal to $2 \cdot n^{2/5}$ which is 50 in our case. This formula is given for example in the R–package "nortest".

As we can see from Table 11.8, the assumption of a normal distribution for each variable has to be rejected for four out of five variables. The output was generated by the complex sample module in IBM SPSS 19.0. If we had considered the reduced degrees of freedom due to the fact that two parameters of the normal distribution had to be estimated, the assumption of normality would have had to be rejected for all five variables.

Table 11.8 Test of homogeneous proportions (unweighted)

		Chi-Square	Adjusted F	df1	df2	Sig.
zopeness	Pearson	60.891	1.242	49	144,158	.119
	Likelihood Ratio	63.054	1.286	49	144,158	.086
zconscientiousness	Pearson	243.086	4.959	49	144,158	.000
	Likelihood Ratio	243.086	5.670	49	144,158	.000
zextroversion	Pearson	108.869	2.221	49	144,158	.000
	Likelihood Ratio	129.691	2.646	49	144,158	.000
zagreeableness	Pearson	164.866	3.363	49	144,158	.000
	Likelihood Ratio	202.965	4.141	49	144,158	.000
zneuroticism	Pearson	67.007	1.367	49	144,158	.045
	Likelihood Ratio	68.725	1.402	49	144,158	.033

Source: CELLA 2

The adjusted F is a variant of the second-order Rao-Scott adjusted chi-square statistic. Significance is based on the adjusted F and its degrees of freedom.

The question is whether the weighted Pearson's goodness of fit-test and/or the weighted Likelihood Ratio test provide differential results. The corresponding figures are given in Tables 11.8 and 11.9.

Table 11.9 Test of homogeneous proportions (weighted)

		Chi-Square	Adjusted F	df1	df2	Sig.
zopeness	Pearson	54.558	.837	46.396	136,498.174	.778
	Likelihood Ratio	54.439	.835	46.396	136,498.174	.781
zconscientiousness	Pearson	274.078	4.231	47.233	138,959.591	.000
	Likelihood Ratio	311.830	4.814	47.233	138,959.591	.000

Table 11.9 *(continued)*

zextroversion	Pearson	88.177	1.339	46.743	137,517.678	.060
	Likelihood Ratio	95.888	1.457	46.743	137,517.678	.022
zagreeableness	Pearson	147.983	2.288	47.440	139,568.669	.000
	Likelihood Ratio	188.133	2.909	47.440	139,568.669	.000
zneuroticism	Pearson	64.718	.991	47.291	139,130.432	.490
	Likelihood Ratio	66.240	1.014	47.291	139,130.432	.445

Source: CELLA 2

The tables show that additionally considering the sampling process does not enable rejecting the null hypothesis of normality of zopeness, zextroversion, and zneuroticism. Furthermore, the null hypothesis for the variables zconscientiousness and zagreeableness has to be rejected. Consequently, it seems that more powerful normality tests than Pearson's goodness of fit test or the Likelihood Ratio test for survey data with weights need to be developed.

A weighted goodness of fit test using the adjustment weights is not meaningful since the adjustment variables are arbitrary and the weights cannot be interpreted as response probabilities.

11.3 Combined Design and Adjustment Weighting

11.3.1 Excursus: Procrastination

We found evidence that not only the estimates of socio-demographic variables are biased because of nonresponse but also those of other variables of interest. In the following we show this for procrastination since this variable could be interesting as a reason for the participation in surveys. We will show that the measure of procrastination in CELLA 2 is likely to be biased. A comparison of the procrastination variable across survey participants and drop-outs points in this direction.

Procrastination means that high-priority tasks will be replaced with activities of low-priority, with the result that the action does not correspond to one's intention. Procrastination is thus a form of disturbance in one's self-monitoring related to affective, cognitive, and motivational factors. Surprisingly, many people state that they put off important things even if this is disadvantageous and they would like to change their behavior. In a study at Münster University Rist et al. (2006) found that between 10 and 20 % of students consider themselves procrastinators. In our context procrastination is of interest, since it might help explaining participation in surveys.

Procrastination is often found among students. Thus, many survey instruments are designed for an academic context such as the well-known Aitken Procrastination Inventory (Aitken 1982) or the Procrastination Questionnaire for Students – PFS (Glöckner-Rist et al. 2010). Hence, it seems difficult to choose a scale for a survey like CELLA 2, which addresses a sample of the entire population. However, the PFS was designed with the intention of measuring procrastination as cross-situational behavioral disposition. Additionally, the PFS does not contain items that are related in terms of content. Furthermore, a version of this questionnaire is available with a reduced amount of items. This is especially relevant for telephone interviews which are typically of limited duration. Therefore, the procrastination scale used in CELLA 2 is based on the PFS.

In the course of the survey four behaviors that address the individual propensity to postpone tasks were investigated[3]. The respondents were asked to rate how frequently such a behavior occurred during the last two weeks by means of a five-point scale[4]. The individual procrastination tendency results from the sum of the self-reported values.

Regarding the two aspects ‘procrastination tendency’ and ‘participation in interviews’, respondents that complete the survey (participants) can be expected to perceived the call as a welcome change and an opportunity to put off other activities. Accordingly, we may hypothesize that respondents who interrupt surveys (drop-outs) are less prone to procrastinate than persons who complete the whole interview. This raises the question whether drop-outs differ from participants in their tendency to procrastinate. On the basis of the data from CELLA 2, this question will be approached later in this section by making a comparison between drop-outs and participants. Before that, the total number of drop-outs as well as the moments of interruption are briefly discussed.

Out of all calls taken and all interviews realized or discontinued (n = 3,456), 87 % (n = 3,007) of the interviews had been conducted completely and only 13% (n = 449) were aborted. A large number of participants (n = 2,973) provided information on all four items of the question block that addressed procrastination. Out of the drop-outs, 69 gave evaluable answers to this question block. This seemingly small number most likely arises from the position of the respective question block in the middle of the survey (questions 28 - 31 out of 63, see Figure 11.3). Figure 11.3 plots the number of drop-outs that remained during each specific question. Initially, the number of respondents strongly decreased after the first few questions. After question 6, already more than 50 % of the 449 drop-outs had cancelled the survey. The first question block contained an introductory part followed by questions about telephony behavior. The following major set of drop outs occurred in discrete steps after individual question blocks. For example, between question block 1 and 2 (questions 12 and 13) the number of remaining drop-outs further decreased by 12 % but remained constant until question block 2 ended. After block 2 (question 27), the number of remaining drop-outs decreased in a single step of about 13 % to 19 %. These drop-outs participated in question block 3 (questions 28 - 31), which addresses procrastination. The stepwise decrease in the number of drop-outs continued after the procrastination question block: Before question block 4, the number of remaining drop-outs further decreased by 9 % followed by a further

[3] (1) “I put off the beginning of tasks until the last moment.“ (2) “I postpone carrying out activities.” (3) While handling a task, I realise that I could have done it much earlier.” (4) “I start off a task only once I’m under pressure”.

[4] 1 = almost never; 2 = rarely; 3 = sometimes; 4 = frequently; 5 = almost always.

decrease after this question block by 7 %. The drop-outs that remained at question 45 (~ 5 %) continuously decreased to zero towards the end of the questionnaire. In conclusion, respondents were most prone to abort the survey right at the beginning of the questionnaire or after they finished a question block.

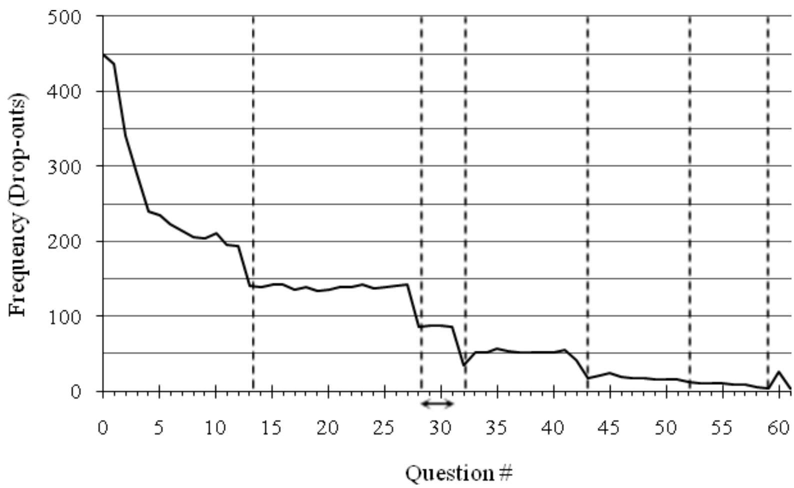


Fig. 11.3 Frequency of the remaining drop-outs per question. Source: CELLA 2.

The first question of each question block is indicated by a dashed line. Note that filter questions were omitted. The double arrow shows the position of the procrastination items within the questionnaire. The increase in frequency at question 60 is due to the fact that this question was placed at the beginning of the questionnaire during the first week of the survey period.

After addressing the characteristics of drop-outs, we now compare the procrastination tendency of the drop-outs and participants. For this purpose, the values per item concerning procrastination behavior mentioned by the respondents must be summed up (the aggregate values range from 4 to 20).

The mean is 9.1, the median is 9 and the most frequent aggregate value is 8 (12.6 %). Regarding the rankings it seems that persons chose rather lower frequencies in view of their procrastination behavior during the last two weeks (see Table 11.10). This is also evident from the fact that only a small proportion (0.7 %) answered all four questions with „almost always" (= 5). It is conspicuous that all of them are participants who completed the interview. In contrast, drop-outs reached lower aggregate values. Nearly 90 % of them show an aggregate value of less than or equal to 12, whereas the proportion of participants with such values is only about 80 %. The aggregate values of the remaining 20 % (and 10 % of drop-outs) range from 13 to 20. This corroborates the assumption that persons who complete the whole interview are more likely to procrastinate than persons who do not complete the interview.

Table 11.10 The tendency to procrastinate according to the level of interview completion: completed (participants) versus broken off (drop-outs)

Tendency to procrastinate		*Participants*			*Drop-outs*		
	Σ	Frequency	%	Cumulative Percent	Frequency	%	Cumulative Percent
almost never	4	285	9.6	9.6	4	5.8	5.8
	5	215	7.2	16.8	9	13.0	18.8
	6	308	10.4	27.2	8	11.6	30.4
	7	290	9.8	37.0	7	10.1	40.5
rarely	8	376	12.6	49.6	8	11.6	52.1
	9	274	9.2	58.8	9	13.0	65.1
	10	277	9.3	68.1	5	7.2	72.3
	11	224	7.5	75.6	6	8.7	81.0
sometimes	12	185	6.2	81.8	6	8.7	89.7
	13	138	4.6	86.4	2	2.9	92.6
	14	133	4.5	90.9	2	2.9	95.5
	15	79	2.7	93.6	0	.0	95.5
frequently	16	100	3.4	97.0	2	2.9	98.4
	17	28	0.9	97.9	0	.0	98.4
	18	25	0.8	98.7	0	.0	98.4
	19	15	0.5	99.2	1	1.4	100.0
almost always	20	21	0.7	100.0	0	.0	100.0
Total		2,973	100	100	69	100	100

Source: CELLA 2

Nevertheless, it has to be emphasized again that the number of drop-outs is rather low and we do not want to over-interpret the results. However, in the situation of total absence of information on nonrespondents, we have to follow the slightest trace. Overall, the results show that there are differences between persons who complete telephone interviews and persons who drop out. On this basis one might assume that there are also differences between respondents and nonrespondents. Consequently, a biased sample might be expected. Nevertheless, it is possible to correct for this error by means of adjustment weighting.

11.3.2 GREG-Estimator: Theoretical Background

The HT-estimator can be used as design unbiased estimator if all units in the gross sample answered the questions of the survey. Unfortunately, we are often in the uncomfortable situation that as consequence of nonresponse the values for the variable of interest only are available for a (sometimes small) part of the gross sample: the respondents.

However, often the values of auxiliary variables are available from registers or other sources for all units of the population or at least for the sample. A well-known way of adjusting the HT-estimator for nonresponse is the generalized regression estimator $\sum_{i\in r} g_i \frac{y_i}{\pi_i}$, where r is the subset of respondents and

$$g_i = \frac{1}{q_i}\left(1 + c_i\left(\sum_{k=1}^{N}\mathbf{x}_k - \sum_{k\in r}\frac{1}{\pi_k q_k}\mathbf{x}_k\right)'\left(\sum_{k\in r}\frac{c_k}{\pi_k q_k}\mathbf{x}_k\mathbf{x}_k{}'\right)^{-1}\right)\mathbf{x}_i.$$

x_k denotes a vector of auxiliary variables for unit k, π_k is the inclusion probability of unit k, q_k denotes the response probability of unit k and c_k is a fixed positive number which is often set to 1.

11.3.3 Practical Application

In the setup of the dual frame approach we present different estimates for the proportions of several variables to observe the consequences of incorporating or neglecting additional information in the estimates. The first estimator is the sample mean, the second the HT-estimator using design weights, and the third is the GREG-estimator. The prerequisite for using the GREG-estimator is the availability of auxiliary variables. In the CELLA 2 study we used gender, age classes, and education as adjustment variables. For some variables of interest we can compare the estimates with the results from the German Microcensus. The first variable we consider is gender (see Table 11.11).

Table 11.11 Different estimators for gender in CELLA 2 and Microcensus

Gender	Sample Mean	HT	GREG	Microcensus
Male	51.9	50.4	48.8	48.8
Female	48.1	49.6	51.2	51.2
Total	100	100	100	100

Typically, the majority of the respondents is male. To improve the estimates it somewhat helps to take the design weights into account. The GREG-estimator provides the same result as the Microcensus since gender is an adjustment variable.

The next example shows that people of lower secondary education are strongly underrepresented while the respondents who completed the Abitur are overrepresented in the sample (see Table 11.12).

Table 11.12 Different estimators for education in CELLA 2 and Microcensus

Education	Sample Mean	HT	GREG	Microcensus
Lower secondary level	18.5	20.7	38.8	38.8
Secondary level	27.1	26.9	21.8	21.8
Secondary level (former GDR)	5.9	6.2	6.7	6.7
Qualified for universities of applied sciences	8.3	7.6	6.0	6.0
Abitur; qualified for all universities	35.1	33.0	19.7	19.7
Other	5.2	5.6	7.0	7.0
Total	100	100	100	100

The table above shows very clearly that nonresponse may cause bias if we neglect weighting and calibration.

In the third example the estimation of the proportions of a non-adjustment variable is given (see Table 11.13).

Table 11.13 Different estimators of household size in CELLA 2 and Microcensus

Household size (persons living in the household)	Sample Mean	HT	GREG	Microcensus
1	24.6	21.0	23.9	22.3
2	34.1	35.0	38.8	38.2
3	18.3	18.9	17.7	18.0
4	14.6	15.3	11.1	15.1
5+	8.4	9.8	8.5	6.5
Total	100	100	100	100

In this case the differences in the estimates are not as dramatic as it is in the next example (see Table 11.14).

Table 11.14 Different estimators of marital status in CELLA 2 and Microcensus

Marital status	Sample Mean	HT	GREG	Microcensus
Married, living together	43.6	46.2	47.3	52.8
Civil homosexual partnership	0.1	0.1	0.1	0.0
Married, living apart together	2.6	2.2	2.5	2.0
Single	38.4	36.9	29.5	29.7
Divorced	9.9	9.1	10.9	7.5
Widowed	4.9	5.0	9.3	8.0
n. a.	0.5	0.5	0.4	0.0
Total	100	100	100	100

The sample mean underestimates the proportion of married people and overestimates the proportion of singles by nearly 10 percentage points. Again in most cases the GREG-estimator provides better results than the HT-estimator except for the proportion of divorced people.

11.4 Summary

Nonresponse in telephone surveys is dramatically increasing in Germany. By now response rates of between 10 % and 20 % are normal (see the chapter of Schneiderat and Schlinzig in this volume). However, the question of who participates in a survey is difficult to answer. Several theories try to explain the varying participation behavior. For instance, participation may be considered a rational decision (Esser 1986). This means that a person who can be contacted by phone might select the course of action that offers maximum benefit. But in what way may a person benefit from participating in a survey? The literature lists many, ranging from the feeling of doing a good deed to receiving incentives (Häder 2010: 194). Furthermore, it is conceivable that respondents decide to participate instead of completing other tasks, i.e. they procrastinate by participating in surveys. Hence, one might assume that those persons who decide to take part in an interview differ from those who reject the interview. According to this, a biased sample is to be expected. However, because of a lack of data on non-participants this assumption is hard to examine. Nevertheless, data about drop-outs provide first indications. In the context of CELLA 2 both participants and drop-outs tended to procrastinate. Although the number of drop-outs was rather small we found indications of differential procrastination behavior amongst both subpopulations. If drop-outs can be considered proxies for nonrespondents, this points toward a biased sample. These findings motivated the adjustment weighting introduced in this chapter.

However, despite the falling response rates telephone surveys are still a frequently used form of data collection in empirical social research in Germany. In 2010, for example, about 35 % of all interviews were conducted via the telephone[5]. However, in recent years this form of data collection faces a new challenge: the number of households with a fixed phone line is declining. Moreover, roughly every tenth household can only be reached by mobile phone. In contrast, 3 - 5 % of all households can solely be reached by landline. Therefore, both groups have to be considered in representative surveys to avoid sampling biases. The Dual Frame Approach combines landline and mobile phone surveys and thereby takes account of these differences. The theoretical background and an empirical example are shown in section 11.2.

Finally, in section 11.3 we explain how design weighting and an additional adjustment weighting can be combined in the GREG-estimator. We demonstrate that weighted estimates of socio-demographic variables are frequently closer to the Microcensus distributions than unweighted estimates. Thus, most of the important biases in telephone samples can be corrected if adequate weighting is applied.

Acknowledgment. We wish to thank Matthias Ganninger (GESIS) very much for his help with the figures.

[5] See www.adm-ev.de.

References

Aitken, M.E.: A personality profile of the college student procrastinator (Doctoral dissertation, University of Pittsburgh, 1982). Dissertation Abstracts International 43, 722 (1982)

Busse, B., Fuchs, M.: One mobile phone = one person? – Findings from two studies addressing consequences of cell phone sharing for sampling in mobile phone surveys. In: Ntts 2011 Conference papers (2011)

Callegaro, M., Ayhan, Ö., Gabler, S., Häder, S., Villar, A.: Combining landline and mobile phone samples: a dual frame approach. GESIS Working Paper 2011/13 (2011)

Dehne, M., Schupp, J.: Persönlichkeitsmerkmale im Sozio-oekonomischen Panel (SOEP) – Konzept, Umsetzung und empirische Eigenschaften (Personality in the Socio-economic Panel – Concept, Application and empirical characteristics). DIW Research Notes 26 (2007)

Esser, H.: Über die Teilnahme an Befragungen (On the participation in surveys). ZUMA-Nachrichten Heft 18, 38–47 (1986)

Gabler, S., Ayhan, Ö.: Gewichtung bei Erhebungen im Festnetz und über Mobilfunk: Ein Dual-Frame Ansatz (Weighting landline and mobile phone survey data. A dual-frame approach). In: Gabler, S., Häder, S. (eds.) Mobilfunktelefonie - Eine Herausforderung für die Umfrageforschung, pp. 39–45. GESIS, ZUMA-Nachrichten Spezial, Band 13, Mannheim (2007)

Glöckner-Rist, A., Engberding, M., Höcker, A., Rist, F.: Prokrastinationsfragebogen für Studierende (PFS) (A questionnaire for the measurement of procrastination for students). In: Glöckner-Rist, A. (ed.) Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS): ZIS 14.00, pp. 1867–5972. GESIS, Bonn (2010)

Häder, M.: Empirische Sozialforschung (Empirical Social Research). VS Verlag für Sozialwissenschaften, Wiesbaden (2010)

Häder, M., Häder, S.: Der Einfluss der Persönlichkeit auf die Antwortqualität – empirische Befunde aus der Umfrageforschung über das Festnetz und den Mobilfunk (The influence of personality on the quality of survey responses – empirical findings from survey research via landline and mobile phones). In: Submitted to Kölner Zeitschrift für Soziologie und Sozialpsychologie (2011)

Häder, M., Häder, S.: Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung (Telephone interviews via mobile phones. Concept, design and implementation of a strategy for data collection). VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

John, O.P., Naumann, L.P., Soto, C.J.: Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In: John, O.P., et al. (eds.) Handbook of Personality: Theory and Research, pp. 114–158. Guilford Press, New York (2008)

Meier, G.: Validierung eines Fragebogens zur Erfassung der Anzahl der Telefonnummern (Validation of a questionnaire for the recording of the quantity of telephone numbers). In: Gabler, S., Häder, S. (eds.) Mobilfunktelefonie - Eine Herausforderung für die Umfrageforschung, pp. 91–104. GESIS, ZUMA-Nachrichten Spezial, Band 13, Mannheim (2007)

Rao, J.N.K., Scott, A.J.: On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated from Survey Data. The Annals of Statistics 12, 46–60 (1984)

Rist, F., Engberding, M., Patzelt, J., Beißner, J.: Aber morgen fange ich richtig an! Prokrastination als verbreitete Arbeitsstörung (But tomorrow I'll really start! Procrastination as a common labor dysfunction). Personalführung 6, 64–78 (2006)

Statistisches Bundesamt Deutschland, Bevölkerung nach Altersgruppen, Familienstand und Religionszugehörigkeit (Population by age, marital status and religious affiliation), http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Bevoelkerung/Bevoelkerungsstand/Tabellen/Content75/AltersgruppenFamilienstand,templateId=renderPrint.psml (Accessed May 2011)

Chapter 12
Weighting to Adjust for Non-observation Errors in Telephone Surveys

Fannie Cobben, Barry Schouten, and Jelke Bethlehem

12.1 Introduction

For its social and demographic surveys, Statistics Netherlands favours CAPI over cheaper modes that employ web or telephone. Due to the persuasive power and assistance of interviewers visiting selected persons or households, nonresponse in CAPI surveys is relatively low and data quality is high. However, the costs of this mode of interviewing are relatively high. A large group of trained interviewers is required. Although the Netherlands is a small country, travel costs make up a considerable proportion of the total costs.

To reduce costs, Statistics Netherlands is changing some of its CAPI surveys into CATI and web surveys. Here, we focus on CATI. By concentrating interviewers in one call centre, a smaller number of interviewers is sufficient. No more time is spent on travel, and this also means no more travel costs are involved. Although a possible change from CAPI to CATI may substantially reduce the costs of surveys, there is also a potential drawback: it may reduce the quality of the produced statistics.

There is a vast amount of literature that compares telephone surveys to face-to-face surveys on different aspects. De Leeuw (1992) performed a meta-analysis of face-to-face and telephone surveys and considered several aspects of data quality. She concluded that differences in data quality between well-conducted face-to-face and telephone surveys are small. This conclusion is in line with Groves (1989), who states that the most consistent finding in studies comparing responses in face-to-face and telephone surveys is the lack of difference between the two modes.

In this chapter we focus on non-sampling errors caused by nonresponse and coverage . We assume that measurement errors are dealt with. Also, we regard the situation where telephone numbers are linked to a sample drawn from a population register since this is current practice at Statistics Netherlands. The sample for a CATI survey is obtained by matching the sample from the population register to the Dutch telephone company KPN for listed telephone numbers. However, links will only be established for sample elements with a listed landline telephone. Currently, the percentage of persons with a listed landline telephone is estimated to be between 60% and 70%. This means that there is a substantial undercoverage of 30% to 40%. For more information on the current practice for telephone surveys at Statistics Netherlands, see the country spotlight by Beukenhorst.

The group without a listed landline telephone actually consists of three categories: persons that have an unlisted landline telephone, persons with only a mobile telephone, and persons that do not have a telephone at all. In our analysis, we cannot distinguish these groups although they are likely to be very different on socioeconomic variables (see e.g. Vehovar et al., 2004 and Callegaro and Poggio, 2004).

The differences between persons with and without a listed landline telephone have been analysed extensively. Cobben and Bethlehem (2005) find an under representation of non-Western, non-native persons and regions where the percentage of non-natives is higher than 20%. Pickery and Carton (2005) analyse the representativeness of telephone surveys in Flanders. They find the same differences in ethnic group as Cobben and Bethlehem (2005). Furthermore, they report differences related to the educational level, the employment situation and ownership of a house. Van Goor and Rispens (2004) find differences in the Netherlands with regard to the ethnic group, the household type and the employment situation. In Finland, Kuusela (2000) finds a difference between persons with and without a listed landline telephone in the degree of urbanization and in the size and type of the household. Ellis and Krosnick (1999) report differences in telephone ownership with respect to education, income and ethnicity.

In the Netherlands, the coverage of listed landline telephones decreased to such an extent that it seriously threatens CATI as a single data collection mode, regardless of other errors. This motivates the two objectives of this chapter: first, we want to assess what the effect of a change from CAPI to CATI is on the estimation of population characteristics. As we will see in section 12.4.1, the estimates based on a CATI survey are different from the estimates in a CAPI survey. Therefore the second objective is to investigate if and how we can adjust for this effect. We thereby focus on coverage errors and nonresponse bias. We introduce several weighting methods based on the generalized regression estimator and on the use of propensity scores and show how these methods can be applied to adjust for coverage errors and nonresponse bias. The data for the analysis come from the Dutch Survey on Living Conditions, denoted by its acronym SLC. This is a CAPI survey. We artificially construct a CATI survey from the CAPI survey by removing persons without a listed landline telephone. By doing so we can analyse directly the impact of undercoverage and we do not have any mode effects in response behavior . Both the response rate and the composition of the response may be different in telephone surveys and face-to-face surveys. We implicitly assume that the CATI response behavior will not be very different from the CAPI response behavior. The outline of this chapter is as follows: in section 12.2 we describe the data from the SLC 2002 survey. In section 12.3 we shortly outline the methods that we applied to the data to answer the research questions. Section 12.4 presents the results from the analysis and section 12.5 concludes.

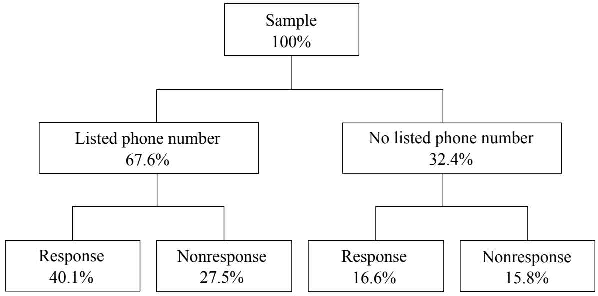


Fig. 12.1 Graphical representation of undercoverage and nonresponse in SLC 2002

12.2 Description of the Data

The data that we use in the analysis is obtained by aggregating the monthly SLC-surveys for the year 2002. SLC is a continuous CAPI survey. Every month a sample of $3,000$ persons is selected and interviewed face-to-face. The survey has a modular structure; there is a base module with questions for all sampled persons and in addition there are a number of modules about specific themes (such as employment situation, health and justice). The sampled persons are selected for one of the thematic modules; the questions in the base module are answered by everyone.

The target population is not the same for every module. However, all target populations consist of consist of persons aged 12 or older. Persons are selected by means of a stratified two-stage sample . In the first stage, municipalities are selected within regional strata with probabilities proportional to the number of inhabitants. In the second stage, an equal probability sample is drawn from each of the selected municipalities (but with different selection probabilities). In this chapter, only persons aged 12 or older are regarded. These persons all have the same first-order inclusion probability. This research focuses on the questions in the base module.

It is difficult to distinguish the contributions of coverage errors and nonresponse bias in the total survey error. Figure 12.1 describes the situation graphically for the SLC 2002 survey. In the ideal situation where the sampling frame exactly covers the population, the bias will only be caused by nonresponse of persons both with and without a listed telephone. CAPI is close to this situation. The overall response rate is 40.1% + 16.6% = 56.7%. (See Luiten (2011) for a description of response rate calculations at Statistics Netherlands.) In case of a CATI survey, the bias is caused both by undercoverage and nonresponse. Only 40.1% of the original sample will respond.

Note that the response rate among the listed telephones is much higher (59.4%) than for persons who do not have a listed telephone (51.1%). Apparently, persons without a listed telephone behave differently from persons with a listed telephone. This is also found by Cobben (2009).

As we already mentioned, the SLC is a CAPI survey. To answer our research questions we need both a CAPI- and a CATI survey. We therefore construct the CATI survey from the CAPI survey. We can do so by matching the sample elements in the

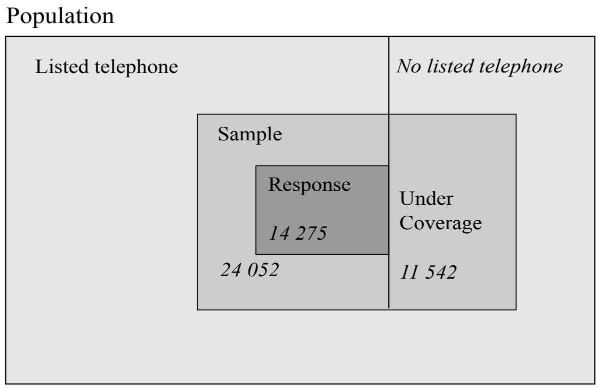


Fig. 12.2 Graphical representation of the division of the population with respect to telephone ownership

CAPI survey to the telephone register provided by the Dutch telephone company KPN. Deleting sample elements without a listed landline telephone provides us with the sample that would have been used had the survey been performed by CATI. An advantage of this artificial way of generating the CATI survey is that possible mode-effects caused by differences in face-to-face and telephone interviewing are avoided.

The CAPI survey sample consists of 35,594 sample elements, 24,052 of which have a listed landline telephone (67.6%). Figure 12.2 graphically displays the population and the two datasets.

There are two types of variables that we use in the analysis: auxiliary variables and survey items. Auxiliary variables are available for both respondents and nonrespondents. These variables come from registers like the population register and the Centre for Work and Income (CWI). The survey items are the answers to the survey questions; these are only available for respondents. The variables that we use in our analysis are displayed in table 12.1.

Table 12.1 Auxiliary variables and survey items in SLC 2002 survey

Auxiliary variable	*Categories*
Gender	Male, Female
Age_3	12–34, 35–54, 55 +
Age_{15}	12–14, 15–17, ..., 70–74, 74 +
Marital status_2	Married, Not married
Marital status_4	Married, Not married, Divorced, Widowed
Ethnic group	Native, Moroccan, Turkish, Suriname, Netherlands Antilles/Aruba, other non-Western non-native, other Western non-native
Region_{15}	Province of residence and three largest cities
Region_4	North, East, South, West
Degree of urbanization	Very low, Low, Average, High, Very high

Continued on next page

Table 12.1 Auxiliary variables and survey items in SLC 2002 survey - continued

Auxiliary variable	*Categories*
Household size	1, 2, 3, 4, 5+
Household type	Single, Couple, Couple with children, Single parent, Other
Interview month	January, February, ... , December
Listed landline telephone	Yes, No
Disability insurance	Yes, No
Social security	Yes, No
Average house value	Missing, 0, 0–50,000, 50,000–75,000, ... , 275,000–300,000, 300,000–350,000, > 350,000 (euro)
% non-natives in 6-digit postcode area	0–5%, 5–10%, ..., 40–50%, 50% and more
Survey item	*Categories*
Employment status	12 hours or more, less than 12 hours, unemployed
Educational level	Primary, Junior general secondary, Pre-vocational, Senior general secondary, Secondary vocational, Higher professional, University, Other
Religion	None, Roman-Catholic, Protestant, Islamic, Other

12.3 The Methods

In section 12.3.1 we describe methods to adjust for undercoverage errors. Section 12.3.2 describes methods that adjust for both undercoverage errors and nonresponse bias. These methods will all be applied to the CATI response of the SLC 2002 survey.

12.3.1 Telephone Coverage Propensity

Let the target population of the survey consist of N sample elements, $i = 1, 2, \ldots, N$. Furthermore, let Y denote one of the survey items described in table 12.1. The aim of the survey is to estimate the population mean

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{12.1}$$

In addition, we denote by $\mathbf{X}$ a vector of auxiliary variables with values $\mathbf{X}_i$ for $i = 1, 2, \ldots, N$. The sample is selected without replacement from the population and can be represented by the N-vector $\delta = (\delta_1, \delta_2, \ldots, \delta_N)'$ of sample indicators, ($\delta_i = 1$) when sample element i is selected in the sample, and ($\delta_i = 0$) otherwise. The expected value of δ is $E(\delta) = \pi$ where $\pi = (\pi_1, \pi_2, \ldots, \pi_N)'$ is the vector of the first order inclusion probabilities. We denote by $d_i = 1/\pi_i$ the design weights. We assume that the values for π are known and nonzero. We can then estimate (12.1) without bias by the Horvitz-Thompson estimator , i.e.

$$\bar{y}_{ht} = \frac{1}{N}\sum_{i=1}^{N}\delta_i d_i Y_i \tag{12.2}$$

Now, let us assume that having a listed landline telephone is the result of a random process. Each element i has a certain unknown probability τ_i of having a listed telephone, for $i = 1, 2, \ldots, N$. Let T denote an indicator, and $T_i = 1$ when element i is sampled and has a listed telephone, $T_i = 0$ if sample element i does not have a listed telephone. Now, the *telephone coverage propensity* is defined as

$$\tau(\mathbf{X}_i^t) = P(T_i = 1|\delta_i = 1, \mathbf{X}_i^t) \tag{12.3}$$

where $\mathbf{X}^t = (\mathbf{X}_1^t, \mathbf{X}_2^t, \ldots, \mathbf{X}_n^t)'$ is the vector of auxiliary variables that is used to estimate the telephone coverage probability τ_i. The values τ_i are unknown. We can estimate τ_i by an appropriate method and the available auxiliary variables, thus obtaining the telephone coverage propensity $\hat{\tau}_i = \tau(\mathbf{X}_i^t)$. Some of the methods that we can use for this purpose are described in the appendix.

The telephone coverage propensities can be used to adjust for undercoverage. To adjust for errors due to undercoverage in the CATI survey, we present two estimators: the propensity weighting estimator (12.4) and the propensity stratification estimator (12.6). The telephone coverage propensity weighting estimator is an adopted version of the Horvitz-Thompson estimator for telephone coverage and can be described as

$$\bar{y}_{ht}^t = \frac{1}{N}\sum_{i\in r}\frac{d_i Y_i}{\tau(\mathbf{X}_i^t)} \tag{12.4}$$

This technique is used by Kalton and Flores-Cervantes (2003) to adjust for nonresponse bias, and Särndal (1981) and Little (1986) have also applied this approach. To obtain the propensity stratification estimator, we stratify the sample based on the telephone coverage propensity score. Then, within strata that have the same value of $\tau(\mathbf{X}^t)$, $P(T_i = 1)$ does not depend on $\mathbf{X}^t$. We thus obtain F strata based on the telephone coverage propensities. Cochran (1968) suggests that it is enough to use five strata, i.e. $F = 5$. The strata are denoted by $s_1, s_2, \ldots, s_5$.

We introduce five dummy variables $X_1, X_2, \ldots, X_5$. For an element in a certain stratum f, the corresponding dummy variable X_f is assigned the value 1, and all other dummy variables are set to 0. The sample size of stratum f is equal to n_f. These sample sizes are random variables and not fixed numbers. Post-stratification assigns the same weight to all elements in the same stratum. The correction weight g_i for element i in stratum f is defined as

$$g_i = \frac{n_f}{n_{r,f}} \tag{12.5}$$

where $n_{r,f}$ is the number of responding sample elements in stratum f. Consequently, the telephone coverage propensity stratification estimator can be expressed as

$$\begin{aligned}\overline{y}_{ps}^{\tau(\mathbf{X}_t)} &= \tfrac{1}{N}\left(\tfrac{n_1}{n_{r,1}}\sum_{i\in s_1} Y_i + \tfrac{n_2}{n_{r,2}}\sum_{i\in s_2} Y_i + \ldots + \tfrac{n_F}{n_{r,F}}\sum_{i\in s_F} Y_i\right)\\ &= \tfrac{1}{N}\sum_{f=1}^{F} n_{r,f}\overline{y}_t^{(f)}\end{aligned} \tag{12.6}$$

where $\overline{y}_t^{(f)}$ is the adopted Horvitz-Thompson estimator for telephone coverage for the survey item in stratum f.

12.3.2 Simultaneous Adjustment of Undercoverage and Nonresponse

Estimators (12.4) and (12.6) adjust for the errors caused by the undercoverage of persons without a listed landline telephone. The question of how we can combine the adjustment for nonresponse and undercoverage now arises. We use the GREG-estimator to adjust for nonresponse. We need to add the adjustment for undercoverage to this method. There are two ways to combine linear weighting and the telephone coverage propensity score method. The first approach is based on the idea of propensity score weighting and comes down to GREG-estimation with adjusted inclusion probabilities. The second approach employs telephone coverage as an additional auxiliary variable in the weighting model for the GREG-estimator. We describe these two approaches separately. We distinguish between auxiliary variables that are used in estimating the telephone coverage propensity denoted by $\mathbf{X}^t$ and the variables that we use to adjust for nonresponse, which we will denote by $\mathbf{X}^r$.

The GREG-estimator can be expressed as

$$\overline{y}_{gr}^{r} = \sum_{i=1}^{n_r} d_i g_i Y_i \tag{12.7}$$

where g_i is the correction weight

$$g_i = 1 + \left(\overline{\mathbf{X}}^r - \overline{\mathbf{x}}_{r,ht}^r\right)'\left(\sum_{i=1}^{n_r} d_i \mathbf{X}_i^r (\mathbf{X}_i^r)'\right)^{-1}\mathbf{X}_i^r \tag{12.8}$$

This estimator only produces consistent estimates if the proper design weights d_i are used. Availability of data in a CATI survey is determined by both the sampling mechanism and the probability of having a listed telephone. Therefore, the design weights d_i should be divided by τ_i and hence be replaced by d_i/τ_i. Unfortunately, the τ_i are unknown, so they have to be estimated by $\hat{\tau}_i = \tau(\mathbf{X}_i^t)$. The estimator in (12.7) then can be expressed as

$$\overline{y}_{gr}^{r} = \sum_{i=1}^{n_r} \frac{d_i}{\hat{\tau}_i} g_{i\hat{\tau}} Y_i \tag{12.9}$$

where $g_{i\hat{\tau}}$ is given by

$$g_{i\hat{\tau}} = 1 + \left(\overline{\mathbf{X}}^r - \overline{\mathbf{x}}_{r,ht}^t\right)'\left(\sum_{i=1}^{n_r} \frac{d_i}{\hat{\tau}_i}\mathbf{X}_i^r (\mathbf{X}_i^r)'\right)^{-1}\mathbf{X}_i^r \tag{12.10}$$

where $\bar{x}^t_{r,ht}$ is the adopted Horvitz-Thompson estimator for telephone coverage for the mean of the auxiliary variables $\mathbf{X}^r_i$. This method is referred to as the *telephone coverage propensity GREG-estimator*.

The other possibility is to include the telephone coverage propensity stratification variable in the weighting model. For the direct telephone coverage propensity stratification estimator we already divided the sample into five strata based on the telephone coverage propensities by introducing the dummy variables $X_1, X_2, \ldots, X_5$. Now, we aggregate these variables into one ordered categorical variable $Telprop_5$ with five classes. Sample elements belong to class f if $X_f = 1$. This method is referred to as the *GREG-estimator with telephone coverage propensity variable*. In the following section we describe the analysis.

12.4 Analysis of the Telephone SLC Survey

When changing the survey interview mode from CAPI to CATI, there are two sources of missing data: nonresponse and undercoverage of sample elements with unlisted telephones, mobile only, or no telephone at all. Similar to nonresponse, undercoverage is only a threat to survey statistics if it is selective. Cobben and Bethlehem (2005) show that having a listed landline telephone is especially selective with respect to the variables ethnic group, % non-natives, and household type.

To obtain insight into the consequences of undercoverage we apply the methods discussed in the previous sections to our data sets. First, we compare the response means of both the CAPI survey and the CATI survey. For this we apply the GREG-estimator based on the CAPI response, that is proposed by Schouten (2004). The variables used in the estimation are

$$\begin{aligned} &Age_{15} + Housevalue_{14} + \%Non-Natives_8 \\ &+Ethnicgroup_7 + Region_{15} + Householdtype_4 + Telephone_2 \end{aligned} \tag{12.11}$$

The subscripts denote the number of categories. Model (12.11) is applied to both surveys. When applying the model to the CATI survey, the variable $Telephone_2$ becomes redundant. The CAPI survey estimates are used as a benchmark for judging the quality of the survey estimates in the CATI survey. For this we need the CAPI estimates with their best nonresponse adjustment. The weighting model is applied to the CATI survey to see whether the survey estimates are biased if we do not account for undercoverage.

Subsequently, we apply the direct telephone coverage propensity weighting estimator (12.4) and - stratification estimator (12.6) to the CATI survey to adjust for undercoverage. In case the CATI survey estimates and CAPI survey estimates are similar, switching from CAPI to CATI does not introduce an additional bias. Otherwise, the selective undercoverage in CATI cannot be ignored and has to be adjusted as well.

And finally, the combination of linear weighting for nonresponse and the telephone coverage propensity is applied to the CATI survey to simultaneously adjust for nonresponse and undercoverage. For the telephone coverage propensity

GREG-estimator described by (12.10) the variables for the nonresponse adjustment $\mathbf{X}^r$ are the same variables as in the GREG-estimator described in (12.11) applied to the CATI survey, thus without the variable $Telephone_2$. For the GREG-estimator with telephone coverage propensity variable the weighting model becomes

$$\begin{aligned}&Age_{15} + Housevalue_{14} + \%Non - Natives_8\\&+Ethnicgroup_7 + Region_{15} + Householdtype_4 + Telprop_5\end{aligned} \quad (12.12)$$

The results for the nonresponse adjustment, the undercoverage adjustment, and the simultaneous adjustment are discussed in the following sections.

12.4.1 Nonresponse Adjustment

In table 12.2 we show the response means and the adjusted estimates when applying the nonresponse adjustment; weighting model (12.11).

Table 12.2 Comparison of unweighted and weighted response means for the CAPI and the CATI survey (in %)

Variable	*Response mean CATI*	*Response mean CAPI*	*GREG CATI*	*GREG CAPI*
Employment				
12 hours or more	51.8 (0.42)	52.4 (0.35)	54.4 (0.27)	53.7 (0.18)
Unemployed	7.0 (0.22)	6.7 (0.18)	6.6 (0.16)	6.4 (0.11)
Less than 12 hours	41.2 (0.41)	40.9 (0.35)	38.9 (0.25)	39.9 (0.17)
Education				
Primary	6.8 (0.21)	7.2 (0.18)	6.1 (0.13)	6.3 (0.09)
Junior general secondary	12.2 (0.27)	12.1 (0.23)	11.9 (0.22)	12.0 (0.15)
Pre-vocational	19.6 (0.33)	19.7 (0.28)	19.6 (0.27)	19.9 (0.18)
Senior general secondary	6.8 (0.21)	7.1 (0.18)	7.1 (0.18)	7.2 (0.12)
Secondary vocational	30.8 (0.39)	30.8 (0.32)	31.2 (0.31)	31.1 (0.21)
Higher professional	17.3 (0.32)	16.7 (0.26)	17.5 (0.26)	16.9 (0.17)
University	6.3 (0.20)	6.3 (0.17)	6.5 (0.17)	6.4 (0.11)
Other	0.2 (0.04)	0.2 (0.03)	0.2 (0.04)	0.2 (0.02)
Religion				
None	36.3 (0.40)	37.7 (0.34)	37.8 (0.32)	38.5 (0.21)
Roman-Catholic	35.4 (0.40)	33.5 (0.33)	32.7 (0.28)	32.4 (0.19)
Protestant	22.7 (0.35)	21.0 (0.29)	21.0 (0.25)	20.4 (0.17)
Islamic	1.2 (0.09)	2.5 (0.11)	3.2 (0.10)	3.3 (0.06)
Other	4.5 (0.17)	5.2 (0.16)	5.2 (0.16)	5.4 (0.01)

The standard errors are given in parentheses. For the response means, the standard errors are calculated by $\sqrt{p(1-p)/n}$, with p the percentage of the survey item category and n the number of respondents ($n = 20,168$ for CAPI and $n = 14,275$ for CATI). For the weighted estimates in columns 4 and 5, the standard errors are calculated by a first order Taylor approximation of the linear regression estimator.

The second and third column in table 12.2 display the unweighted response means for both surveys. There are quite some differences. Respondents with a listed, landline telephone tend to work less, have a higher education and are more often non-religious than respondents without a listed landline telephone. However, it is possible that adjusting for nonresponse also handles the selectivity of telephone coverage and reduces these differences. To see if this is indeed so we apply weighting model (12.11) to both surveys. The results are displayed in columns four and five in table 12.2.

To assess the quality of the adjusted estimates based on CATI, we compare these with the estimates from the CAPI survey. We assume that the results from weighting the CAPI survey are unbiased, i.e. equal to the true population characteristics.

Table 12.2 shows that the estimates from CATI and CAPI are different. The CATI survey estimates tend to be adjusted in the right direction, but then over- or underestimate the characteristic when compared to the CAPI survey estimates. For instance the response mean of the percentage of persons that work 12 hours or more for the telephone survey is 51.8%. The CAPI survey estimate is 53.7%. The CATI survey estimate indeed increases the response mean but overestimates the CAPI survey estimate by 0.7%. The differences may seem minor, but 0.7% of 14 million persons (the population aged 12+) is still approximately 100,000 persons.

These differences are remarkable since model (12.11) incorporates the variables that cause the largest selectivity (Ethnic group, % non-natives and Region). The results suggest that on average answers by sample elements with a listed telephone are different from answers by sample elements without a listed telephone, even when accounting for the variables that are correlated with telephone ownership. We thus have to account for the selective undercoverage in the CATI survey. Therefore, we applied the adjustment methods using the telephone coverage propensity discussed in section 12.3.1.

12.4.2 Undercoverage Adjustment

We model the telephone coverage propensities by means of a logit model. Other models can be used too, but Dehija and Wahba (1999) for example conclude that different models often produce similar results. In the appendix more information about estimating these probabilities is provided. Using the CAPI survey sample, the propensities are modelled with the software package Stata. By stepwise excluding insignificant variables, the variables $\mathbf{X}^t$ in the final model are

$$\begin{aligned}&\%Non-Natives_8 + Region_{15} + EthnicGroup_7 + Urbanisation_5 +\\ &MaritalStatus_2 + HouseholdType_4 + HouseValue_{14} + Age_3 +\\ &DisabilityAll_2 + SocialAll_2\end{aligned} \tag{12.13}$$

The subscripts denote the number of categories. We estimate the model parameters by Maximum Likelihood Estimation. The value of the pseudo R^2 for this model is 9.1%. This is rather low, which is an indication that there still is a lot of unexplained variance in this model. This small value for the pseudo R^2 possibly results from the

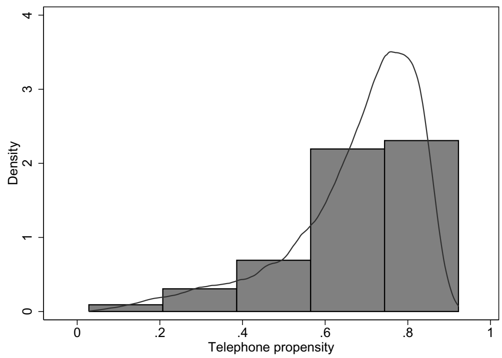


Fig. 12.3 Telephone coverage propensities divided into five strata.

fact that the persons without a listed landline telephone actually consist of three very different socio-demographic groups.

Based on this model the telephone coverage probabilities can be estimated. The telephone coverage propensity strata are displayed graphically in figure 12.3. In addition to the five strata, a kernel density line has been estimated. This represents the continuous distribution of the telephone coverage propensity in the sample. The higher the number of strata, the more the distribution of the strata will resemble the continuous distribution.

We then use the propensities $\hat{\tau}_i$ to adjust for errors due to undercoverage as described in section 12.3.1. The results are displayed in table 12.3.

The standard errors are given in parentheses. The first two columns are again the response mean for CATI and CAPI respectively, like in table 12.2. For the telephone coverage propensity methods in columns 4 and 5, the standard errors are calculated in R by non-parametric bootstrap estimation with $B = 1,000$ bootstraps.

Table 12.3 Estimates for the CATI survey based on telephone coverage propensity weighting and stratification

Variable	*Response mean CATI*	*Response mean CAPI*	*Propensity weighting*	*Propensity stratification*
Employment				
12 hours or more	51.8 (0.42)	52.4 (0.35)	52.2 (0.50)	52.1 (0.48)
Unemployed	7.0 (0.22)	6.7 (0.18)	7.0 (0.24)	7.0 (0.22)
Less than 12 hours	41.2 (0.41)	40.9 (0.35)	40.7 (0.49)	40.9 (0.45)
Education				

Continued on next page

Table 12.3 Estimates for the CATI survey based on telephone coverage propensity weighting and stratification - continued

Variable	*Response mean CATI*	*Response mean CAPI*	*Propensity weighting*	*Propensity stratification*
Primary	6.8 (0.21)	7.2 (0.18)	7.1 (0.24)	7.1 (0.21)
Junior general secondary	12.2 (0.27)	12.1 (0.23)	12.2 (0.32)	12.2 (0.28)
Pre-vocational	19.6 (0.33)	19.7 (0.28)	19.4 (0.39)	19.5 (0.34)
Senior general secondary	6.8 (0.21)	7.1 (0.18)	6.9 (0.25)	6.9 (0.22)
Secondary vocational	30.8 (0.39)	30.8 (0.32)	30.6 (0.44)	30.7 (0.40)
Higher professional	17.3 (0.32)	16.7 (0.26)	17.1 (0.36)	17.2 (0.33)
University	6.3 (0.20)	6.3 (0.17)	6.3 (0.23)	6.3 (0.20)
Other	0.2 (0.04)	0.2 (0.03)	0.2 (0.04)	0.2 (0.04)
Religion				
None	36.3 (0.40)	37.7 (0.34)	37.2 (0.46)	37.0 (0.43)
Roman-Catholic	35.4 (0.40)	33.5 (0.33)	33.5 (0.46)	33.8 (0.41)
Protestant	22.7 (0.35)	21.0 (0.29)	21.5 (0.39)	21.7 (0.37)
Islamic	1.2 (0.09)	2.5 (0.11)	2.7 (0.11)	2.5 (0.09)
Other	4.5 (0.17)	5.2 (0.16)	5.1 (0.21)	5.1 (0.17)

To see how the techniques perform, the results from telephone coverage propensity weighting and -stratification are compared to the response mean from the CAPI survey. In columns 4 and 5, the results from telephone coverage propensity weighting and stratification are shown. Both telephone coverage propensity weighting and stratification perform well. The estimates based on the CATI survey are very close to the CAPI response means. Stratification based on the telephone coverage propensities has slightly lower standard errors than weighting. The method appears to be successful in adjusting for the selective undercoverage in CATI surveys.

12.4.3 Simultaneous Adjustment

Now we can proceed to adjust the final estimates for nonresponse bias and coverage errors simultaneously. For this we apply the methods presented in section 12.3.2. The results are displayed in table 12.4.

The standard errors are given in parentheses and calculated by a first order Taylor approximation of the linear regression estimator. For the methods using the telephone coverage propensity in columns 3 and 4, we make the assumption that these propensities are deterministic values that introduce no additional variance.

The results are again compared to the estimates from the CAPI survey, displayed in the second column. The third and fourth column display the results from adjusting the CATI survey for nonresponse and undercoverage simultaneously with the GREG-estimator for nonresponse adjustment with an additional telephone coverage propensity variable and the GREG-estimator to adjust for nonresponse bias with adjusted inclusion probabilities for telephone coverage presented in section 12.3.2.

Table 12.4 Adjustment of the CATI survey estimates for both coverage errors and nonresponse bias

Variable	*GREG-estimator CAPI*	*GREG-estimator + propensity variable*	*Propensity GREG-estimator*
Employment			
12 hours or more	53.7 (0.18)	54.0 (0.27)	54.8 (0.35)
Unemployed	6.4 (0.11)	6.7 (0.16)	6.6 (0.20)
Less than 12 hours	39.9 (0.17)	39.3 (0.26)	38.7 (0.33)
Education			
Primary	6.3 (0.09)	6.1 (0.14)	6.1 (0.17)
Junior general secondary	12.0 (0.15)	11.9 (0.22)	11.9 (0.29)
Pre-vocational	19.9 (0.18)	19.6 (0.27)	19.6 (0.36)
Senior general secondary	7.2 (0.12)	7.1 (0.18)	7.0 (0.23)
Secondary vocational	31.1 (0.21)	31.3 (0.31)	31.3 (0.41)
Higher professional	16.9 (0.17)	17.5 (0.26)	17.5 (0.34)
University	6.4 (0.11)	6.5 (0.17)	6.4 (0.22)
Other	0.2 (0.02)	0.2 (0.04)	0.2 (0.04)
Religion			
None	38.5 (0.21)	37.8 (0.32)	37.9 (0.42)
Roman-Catholic	32.4 (0.19)	32.7 (0.28)	32.8 (0.37)
Protestant	20.4 (0.17)	21.0 (0.25)	21.1 (0.34)
Islamic	3.3 (0.06)	3.2 (0.11)	3.1 (0.13)
Other	5.4 (0.01)	5.3 (0.17)	5.2 (0.21)

Both methods seem to perform well. The standard errors for the combination of linear weighting and adjusted inclusion probabilities for telephone coverage are larger than for linear weighting including an additional telephone coverage propensity stratification variable. For the variables educational level and religion, the estimates for both methods are very similar. With respect to the employment situation, the estimates based on linear weighting with an additional telephone coverage propensity stratification variable are closer to the benchmark estimates from the CAPI survey. The best adjustment technique to reduce the bias caused by telephone interviewing for this case, appears to be linear weighting including an additional telephone coverage propensity stratification variable.

12.4.4 Summary

In this chapter we consider the influence of the data collection mode on errors related to undercoverage and nonresponse, and compare various techniques that aim at adjusting for the bias caused by these errors. These techniques are partly based on the GREG-estimator, and partly on using propensity scores. The aim of our research is to answer the two following questions: What is the effect on the quality of estimates of population characteristics, when changing a CAPI survey into a CATI survey and, consequently, can we adjust for this effect?

We explore to what extent adjustment techniques can reduce the bias caused by telephone interviewing. First, the CATI survey is adjusted for undercoverage of sample elements without a listed landline telephone. Two methods are used: Telephone coverage propensity weighting and -stratification. No nonresponse bias is considered yet and the results are compared to the unweighted response mean of the CAPI survey. Second, the nonresponse bias is taken into account as well. Two combinations of the GREG-estimator and the telephone coverage propensities are considered.

A comparison of the response means in the CATI- and the CAPI survey shows that indeed these two surveys differ in the population characteristics of interest. Respondents that own a listed landline telephone tend to work less, have a higher education and are more often non-religious than respondents that do not own a listed landline telephone. Ignoring the undercoverage in the CATI survey, i.e. adjusting for nonresponse only, does not take away all the bias. We use the telephone coverage propensities to adjust for undercoverage in the CATI survey. Both telephone coverage propensity weighting and -stratification perform well. The estimates based on the CATI survey are very close to the CAPI response means. Stratification based on the telephone coverage propensities has slightly lower standard errors than weighting. Both methods appear to be appropriate for adjusting for selective undercoverage in CATI surveys.

Subsequently, we simultaneously adjust for undercoverage and nonresponse bias. For this we applied two combinations of the GREG-estimator for nonresponse and the telephone coverage propensity score method. The best adjustment technique to reduce the bias caused by telephone interviewing for this case, appears to be the GREG-estimator with the inclusion of an additional telephone coverage propensity stratification variable. However, the ultimate estimates are still biased. There seems to be a relationship between telephone ownership and the questions in the survey that we cannot explain with the available auxiliary variables. We lack variables that are sufficiently informative to explain telephone ownership. This leads to biased estimates, especially for those survey questions that are related to education and income. The variables that we try to estimate are actually the variables that we would like to use in our model.

12.5 Concluding Remarks

The proportion of explained variance of the logit model that is used to estimate the telephone coverage propensities is only 9.1%. Despite this low level of explained variance, the telephone coverage propensity method is still able to adjust for selective undercoverage in the CATI survey. However, more research is needed to determine the importance of the model fit when using estimated propensities in the methods described in section 12.3.

The steady decrease in coverage of listed landline telephones raises the question of whether CATI, as conducted at Statistics Netherlands, is still a viable single data collection mode. The results of our research indicate that the omission of sample elements without a listed landline telephone implies a bias for certain survey topics, like income and education, that cannot be adjusted for sufficiently. However, in a mixed mode data collection design the undercoverage of persons without a telephone can be compensated by another mode and, hence, CATI can still be a very important mode in a mixed mode design. A necessary condition for mixed mode data collection is that the survey questionnaire is cognitively equivalent for multiple modes. In that case, the advantages of CATI can be retained while at the same time the disadvantage of undercoverage can be overcome. The use of CATI in a mixed mode design has been analysed but deserves more attention, especially with respect to the coverage problem.

Appendix

12.A Estimating Telephone Coverage Probabilities

The telephone coverage probability τ_i is a latent variable. Furthermore, its value lies between 0 and 1. To estimate this type of variable we can use a probability model that restrains the predicted outcome to lie between 0 and 1. In other words, we want $P(T_i = 1) = F(\mathbf{X}_i'\beta)$ with some function F translating $\mathbf{X}_i'\beta$ into a value between 0 and 1. $T_i = 1$ when sample element i has a telephone, otherwise $T_i = 0$. $\mathbf{X}_i$ represents a $J \times 1$-vector of auxiliary variables with values for sample element i, and β is a $J \times 1$-vector of coefficients that correspond to the different auxiliary variables in $\mathbf{X}_i$.

The telephone coverage probabilities τ_i are theoretical quantities. The individual probabilities can only be estimated by looking at the aggregate sample level. For each sample element there simply is only one replication. These probabilities can hence be estimated based on the sample. By using an appropriate model based on auxiliary information, we can compute sample-based estimates of the telephone coverage probabilities, i.e.

$$\hat{\tau}_i = \rho(\mathbf{X}_i) = P(T_i = 1|\delta_i = 1; \mathbf{X}_i) \tag{12.14}$$

for $i = 1, 2, \ldots, n$. Hence, $\hat{\tau}_i = \hat{\tau}_j$ if $\mathbf{X}_i = \mathbf{X}_j$, i.e. person i has the same probability of having a telephone as all other persons in the same strata defined by $\mathbf{X}$. We refer to $\hat{\tau}_i$ as the *telephone coverage propensity*. The telephone coverage propensity is the estimated telephone coverage probability conditional on the sample and the individual characteristics $\mathbf{X}_i$.

We consider three models for computing the telephone coverage propensities $\tau(\mathbf{X}_i)$, each using a different transformation function F, or, in the framework of Generalized Linear Models (GLM, see for example Greene, 2003), a different known link function. The first choice of F is the identity function, which results in the linear probability model $P(T_i = 1) = \mathbf{X}_\mathbf{i}'\beta$. This model, however, does not restrict the propensities to lie between 0 and 1. Another transformation is the standard normal distribution, which does restrict the probability to lie between 0 and 1. This results in the probit model

$$\tau(\mathbf{X}_i) = \Phi(\mathbf{X}_\mathbf{i}'\beta) = \int_{-\infty}^{\mathbf{X}_\mathbf{i}'\beta} \frac{1}{\sqrt{2\pi}} exp(\frac{-z^2}{2})dz \tag{12.15}$$

When choosing F to be the logistic distribution, this leads to the logit model

$$\tau(\mathbf{X}_i) = \Lambda(\mathbf{X}_\mathbf{i}'\beta) = \frac{exp(\mathbf{X}_\mathbf{i}'\beta)}{1+exp(\mathbf{X}_\mathbf{i}'\beta)} \tag{12.16}$$

The probit and the logit function are the most common models but in fact any model with the right properties can be used. The logit transformation leads to the logistic regression model

$$log(\frac{T_i}{1-T_i}) = logit(\tau(\mathbf{X}_i)) = \mathbf{X}_\mathbf{i}'\beta \tag{12.17}$$

whereas the probit transformation is defined as the solution of (12.15). Note that the logit regression model does not have an error term. The probit and the logit models can be described as latent variable regression models. To see that, we define a latent variable $\tau_i^*(\mathbf{X}_i)$ as

$$\tau^*(\mathbf{X}_i) = \mathbf{X}_i'\beta + \varepsilon_\mathbf{i} \tag{12.18}$$

The latent variable $\tau^*(\mathbf{X}_i)$ can be regarded as a transformation of the telephone coverage propensity into the $(-\infty,\infty)$-interval. The error terms are assumed to be identical and independently normally distributed, i.e. $\varepsilon_i \sim NID(0,\sigma^2)$. We do not observe $\tau^*(\mathbf{X}_i)$, but rather we observe the telephone indicator T_i which takes on values of 0 or 1 according to

$$T_i = \begin{cases} 1, \; if \tau^*(\mathbf{X}_i) > 0 \\ 0, \; else \end{cases} \tag{12.19}$$

The latent variable $\tau^*(\mathbf{X}_i)$ is not equal to the telephone coverage probability τ_i because its value is not restricted to the interval $(0,1)$. However, the probit or logit transformation ensures that the estimated values, the telephone coverage propensities $\hat{\tau}_i$, can be interpreted as probabilities.

For more information on these models see for instance Johnston and Dinardo (1997) or Greene (2003). Does it matter which model is used to compute the propensities? Some research has been performed to answer this question for the estimation of response propensities. According to Dehija and Wahba (1999) it does not matter which model is used. They use a logistic model and state that other models yield similar results. However, Laaksonen (2006) provides some examples in which it does matter which model is used. He evaluates logit, probit, complementary log-log and log-log models. Furthermore, there is no benchmark with respect to how well the model should fit the data in order to be used effectively in adjustment for nonresponse.

References

1. Callegaro, M., Poggio, T.: Where can i call you? the ‘mobile’ revolution and its impact on survey research and coverage error: Discussing the italian case. In: RC 33 Sixth International Conference on Social Science Methodology (2004)
2. Cobben, F.: Nonresponse in sample surveys. methods for analysis and adjustment, Ph.D. thesis (2009)
3. Cobben, F., Bethlehem, J.G.: Adjusting undercoverage and nonresponse bias in telephone surveys. Discussion paper 05006, Statistics Netherlands (2005), http://www.cbs.nl
4. Cochran, W.: The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 24, 205–213 (1968)
5. De Leeuw, E.D.: Data quality in mail, telephone, and face to face surveys, Ph.D. thesis (1992)
6. Dehija, R., Wahba, S.: Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. Journal of the American Statistical Association 94(448), 1053–1062 (1999)
7. Ellis, C.H., Krosnick, J.A.: Comparing telephone and face-to-face surveys in terms of sample representativeness: A meta-analysis of demographic characteristics, Technical Report Series nes010871, NES (1999)
8. van Goor, H., Rispens, S.: A middle class image of society. Quality and Quantity 38(1), 35–49 (2004)
9. Greene, W.H.: Econometric analysis, 5th edn. Prentice-Hall (2003)
10. Groves, R.M.: Survey errors and survey costs. Wiley, New York (1989)
11. Johnston, J., Dinardo, J.: Econometric methods. McGraw-Hill (1997)
12. Kalton, G., Flores-Cervantes, I.: Weighting methods. Journal of Official Statistics 19(2), 81–97 (2003)
13. Kuusela, V.: A survey on telephone coverage in finland. Zuma Nachrichten Special, 113–119 (August 1998)
14. Laaksonen, S.: Does the choice of link function matter in response propensity modelling? Model Assisted Statistics and Applications I, 95–100 (2006)
15. Little, R.J.A.: Survey nonresponse adjustments for estimates of means. International Statistical Review 54, 139–157 (1986)
16. Luiten, A.: Methodenreeks; thema veldwerkorganisatie; deelthema responsberekening sociale statistieken, Methodenreeks, Statistics Netherlands (2011) (in Dutch)
17. Pickery, J., Carton, A.: Hoe representatief zijn telefonische surveys in vlaanderen?, Nota 4, Administratie Planning en Statistiek, Ministerie van de Vlaamse Gemeenschap (2005) (in Dutch)
18. Särndal, C.E.: Frameworks for inference in survey sampling with application to small area estimation and adjustment for non-response. Bulletin of the International Statistical Institute 49, 494–513 (1981)
19. Vehovar, V., Belak, E., Batagelj, Z., Cikic, S.: Mobile phone surveys: The slovenian case study. Methodološki sveski. Advances in Methodology and Statistics 1(1), 1–19 (2004)

Chapter 13
Nonresponse in an Individual Register Sample Telephone Survey in Lucerne (Switzerland)

Oliver Lipps and Kathrin Kissau

13.1 Introduction

Surveying people on the telephone has a number of advantages compared to face-to-face interviews, especially if the survey is carried out by telephone centres (Häder 2009):

- Large samples can be realized within short periods.
- It is easy to call households that are difficult to access frequently.
- Monitoring interviewers (defining cases as ineligible, interviewing the actually sampled individuals, conducting the interview according to defined protocols, etc.) is easy.
- The influence of third persons is likely to have no or only minor effects.
- Interviewer effects are likely to be smaller.
- The fear of letting an interviewer into the house does not play a role in telephone surveys.
- Many face-to-face surveys need (regional) clustering of sample cases which causes design effects over and above those caused by interviewers who work (regional) sample points. The consequence is that a higher number of interviews are necessary in face-to-face surveys to obtain the same precision.

In Switzerland telephone surveys are particularly popular. This is due to cost issues as well as the fact that up to now samples for "representative" national social science surveys were usually drawn from the telephone directory (Jann 2007). Concerning the surveys conducted by the Lausanne based Swiss Centre of Expertise in Social Sciences (FORS) alone, this frame is used for both telephone surveys (e.g. the Swiss electoral studies (SELECTS)[1] (Lutz 2008) or the Swiss Household Panel (SHP)[2] (Voorpostel et al. 2010)) and face-to-face surveys (e.g. the Survey of Health, Aging and Retirement in Europe (SHARE)[3] (Klevmarken 2005) or the first three rounds of the European Social Survey (ESS)[4] (ESS 2002, 2004, 2006)).

[1] http://www.2.unil.ch/selects/spip.php?rubrique3&lang=en
[2] http://www.swisspanel.ch/?lang=en
[3] http://www.unil.ch/share
[4] http://www.europeansocialsurvey.org/

Problems with this sampling frame are increasing under-coverage and bias (see Ernst-Stähli in this volume). A way to react to these challenges of drawing representative random telephone samples is to change the sampling frame. One possibility in Switzerland is to make use of address-based registers instead of telephone directories. The advantage of registers is their almost[5] full coverage of the sample population. For the fourth round of the ESS (ESS 2008), for example, a file maintained by the Swiss Post, which includes all buildings in Switzerland and an indication of the number of private letter boxes in each, was used. The sample was randomly drawn from all private letter boxes. The ESS surveys only targeted individuals, which makes the screening of the household in multiple person homes necessary. Therefore, this two-step process adds the problem that gatekeepers may refuse to cooperate and provide information about the household or give access to the targeted person.

To take advantage of the benefits of the telephone once an address sample is drawn, the telephone survey mode can still be used (Johnson and Williams 2010). This of course involves the task of finding suitable information sources with telephone numbers and matching them with the sampled addresses. The same is necessary if persons are drawn from an individual register instead of an address-/household-based register. In this case register-based information is available about all individual sample members, i.e. also the nonrespondents. If address frames are used, matching is of course not necessary when the survey mode is directly based on addresses (Link et al. 2006 who use a mail survey).

Since the end of 2010, a nationwide harmonized individual register is available in Switzerland[6] and most surveys run by FORS will make use of this sampling frame. To our knowledge sampling bias has not been analyzed in Switzerland in samples drawn from an individual register frame and matching this sample with fixed line telephone numbers using different information sources successively.

In this chapter we will therefore

1. present efforts of matching telephone numbers from different information sources to individuals sampled from an address-register for a telephone survey that is representative of the general population.
2. analyze sampling effects by looking into cases where a telephone number could be matched and cases for which no telephone number could be found.
3. analyze the characteristics of people from the activated (matched with a telephone number and released for the fieldwork) sample according to their final response status.

We use data from a telephone survey that was carried out in the Canton of Lucerne from October to December 2009. The innovation of this survey is that it is based on a sample that was drawn from individual registers. Starting from a random sample of individuals, the primary aim of this research is to analyze which socio-demographic groups are over- or under-represented in telephone samples (gross sample) with various efforts invested to match telephone numbers. The

[5] Excluding those without a permanent residence, or when changes of addresses are reported with a time delay.

[6] http://www.bfs.admin.ch/bfs/portal/de/index/news/00/08.html (in German).

secondary aim is to reveal different final response statuses of socio-demographic groups and, if relevant, diverse reasons given for refusals. For both parts of our analysis socio-demographic information available in the registers on all individuals (respondents and nonrespondents) can be used, which is an advantage compared to other nonresponse studies. It is expected that this analysis will give new and valuable insights into possible sampling improvements that are possible with the availability of the Swiss nationwide individual register.

13.2 Methodological "Pilot Study" in the Canton of Lucerne

In 2009 a population survey with roughly 4,000 CATI interviews was carried out within the Swiss canton of Lucerne ("Einwohnerbefragung Kanton Luzern 2009"). This canton with a population of about 370,000 is located at the centre of Switzerland in the German-speaking part of the country. The cantonal capital is the city of Lucerne. The survey was initiated by the government of the canton of Lucerne and covered a broad spectrum of questions on the living conditions of Lucerne citizens and their opinions about their living environment. The idea behind this survey was to provide policy makers with information about the interests, preferences, appraisal and satisfaction of the inhabitants, in order to evaluate the current situation and to aid the definition of future goals and measures (LUSTAT 2010). In this paper, however, we will not focus on this substantive data, but rather on the sample and the effects of the sampling procedure.

The sample that was representative of the residential population of the canton of Lucerne was drawn by the cantonal statistical office ("LUSTAT") from the communal individual registers of citizens. The population was defined as all permanent[7] residents above the age of 15 living in private households in the canton. The sample was stratified by four age groups, gender, four community size groups, and nationality (Swiss or non-Swiss). For each of the 64 strata a proportional random sample was drawn. In total, 12,501 addresses were selected, comprising the gross sample. Because the communal registers do not include telephone numbers, telephone numbers of the households, where the sampled individuals live, had to be found and matched with the individuals.

In a first step, the addresses of the sample members were matched electronically to the electronic telephone directory (ETV - "Elektronisches Telefon Verzeichnis") which does not include unlisted numbers. 9,400 (75.2 %) individuals could thus be assigned a telephone number[8]. In a second step the assignment was cleaned and, if possible, supplemented by LUSTAT by "manually" matching additional 677 (5.4 %) cases using different telephone directories. Third, 52 (0.4 %) further citizens could be assigned a number with the help of information from other members of the household. The rest of the gross sample (n = 2,386; 19.1 %) received a postcard, in which the survey was described

[7] This excludes foreigners with a permit of less than one year, and individuals in an asylum procedure.

[8] We assigned 14 landline telephone numbers where the origin of the number was not published to this group.

and the target person was asked to deliver a telephone number where he/she could be reached. Of those addressed, 588 (4.7 %) returned the postcard with a valid telephone number in a fourth step[9]. Some of these people disclosed a mobile or an office telephone number. 1,713 (13.7 %) individuals could not be assigned a valid telephone number. These individuals are citizens who did not answer the postcard (though they possess a mobile or/and a fixed line telephone), do not have access to a telephone at all, or could be matched neither electronically nor manually. Altogether, a valid telephone number could be identified for 10,788 (86.3 %) individuals[10]. Altogether, 5,769 sample members were randomly activated for the survey (un-cleaned gross sample) and sent an advance letter signed by the president of Lucerne´s governing council which informed them about the upcoming survey, about two weeks before they were called. Subsequently, 71 of these sample members had to be deleted because they or household members refused participation due to illness, disease, etc. in a written form[11].

We list the sources of the telephone numbers of the sample members in Table 13.1 including those for whom a telephone number could not be found and those cases where the telephone number had to be deleted.

Table 13.1 Source used to identify Telephone number. Data: Population Survey Canton of Lucerne 2009

Source of telephone number	n	%
Automatic telephone number match using electronic telephone directory ETV	9,400	75.2
Telephone number manually matched	677	5.4
Telephone number found by means of other household members	52	0.4
Telephone number delivered by postcard	588	4.7
No telephone number available	1,713	13.7
Telephone number deleted due to written refusal after receipt of advance letter	71	0.6

For the analyses, we combine the two groups "Telephone number manually matched" and "Telephone number found by means of other household members". The reasons are that both sources are usually not used and that the second group is very small.

We briefly describe the final response status of the 5,769 activated sample cases (telephone numbers), excluding the ineligible sample cases (n = 39; 0.7 %: deceased, moved out of the canton, collective households). 373 sample members (6.5 %) could not be reached. Of those that refused to participate, 221 individuals stated health problems and 121 language problems. Because sample members not

[9] We assigned seven mobile telephone numbers where the origin of the number was not published to this group.

[10] Including the 71 numbers that have to be deleted due to written refusal by a household member after receipt of the advance letter, see below.

[11] We treat all these cases as refusals, although some (e.g., deceased) might stem from out-of-date individual registers. However, it is not clear whether the reasons stated are indications of non-eligibility or real refusals.

able to answer the survey in (Swiss-)German were offered to answer in Italian, Portuguese, Serbo-Croatian, Bosnian, or in Albanian, the latter figure is surprisingly high. 475 individuals declined to participate twice having been contacted a second time in an attempt of refusal conversion. 602 individuals never made an appointment ("call later") or broke a fixed date. The response rate amounts to 68.7 %. In Table 13.2 we list the distribution of the reasons for eventual nonresponse in the activated sample of eligible citizens.

Table 13.2 Final response status of activated sample. Data: Population Survey Canton of Lucerne 2009

Final response status of activated sample members	n	%
Not reached	373	6.5
Health problems	221	3.9
Language problems	121	2.2
Not fixed or broken appointment	602	10.5
Refusal after refusal conversion attempt	475	8.3
Participants	3,938	68.7
Total	5,730	100.0

13.3 Socio-demographic Characteristics of Sample by Source of Telephone Number

In this section we analyze characteristics of all 12,501 sample members by stage of identification of a telephone number, if ever, according to the variables available from the register. We report socio-demographic characteristics by the stage of telephone number identification first separately and then cumulatively. The following socio-demographic variables are available from the register:

- Age groups: 15 - 29 years (base category), 30 - 44 years, 45 - 59 years, 60 + years
- Region: city of Lucerne (base category), agglomeration of Lucerne, large villages, small villages
- Civil status: unmarried, married (base category), divorced, widowed
- (First) nationality: Swiss (base category), from a neighboring country (Germany, France, Italy, Austria), from another country. The most important reason for making this distinction is that these countries share one of the national languages in Switzerland (Lipps et al. 2011)
- Gender

We first present graphs with the distributions of the sample socio-demography by information source. For the sake of completeness we include written refusals. Next, we check whether the bias in the socio-demographic variables can be reduced by gradually including those samples, whose telephone numbers were retrieved from the additional information sources, and illustrate the changed sample compositions. Finally, we run a series of logit models comparing samples that result when in addition to the base (the automatic telephone number matching) the telephone numbers from the other information sources are sequentially added.

Before we turn to this, we briefly analyze the socio-demography of those who reported a mobile number versus a fixed line telephone among those who sent back the return card. This subsample is of course not representative of the residential population, but it might nonetheless provide some insights into the preferences for being contacted on a mobile or landline phone of cooperative sample members, who have an unlisted and/mobile phone number.

13.3.1 Returned Card Sample: Telephone Type by Socio-demography

When we look at the age distributions, we find that, not surprisingly, younger citizens tend to report a mobile number. Individuals living in the agglomeration and especially in the city of Lucerne more often report mobile phone numbers compared to those living in large and small villages (see Table 13.3).

Table 13.3 Total number and percentage of mobile numbers reported in return card by socio-demography

Telephone type reported in return cards [% mobile]	n (return card)	% mobile
Age groups		
15 - 29	190	59
30 - 44	198	45
45 - 59	127	40
60 +	73	18
Region		
City of Lucerne	167	55
Agglomeration of Lucerne	172	51
Large Villages	137	35
Small Villages	112	36
Civil Status		
Unmarried (never married)	290	57
Married	215	29
Widowed	15	20
Divorced	68	50
Nationality		
Swiss	452	45
From neighbouring country	52	36
Foreigner	77	53
Gender		
Female	281	33
Male	307	56
Total	588	45

Data: Population Survey Canton of Lucerne 2009.

The difference in the civil status of landline or mobile phone number reporters is evident. Unmarried individuals report mobile numbers more often, married individuals less often. As to the nationality, while the Swiss do not have a special preference, those with a nationality from a neighboring country tend to report a landline number, and those with a nationality from another country a mobile number. When it comes to gender differences, we find that men report a mobile number more often than women.

Overall, this provides us with a tentative profile of citizens whose phone numbers cannot be found out through public telephone directories or through other household members. Of these "difficult to reach" individuals, those that prefer being contacted by mobile phone tend to be younger, unmarried, male, not of a nationality from one of the neighboring countries and living in a city or agglomeration. After this short excursus focusing on the demographics of those that sent a return card back, we now turn to the analysis of the various steps of the process of matching telephone numbers with the sampled individuals, and the respective socio-demographic distributions.

13.3.2 Age Group by Telephone Status

In our first analysis of the sample demography by information source, we focus on age. We note that the higher proportion of written (pre-) refusals among older persons (or someone else in the household on their behalf) is related to the reasons given (age, health, disease). There are more young people and fewer elderly in the manually matched sample, when compared with the automatically matched sample (see Figure 13.1a and 13.1b).

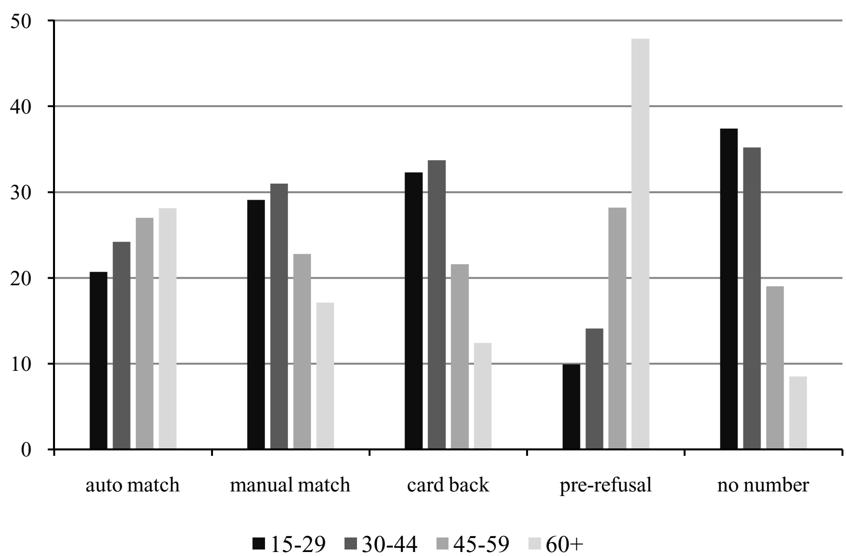


Fig. 13.1a Age group distributions by source of telephone number [%]. Data: Population Survey Canton of Lucerne 2009

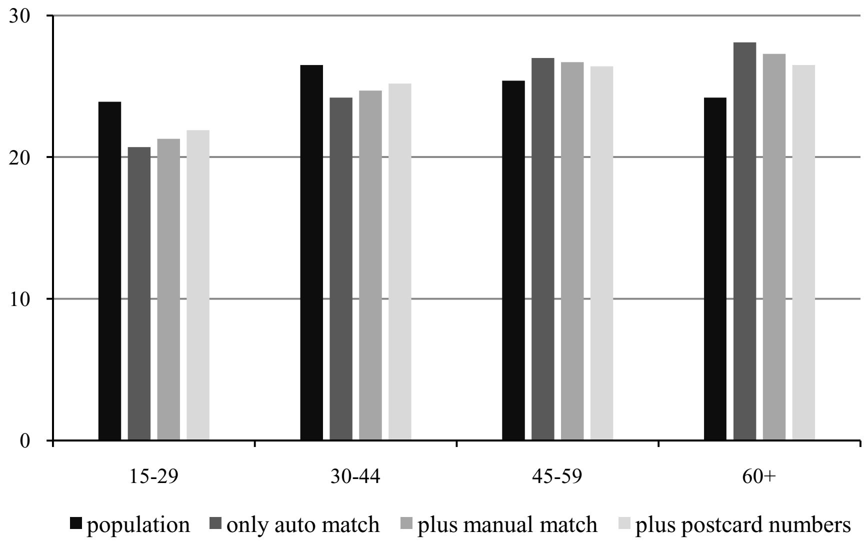


Fig. 13.1b Age group distributions by step in the matching process [%]. Data: Population Survey Canton of Lucerne 2009

This discrepancy is even larger in the sample that sent back a card with their telephone number. As expected, the youngest age group is by far the largest and the oldest age group is by far the smallest in the "no number" group.

To check if the bias in the socio-demographic variables can be reduced by gradually taking into consideration additionally matched telephone numbers, we portray the age distribution in the total sample (population) and in the sample that could be automatically matched. Stepwise we then add those that were manually matched, and those who reported a valid telephone number. The proportion of the 15 - 29 year-olds in the total (gross) sample amounts to 23.9 %, of the 30 - 44 year-olds to 26.5 %, of the 45 - 59 year-olds to 25.4 %, and of the 60 years and older to 24.2 %. The older the sample members the more are they overrepresented among the automatically matched, and vice versa. This distortion can be reduced by including those manually matched and the telephone numbers delivered with the return card. However, the initial bias mostly remains.

13.3.3 Region by Telephone Status

With regard to the information sources for telephone numbers of inhabitants of the four regions in the canton, we observe that while by manual searches comparatively many telephone numbers can be found in small villages, individuals from the city of Lucerne send back the postcard more often and those living in large villages less frequently (see Figure 13.2a and 13.2b).

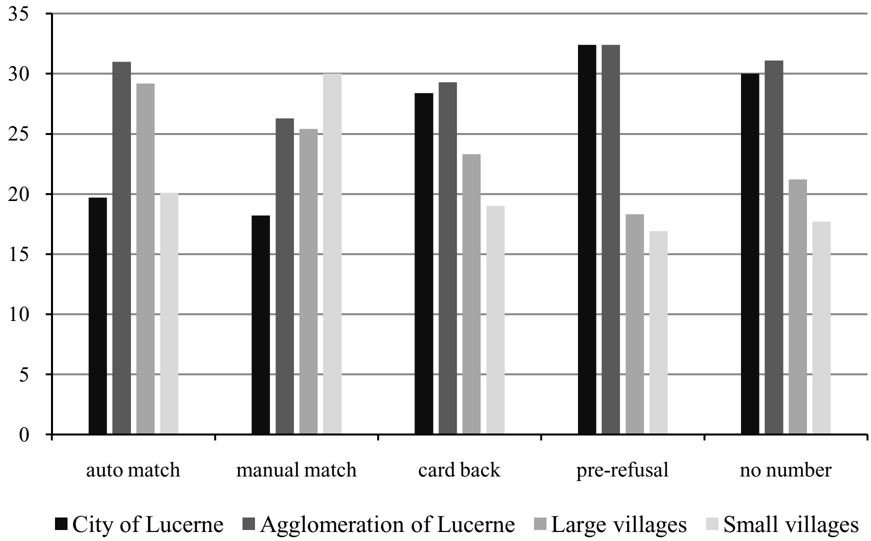


Fig. 13.2a Regional distributions by source of telephone number [%]. Data: Population Survey Canton of Lucerne 2009

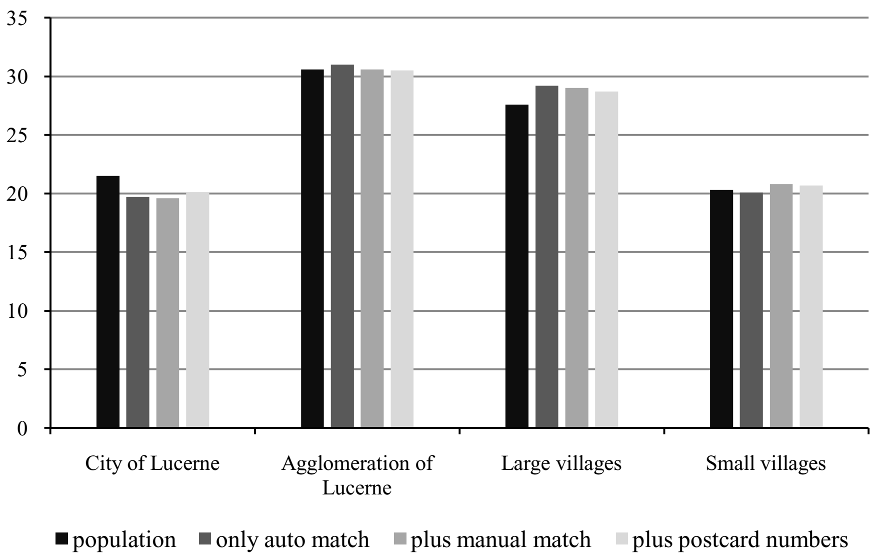


Fig. 13.2b Regional distributions by step in the matching process [%]. Data: Population Survey Canton of Lucerne 2009

As opposed to individuals from large villages, individuals from the city of Lucerne are less likely to have a telephone number matched. Also written refusals occur more often among citizens of the city of Lucerne, especially when compared to inhabitants of large villages.

The proportion of individuals from the city of Lucerne in the total sample amounts to 21.5 %, from the agglomeration to 30.6 %, from larger villages to 27.6 %, and from smaller villages to 20.3 %. When comparing the distribution of citizens from the four regions within the population with the distribution within various steps of the sample compilation, it becomes clear that the regional bias is much smaller than that of the age group when only an automatic match is conducted. We find that the underrepresentation of inhabitants of the city of Lucerne based on the automatically matched cases can be slightly improved by the inclusion of the manually matched, and especially of those who report their telephone number in the return card.

13.3.4 Civil Status by Telephone Status

When it comes to the telephone number sources by civil status, two patterns are noticeable. Many more unmarried people than people with another civil status are in the groups "card back", "written refusal" and "no number". In contrast, there are many married sample persons in the automatically and manually matched groups and rather few in the group of written refusals, where, expectedly, many are widowed (see Figure 13.3a and 13.3b).

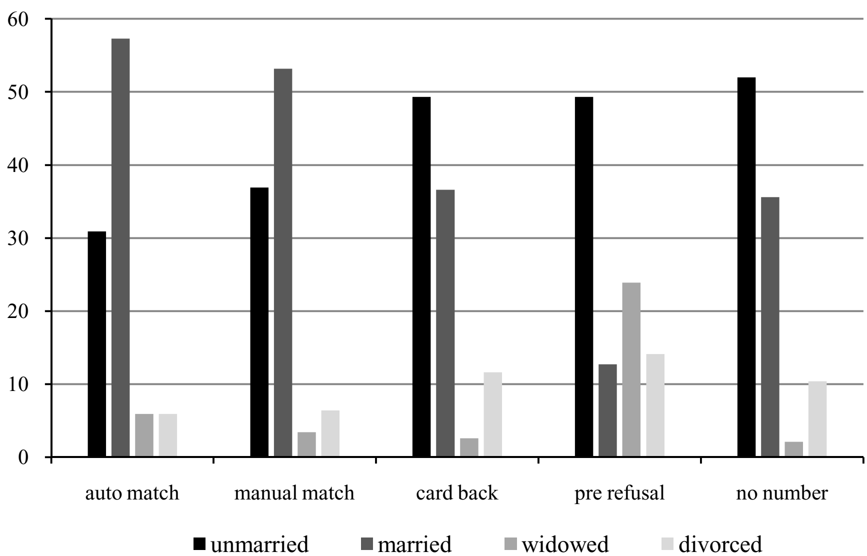


Fig. 13.3a Civil status distributions by source of telephone number [%]. Data: Population Survey Canton of Lucerne 2009

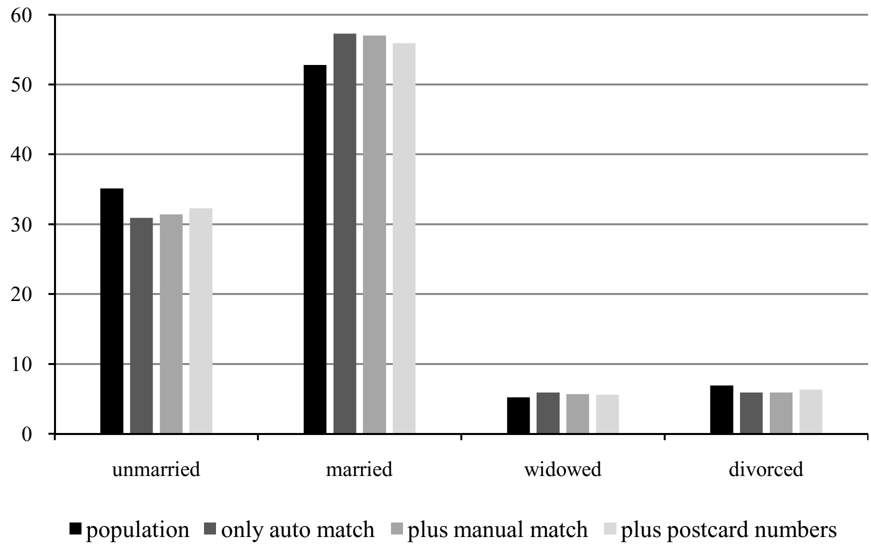


Fig. 13.3b Civil status distributions by step in the matching process [%]. Data: Population Survey Canton of Lucerne 2009

The proportion of unmarried individuals in the total sample amounts to 35.1 %, of the married to 52.8 %, of the widowed to 5.2 %, and of the divorced to 6.9 %. The comparison of the population composition with the samples reveals that a bias exists especially regarding the married and unmarried inhabitants if only the automatically matched are considered. The insufficient number of unmarried individuals and the disproportionately large number of married persons in the automatically matched sample can to some extent be balanced by including those who report their telephone number in the return card.

13.3.5 Nationality by Telephone Status

When focusing on the telephone number sources of individuals according to their nationality, we find the lowest proportion of foreigners in the automatically matched directory, the highest in the group without a telephone number. Telephone numbers of citizens from neighboring countries are most frequently matched through a returned postcard, of foreigners from other countries when manually searched (see Figure 13.4a and 13.4b)

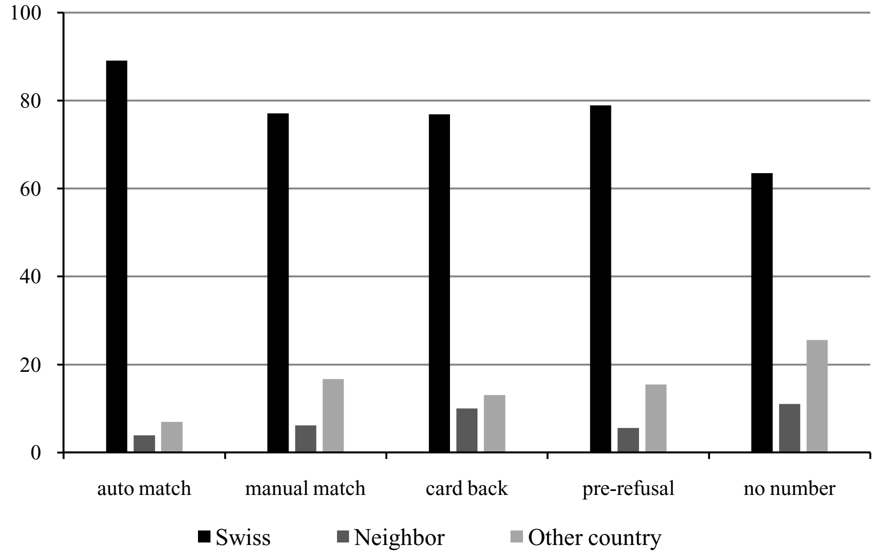


Fig. 13.4a Nationality distributions by source of telephone number [%]. Data: Population Survey Canton of Lucerne 2009

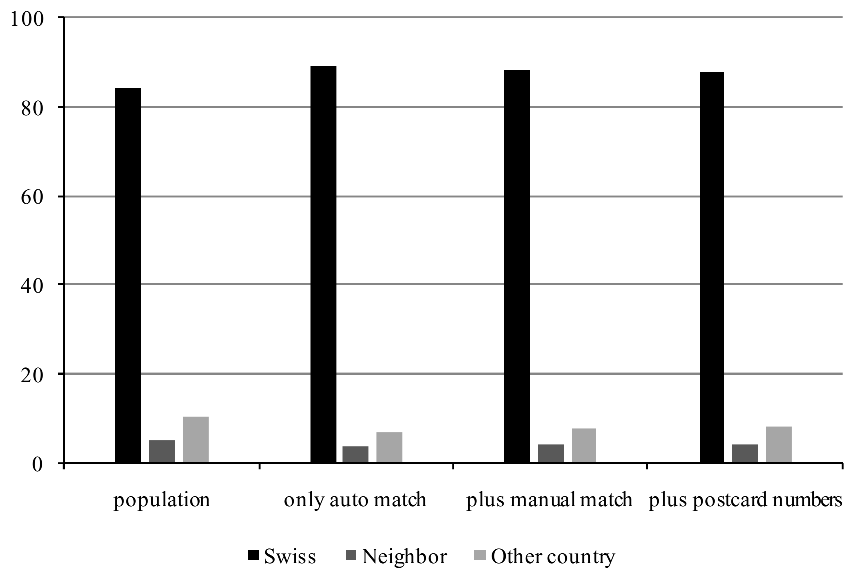


Fig. 13.4b Nationality distributions by step in the matching process [%]. Data: Population Survey Canton of Lucerne 2009

The proportion of individuals with Swiss first nationality in the total sample amounts to 84.3 %, with a nationality of one of the neighboring countries to 5.3 %, and with a nationality of another country to 10.4 %. When comparing the distribution of nationalities in the population with that of the automatically matched sample, it is observable that foreigners are generally underrepresented when their numbers are merely retrieved automatically. The insufficient portion of foreigners can to some degree be compensated by including other telephone numbers than those that were automatically matched and especially by including cases where a return card was sent back, as far as foreigners from neighboring countries are concerned.

13.3.6 Gender by Telephone Status

The proportion of men in the total sample amounts to 49.2 %. There are only small differences in the various samples apart from more women among those who sent a written refusal (see Figure 13.5a and 13.5b).

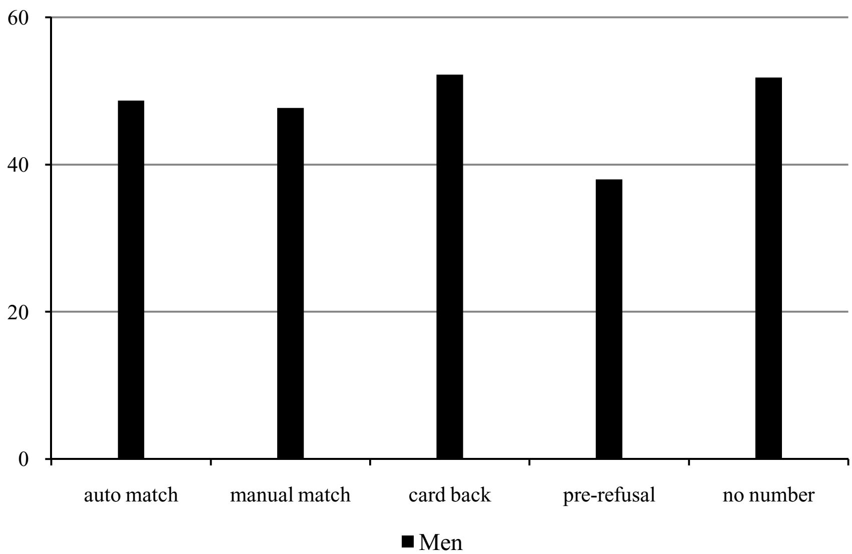


Fig. 13.5a Gender distributions by source of telephone number [%]. Data: Population Survey Canton of Lucerne 2009

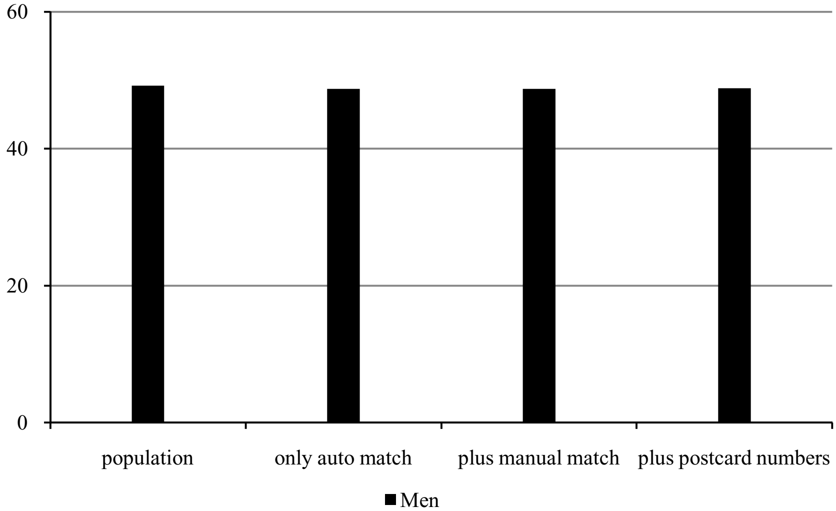


Fig. 13.5b Gender distributions by step in the matching process [%]. Data: Population Survey Canton of Lucerne 2009

Generally the distribution of gender within the automatically matched sample hardly changes through the inclusion of the other samples.

13.3.7 Socio-demography by Telephone Status: Multivariate Analysis

After this descriptive assessment of the sample characteristics by the source of the individual's telephone number, we now turn to a multivariate analysis. We present the results of a series of multivariate logit models, showing the marginal effects of socio-demographic register variables at each step of the matching process (see Table 13.4). The cumulative sources used to match the individual's telephone number are compared with the sampled citizens that could not be matched through the three sources used. The advantage of this procedure is that the "no number available" (NN) category contains the same individuals in each logit model[12]. This NN group is first compared with the automatically matched (AM) group, then with the combined AM and manually matched (MM) groups, and finally with the combined AM, MM, and the return card (RC) groups. As a consequence, the total sample analyzed is the sum of the AM, MM, RC, and NN groups. However, all four groups are only used in the last model (AM + MM + RC versus NN).

[12] We exclude the "written refusals" group, first because of its small size and second because it is not comparable with the other subsamples.

Statistical differences between the members in the currently matched group and NN show:

1. the socio-demographic variable category that is differently distributed in the sample group (AM, AM + MM, or AM + MM + RC) considered and
2. given that there is bias, the models reveal if the bias increases or decreases with additional sample groups included.

If the final sample (AM + MM + RC) were unbiased, its socio-demographic distribution would have to correspond to the sampled group that remained without a telephone number (NN) and also to the total sample (AM + MM + RC + NN).

In addition to the coefficients we list the results of significance tests which compare the coefficients with the sample from the previous stage (in brackets).

Table 13.4 Regression Coefficients [logits] and test of sample coefficient differences

Regression Coefficients [logits]	*Model 1* Automatic match (AM) vs. no number available (NN)	*Model 2* Manual match (AM + MM) vs. NN	*Model 3* Return card (AM + MM + RC) vs. (NN)
30-44 years old (vs. 15-29)	-.06	-.08	-.05
45-59 years old (vs. 15-29)	.56**	.50** (AM**)	.49**
60+ years old (vs. 15-29)	1.27**	1.19** (AM**)	1.15** (MM**)
Agglomeration Lucerne (vs. City L.)	.35**	.35**	.30**
Large village (vs. City of Lucerne)	.60**	.57**	.52** (MM**)
Small village (vs. City of Lucerne)	.35**	.37**	.32** (MM**)
Unmarried (vs. married)	-.84**	-.84**	-.78** (MM**)
Widowed (vs. married)	-.40*	-.41*	-.37
Divorced (vs. married)	-1.37**	-1.32** (AM*)	-1.24** (MM**)
From neighbour country (vs. Swiss)	-1.28**	-1.22** (AM**)	-1.13** (MM**)
From another country (vs. Swiss)	-1.67**	-1.55** (AM**)	-1.49** (MM**)
Male	-.05	-.06	-.06
n	11,113	11,842	12,430
Pseudo R^2	.141	.129	.117

** (*) = significant on 1 (5) % level. Data: Population Survey Canton of Lucerne 2009.

The comparatively high pseudo R^2 of 14 % of the first model shows that the AM and the NN samples are different with respect to the socio-demographic variables investigated. With the exception of gender, all socio-demographic variables are different, which indicates that the automatically matched sample of the Lucerne survey is biased with regard to age, region, civil status and nationality. Adding the MM and especially the RC telephone numbers, however, reduces the pseudo R^2, decreases the absolute values of the coefficients and thus lowers the bias to a certain degree. This is also confirmed by many significant differences between the sample coefficients when adding the manually matched sample group as well as the group matched using return cards. Nevertheless, the final telephone sample (AM + MM + RC) is still quite different from the NN sample in terms of the socio-demographic variables included.

We now discuss the sample bias and its changes when additional subsamples are incorporated with respect to the socio-demographic variables included.

As for age groups (base: 15 - 29 years old) the probability to belong to the 30 - 44 year-olds is not biased and also does not change across the subsamples. The 45 - 59 years and especially 60 + years old are, however, strongly overrepresented in all telephone subsamples. Nevertheless the absolute value of the coefficient decreases which shows that the bias decreases, when the subsamples are added.

Concerning region (base: city of Lucerne) we observe that while people from the city of Lucerne are underrepresented, all other areas and especially large villages are over-represented. This bias can also be reduced by adding the additional telephone subsamples.

Regarding civil status (base: married) it is evident that the unmarried and especially the divorced are constantly underrepresented. Again, this under-representation can be reduced by adding further telephone samples. The insufficient proportion of widowed (note: even if controlled for age!) can even be turned insignificant by adding the RC sample.

As far as nationality is concerned (base: Swiss) both foreigner groups are strongly underrepresented. This underrepresentation can also be reduced by adding MM and RC telephone numbers.

Gender is the only variable without significant bias throughout all samples.

13.4 Socio-demographic Characteristics by Response Status

In this section we briefly analyze the socio-demographic differences of the activated sample excluding the ineligibles (n = 5,127) according to their final response status. In addition, we present the final response statuses for the different telephone samples. We distinguish six final statuses:

1. the participants
2. those who could not be reached
3. those who did not fix an appointment ("call later") until the end of the fieldwork or broke one
4. those who state language problems
5. those who state health problems
6. those who refuse twice (also after refusal conversion attempts)

In the cross-tables (Table 13.5 and 13.6) we list the cell- and total Chi^2 contributions for each category of socio-demographic variables and the telephone samples by the different response status groups[13]. In addition, we indicate a smaller than expected frequency with a "(-)" sign behind the value. We find total Chi^2 values that are highly significant for the response groups by all socio-demographic variables. These tables give hints of which final status (and which reason for refusal is used if applicable) is more likely for which socio-demographic group. Suitable measures to increase response rates can be better tailored (Groves and Couper 1998) to specific groups if their profile is better known.

13.4.1 Age Group by Response Status

When focusing on the response groups by age categories we find some expected patterns: fewer not fixed (no appointment) or broken appointments (not reached) and many more health related refusals among the elderly, and more not reached among the young people.

Table 13.5 Chi^2 contribution of age groups by response status. "(-)" indicates a smaller than expected frequency. Data: Population Survey Canton of Lucerne 2009.

Chi^2 contribution	15 - 29	30 - 44	45 - 59	60 +	Total
No appointment	5.6	.9 (-)	4.0	10.7 (-)	21.1
Health problem	33.3 (-)	22.4 (-)	19.0 (-)	209.8	284.4
Language problem	2.6 (-)	11.6	.5	7.3 (-)	22.1
Not reached	77.3	.8 (-)	6.5 (-)	21.0 (-)	105.7
Participants	.3 (-)	.8	.2	.6 (-)	1.9
Refusers	9.2 (-)	.0	8.9	.3	18.5
Total	128.3	36.5	39.1	249.8	453.8

13.4.2 Region by Response Status

When it comes to response groups by region, we generally observe fewer differences. We find more not fixed or broken appointments and more language problems in the city of Lucerne but fewer refusers. People seem to refuse by not

Table 13.6 Chi^2 contribution of region by response status. "(-)" indicates a smaller than expected frequency. Data: Population Survey Canton of Lucerne 2009

Chi^2 contribution	City of Lucerne	Agglomeration of Lucerne	Large village	Small village	Total
No appointment	4.0	.1	1.5 (-)	.7 (-)	6.3
Health problem	.0	.7 (-)	.0	2.2	3.0
Language problem	20.1	.1	5.6 (-)	4.9 (-)	30.7
Not reached	1.6 (-)	1.5	.1	.3	3.5
Participants	.0	.2 (-)	.0	.2	.5
Refusers	3.9 (-)	.4	1.6	.1 (-)	6.0
Total	29.7	3.1	8.9	8.4	50.0

[13] Almost all cells have expected frequencies of more than ten.

fixing a time rather than explicitly refusing in bigger cities. That there are fewer not reached in the city of Lucerne is surprising.

13.4.3 Civil Status by Response Status

In the analysis of response groups by civil status we find similar patterns to those of the age groups. High occurrence of health problems exist among the widowed, and more not reached among the unmarried (see Table 13.7).

Table 13.7 Chi^2 contribution of civil status by response status. "(-)" indicates a smaller than expected frequency. Data: Population Survey Canton of Lucerne 2009.

Chi^2 contribution	unmarried	married	widowed	divorced	Total
No appointment	5.1	.4 (-)	4.2 (-)	1.4 (-)	11.0
Health problem	10.6	1.0 (-)	144.0	2.4 (-)	158.0
Language problem	15.7-	7.1	1.9	.05	25.1
Not reached	51.2	16.3 (-)	4.8 (-)	2.3 (-)	74.7
Participants	.3 (-)	.7	3.7 (-)	.5	5.2
Explicit refusers	3.6 (-)	1.1	.0 (-)	2.1	6.9
Total	86.4	26.5	158.8	9.3	281.0

13.4.4 Nationality by Response Status

The analysis of response group by nationality reveals that the comparatively high total R^2 is almost exclusively due to the high frequency of language problems of foreigners from another than a neighboring country (see Table 13.8).

Table 13.8 Chi^2 contribution of nationality by response status. "(-)" indicates a smaller than expected frequency. Data: Population Survey Canton of Lucerne 2009.

Chi^2 contribution	Swiss	Neighbour	Other country	Total
No appointment	2.3 (-)	.2	12.9	15.3
Health problem	.8	4.3 (-)	.7 (-)	5.8
Language problem	47.6 (-)	7.5 (-)	384.7	439.8
Not reached	.0	.2-	.0	.2
Participants	1.1	1.4	12.3 (-)	14.8
Explicit refusers	.4	.3 (-)	1.5 (-)	2.2
Total	52.2	13.8	412.1	478.2

Because there are almost no gender differences with respect to response status, we do not list the table for gender.

13.4.5 Telephone Sample Group by Response Status

When turning to the different subsamples we observe that there are fewer persons stating health problems among those who sent a return card (See Table 13.9). However, language problems occur more often in this group, and especially in the manually matched one. This is presumably because a higher proportion of foreigners belong to these groups, see Figure 13.4a and 13.4b. A higher proportion of those who sent a return card back are participants rather than refusers or not reached, which could be expected.

Table 13.9 Chi^2 contribution of telephone number matching stage by response status. "(-)" indicates a smaller than expected frequency. Data: Population Survey Canton of Lucerne 2009.

Chi^2 contribution	Automatic match (AM)	Manual match (MM)	Return cards (RC)	Total
No appointment	.0	3.1	2.6 (-)	5.7
Health problems	.3	1.5	10.9 (-)	12.7
Language problems	4.4 (-)	22.3	9.0	35.7
Not reached	.2	5.8	19.7 (-)	25.8
Participants	.1 (-)	2.9 (-)	9.3	12.3
Explicit refusers	2.1	2.9 (-)	14.6 (-)	19.4
Total	7.1	38.2	66.2	111.5

13.5 Summary and Conclusion

In the presented research we analyze which socio-demographic groups are over- and under-represented in a telephone survey where the sample was drawn from an individual register in Switzerland. The primary aim is to examine bias if different sources of telephone numbers involving different degrees of effort to match names and addresses from the sample are used. The secondary aim is to check socio-demographic differences in the final response status and, if relevant, different reasons given for refusal. We also examine the final response status by different telephone number sources. Although we use a sample from the canton of Lucerne, it is very likely that our results apply to nationwide surveys.

Our starting point is a sample which was randomly drawn from stratified communal inhabitant registers in the canton of Lucerne. These registers include a number of socio-demographic variables. Because there is no telephone number available in the registers, this must be identified using other information sources. In the survey considered here, three steps were taken:

1. Electronic (automatic) match (AM) of names and addresses from the registers with the electronic telephone directory (ETV), which does not include unlisted telephone numbers. This procedure can be put on the same level as current standard scientific or administrative telephone surveys in Switzerland, where a sample is drawn from this telephone directory or the telephone numbers are matched to a sample drawn from another sampling frame, as is the case here.
2. Manual match (MM) with other telephone directories and external sources, including using information from other household members.
3. Sending return cards (RC) to those whose telephone number could not be found.

Our comparison category is the group of persons for whom no telephone number could be found (no number matched (NN)).

We address our first research focus with bivariate comparisons of differences between the samples from the various telephone number sources with respect to the socio-demographic variables available in the individual inhabitants registers: age (we use four groups: 15 - 29, 30 - 44, 44 - 59, 60 + years old), region (city of Lucerne, agglomeration Lucerne, large villages, small village), civil status (unmarried, married, widowed, divorced), nationality (Swiss, from one of the neighboring countries, from another country), and gender.

We find that in the automatically matched sample the young age group is under-represented; the oldest age group is overrepresented. This bias can be reduced by including the telephone number samples from the additional sources. The same applies to the initial under(over)-representation of inhabitants of the city of Lucerne (large villages), the under(over)-representation of the unmarried (married), and the underrepresentation of foreigners. Including telephone numbers from the RC has the highest bias reducing effects. These results are confirmed by a series of multivariate models: the stepwise reduction of the (absolute) value of the socio-demographic (logit) coefficients comparing the (logarithm of the) probability to belong to one of the telephone samples (AM or MM or RC) with the (logarithm of the) probability of not being matched with a telephone number shows the improved representativeness of the samples, when additional telephone number sources are included. As it appears the representativeness of all socio-demographic groups is improved.

If the relative reduction of the pseudo R^2 with additional sources is taken as a measure for improved representativeness, 17 % improvement is achieved by including all three information sources (AM + MM + RC), compared with AM alone. We identify the youngest age group, the divorced, and foreigners as the socio-demographic groups that are notoriously difficult to match by telephone. It is difficult to assess whether this problem can partially be solved by even more effort in retrieving additional telephone numbers, for example by personally visiting the household to ask for a contact telephone number or matching mobile phone numbers. Recent research in Germany has shown that is possible to both identify mobile numbers and to conduct surveys over mobile phones (Häder and Häder 2009). Thus, a challenge that remains for the future to improve the representativeness of samples is to match mobile numbers with information from individual registers. This is important because we believe that more and more samples will be drawn from (national or communal) inhabitants registers, and while the landline telephone will remain a dominant survey mode the mobile telephone will probably become a more popular survey mode in Switzerland and worldwide. An alternative pathway would be to combine telephone surveys with other survey modes for those who cannot be reached by telephone. However, survey mode effects are a comparatively new issue for survey methodologists in Switzerland.

As far as our second research focus is concerned, the final response status (and reasons for refusal) of the survey sample activated for the fieldwork by socio-demographic group, we find patterns that could be expected: older and widowed people more often fix or keep appointments, but state many more health or age related reasons for refusal. Younger and unmarried people are not reached more frequently. Generally the young and the old are less likely to participate. Language problems are more frequently stated by inhabitants of the city of Lucerne, and especially by foreigners from countries other than one of the neighboring countries Germany, Austria, France, or Italy. This corresponds to a higher proportion of language problems among those not automatically matched. Those who returned the postcard had fewer health problems, are reached more often, participate more often, and rarely refuse the survey. These facts could be expected, because this group shows a certain motivation by sending back the card.

Generally, we believe that drawing samples from registers and then matching phone numbers results in more representative samples then when sampling from a telephone directory. This will probably also be true once the Swiss Federal Statistical Office has access to unlisted telephone numbers; this is currently prepared in the "CASTEM" project[14].

Sample improvements are especially evident, when additional efforts are undertaken to retrieve telephone numbers not automatically matched in the directory, as was done in this population survey in Lucerne. The most important reason is the improved socio-demographic representation of the population. However, it is also likely that those who sent back a return card are, within this "difficult to find" sample group, more representative of the "easier" sample members due to self-selection of sending back the card. As was shown above, the return card group often showed slightly different characteristics compared to the matched sample members, supporting this expectation. While there is no danger of including sample members that could otherwise not be reached and of approaching them in the usual way, more research is needed with respect to bias effects of those who self-select into the survey because they are approached in a different way. It is possible that inclusion of this group of persons biases both localization and dispersion measures of important research variables.

References

ESS. ESS 2002 Documentation Report, Edition 6.1, European Social Survey (2002), `http://ess.nsd.uib.no/ess/round1/`

ESS, ESS 2004 Documentation Report, Edition 3.2, European Social Survey (2004), `http://ess.nsd.uib.no/ess/round2/`

ESS, ESS 2006 Documentation Report, Edition 3.2, European Social Survey (2006), `http://ess.nsd.uib.no/ess/round3/`

ESS, ESS 2008 Documentation Report, Edition 3.01, European Social Survey (2008), `http://ess.nsd.uib.no/ess/round4/`

Häder, M.: Befragungen über das Mobilfunknetz: Eine neue Technik im sozialwissenschaftlichen Methodenarsenal (Surveys over the mobile network: a new technique in the arsenal of social science methods). In: Häder, M., Häder, S. (eds.) Telefonbefragungen über das Mobilfunknetz, pp. 13–16. VS Verlag, Wiesbaden (2009)

[14] http://www.bfs.admin.ch/bfs/portal/de/index/news/00/08.html (in German).

Häder, M., Häder, S.: Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung (Telephone interviews over the mobile phone. Concept, design and implementation of a strategy for data collection). VS Verlag, Wiesbaden (2009)

Groves, R., Couper, M.: Nonresponse in Household Interview Surveys. Wiley, New York (1998)

Jann, B.: Reflections on the Bernese Sampling Plan. Swiss Journal of Sociology 33(2), 307–325 (2007)

Johnson, P., Williams, D.: Comparing ABS vs. Landline RDD Sampling Frames on the Phone Mode. Survey Practice (June 2010)

Klevmarken, N.: Sample Design. In: Börsch-Supan, A., et al. (eds.) Health, Ageing and Retirement in Europe – First results from the Survey of Health, Ageing and Retirement in Europe. Mannheim Research Institute for the Economics of Aging, Mannheim (2005)

Link, M., Battaglia, M., Frankel, M., Osborn, L., Mokdad, A.: Address-based versus Random-Digit-Dial Surveys: Comparison of Key Health and Risk Indicators. American Journal of Epidemiology 164, 1019–1025 (2006)

Lipps, O., Laganà, F., Pollien, A., Gianettoni, L.: National minorities and their representation in Swiss surveys (I): Providing evidence and analysing causes for their under-representation. FORS Working Paper Series, 2011-2, FORS Lausanne (2011)

LUSTAT (Statistical Office Canton of Luzern) (2010). Einwohnerbefragung 2009. lustat aktuell 2010/03

Lutz, G.: Eidgenössische Wahlen 2007 Wahlteilnahme und Wahlentscheid (Federal elections in 2007 Electoral participation and electoral vote). Lausanne: Selects – FORS (2008)

Voorpostel, M., Tillmann, R., Lebert, F., Weaver, B., Kuhn, U., Lipps, O., Ryser, V.-A., Schmid, F., Wernli, B.: Swiss Household Panel Userguide (1999-2008). In Wave 10. FORS, Lausanne (2010)

Part IV
Data Quality

Chapter 14
Factors Affecting Measurement Error in Mobile Phone Interviews

Peter Lynn and Olena Kaminska

14.1 Introduction

Imagine a survey question that asks, “How many hours do you spend watching TV in a typical week?” A respondent may answer “16”. Whether that is the ‘correct’ answer depends on what exactly the researcher intended the question to measure. The respondent may have come up with his answer by thinking along the following lines: “The TV is on for about two hours each evening, so that’s 14 hours, plus I usually watch the sport for a couple of hours on Saturday afternoon, so that makes about 16 in total.” But maybe the researcher only had in mind time spent actually paying attention to the content of a TV broadcast, so having the TV turned on while devoting one’s attention to something different should not be counted. If, during the two hours each evening that the respondent’s TV is on, he is also eating dinner and talking to his children, then he may only be paying attention to the TV for half of that time. So, in the researcher’s terms the ‘correct’ answer would have been nine hours. There may be many people in the survey sample who, like this respondent, have included in their response to this question time that the TV is switched on even if they are not really paying attention to it. This will result in the researcher over-estimating the time people spend attentively watching TV. This error in estimation could have important consequences.

Thus, measurement error arises when a response given to a survey question is not an accurate representation of the respondent’s attitude or behavior that the researcher intended the question to measure. If this happens repeatedly, i.e. many survey respondents give a response that is “in error”, then survey estimates can be affected and incorrect conclusions may be drawn. But the scenario described above is just one of many ways in which measurement error can arise. Measurement error can have many causes and can take many forms.

The likelihood of measurement error arising, and the nature of the error, depends on the characteristics of the respondent. For example, people with regular and attentive TV viewing patterns may be more likely than others to be able to provide an accurate answer. This may apply particularly to retired people. And people with young children may tend to overstate their attentive viewing hours.

This could lead a survey to erroneously conclude that people with young children watch TV as often as retired people.

But the likelihood of measurement error arising, and the nature of the error, also depends on the characteristics of the survey, such as the mode of data collection, the survey procedures and the design of survey questions. In particular, telephone surveys have a number of characteristics that are distinct from other survey modes. Compared with self-completion surveys, the presence of an interviewer provides opportunities to clarify the meaning of a question or the form in which an answer is required. The interviewer also is in control of which questions are asked and in which order. However, a telephone interviewer is more constrained than a face-to-face interviewer. The latter is additionally able to use visual aids and non-verbal communication to assist in the smooth administration of the interview and she may also benefit from the social norms associated with personal contact.

Furthermore, measurement error in interviews conducted on a mobile phone may differ in some respects from those conducted on a fixed phone. Respondents on mobile phones may be more likely to be in a public place and to be engaged in other activities at the time of the interview. This could influence their responses to survey questions. Additionally, the quality of the line may be imperfect due to poor network coverage or interference. This too could influence the understanding of questions and the transmission of responses.

14.2 Factors Affecting Measurement Error in Mobile Phone Interviews: A Framework

A theoretical framework of the factors that may affect measurement error differently between mobile phone and fixed phone interviews is developed by Lynn and Kaminska (2011). Here we broaden the framework to provide a summary of all the factors that may influence measurement error in mobile phone interviews. Figure 14.1 shows the features of a mobile phone interview that determine whether measurement error will arise. Rectangular boxes indicate features of the interviewer-respondent interaction with respect to a specific survey question; the four oval boxes represent the components of the question-answering process; arrows indicate causal effects. For an answer to be of good quality, the respondent has to perceive and understand the question as intended (perception and comprehension stage), retrieve all relevant information (recall stage), evaluate the retrieved information with regard to its relevance to the requested information (evaluation stage) and map the answer to an answer scale and report it to the interviewer (reporting stage) (Tourangeau et al 2000). The measurement error associated with the response to a survey question is determined by the quality of execution of each of the four components of the survey-answering process. The extent to which a question is correctly perceived and understood can be influenced by the respondent's ability to hear as well as by their personal interpretation of the question (Schwarz 1996). Comprehension, recall, evaluation and reporting could all be compromised by a failure of the respondent to pay sufficient attention to the response task or to make sufficient effort to complete the task well (Krosnick 1991). Whether an answer is reported accurately will additionally depend upon the respondent's willingness to reveal the answer (Tourangeau and Smith 1996). Ideally, then, we want the respondent to be able to hear the question and the response options, to interpret these as the researcher intended, to pay attention to the task of responding, to make sufficient effort to complete the task, and to fully reveal his or her answers. However, a number of factors may prevent these ideal circumstances from occurring.

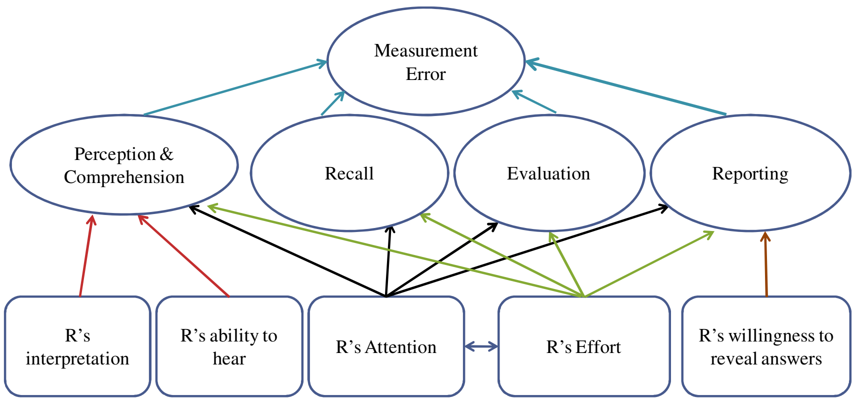


Fig. 14.1 Sources of measurement error in a mobile phone interview

The features of a mobile phone interview that determine whether measurement error will arise are in turn influenced by a number of factors, of which some are attributes of the survey and some are attributes of the respondent. These can either have a direct influence or an indirect influence via the nature of the interview situation (see Figure 14.2). In the following sections we discuss these factors and the implications they may have for survey design and implementation.

14.2.1 The Interview Situation

The circumstances in which a mobile phone interview is carried out can influence the extent to which a survey question is answered in the ideal way. Central considerations are when, where and in what circumstances the respondent is interviewed. These features of the interview situation are likely to determine the quality of the phone line over which the interview is conducted, the likelihood of the respondent being distracted, either visually or aurally, during the interview, the extent and nature of any other tasks in which the respondent may be engaged at the same time as taking part in the interview, and whether other people are present during the interview (AAPOR 2010). The influences of these aspects of the interview situation are summarized in Figure 14.2.

Line quality is a determinant of the respondent's ability to correctly hear the survey question. A failure to hear correctly all utterances that form part of the question or response options leads to the respondent making inferences about the nature of the missed information, which can in turn lead to misinterpretation of the question (Kellogg 2007). The respondent could thus be answering a different question to the one that was posed.

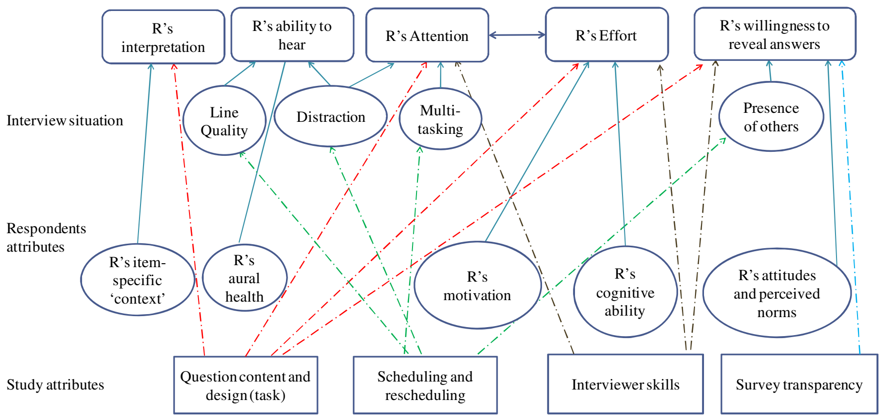


Fig. 14.2 The role of respondent attributes and study attributes

Distraction diverts the respondent's attention from the question answering task, which may lead to mistakes or shortcuts (Kellogg 2007: 75). Distraction could be either aural or visual and could be either active (somebody trying to attract the respondent's attention or to talk to them) or passive (background noise or background movement). While this is true of any kind of survey interview, the likelihood of distraction may be greater for interviews conducted on a mobile phone, due to the greater range of possible locations and circumstances in which an interview may be conducted (AAPOR 2010, Kennedy 2010).

Multi-tasking can similarly divert or divide the respondent's attention. There are many different kinds of tasks that a respondent could be performing while being interviewed and it is likely that different tasks demand a different degree of attention. For example, cooking is likely to require less attention than driving a car. It may also make a difference whether the other task is one over which the respondent more or less has control (e.g. housework) or not (e.g. minding children). In the first case the respondent may be able to pause from the other task when a survey question is particularly demanding. But in the second case, unpredictable demands on his or her attention could occur at any time. It is likely that different sample subgroups are prone to multi-tasking to different degrees, and to different types of multi-tasking.

The presence of other people (within earshot of the respondent) may make the respondent unwilling to reveal certain answers or to respond in a way that reveals the nature of the conversation (Schneiderat and Häder 2007). We suggest that the effect is likely to be stronger if those present are known to him or her (i.e. friends or family rather than strangers) and if they are proximate rather than distant.

Other relevant practical considerations regarding the survey situation may include whether or not the respondent has to pay a fee to receive incoming calls (AAPOR 2010, Vicente et al. 2009). This may be particularly likely if the respondent is outside of his or her provider country ("roaming") at the time the survey call is received. For most providers in most countries roaming charges are greater than standard charges. If the sample member agrees to participate, liability for charges may provide an incentive to hurry their answers and take cognitive shortcuts.

14.2.2 Respondent Attributes

Several respondent attributes can influence the measurement error process in mobile phone interviews. The respondent's lifestyle will interact with the survey calling schedule to determine the interview situation (Döring 2009). Specifically, where and with whom the respondent tends to be at the times that survey interviewers call will determine the potential for measurement error to be affected by distraction, multi-tasking, the presence of others and line quality. Also, the respondent's motivation to respond to the survey question (via saliency, interest, etc) and his or her cognitive ability will, in combination, influence the extent to which they make sufficient effort to respond accurately (Krosnick 1991).

Additionally, ability to adequately hear the question is influenced by the respondent's aural health (directly) and perhaps also by his or her choice of phone service provider (indirectly, via the obtained line quality). Finally, the respondent's attitudes and perceptions of social norms will influence willingness to reveal true answers. When respondents are unwilling to reveal a true answer there is the potential for social desirability bias (Tourangeau et al. 2000).

14.2.3 Survey Attributes

Survey design can have a strong influence on the interview situation and hence on measurement error. For example, the survey call scheduling protocol will determine the days and times that call attempts are made (Weeks et al. 1987), which will in turn influence the situations in which sample members are successfully contacted and interviewed. This includes the survey's protocol for rescheduling an interview if the respondent's current situation is deemed inappropriate (for example, if they are driving a motor vehicle or operating machinery, or even if they are in a public place or with friends).

Design features such as whether the caller's number is displayed or withheld (McCutcheon et al. 2010) and whether or not advance letters are sent (Link and Mokdad 2005) have been shown to be associated with participation rates. This most likely is because these features establish the credentials of the survey organization and thereby influence the level of trust that the respondent has in the survey organization. Consequently, the features could also influence respondent willingness to respond honestly to sensitive questions. Similarly, the experience, ability and training of the survey interviewers could be influential, both on the attention and effort made by the respondent and on the respondent's willingness to reveal socially undesirable answers (van der Zouwen et al. 2004).

A second sphere of influence of survey design is via the content and design of the survey question (Bradburn and Sudman 2004). This determines the difficulty of the task of responding to the question, which may in turn influence the likelihood of the respondent being able and willing to devote sufficient effort to responding. The more difficult the task, the less likely it is that the respondent will make sufficient effort.

14.3 The Nature of Measurement Error

The nature of measurement error is determined by interactions between the three sets of factors described above: the interview situation, respondent attributes and survey attributes. These interactions are depicted in Figure 14.2 by the multiple arrows pointing towards each of the five rectangular boxes representing the relevant features of the respondent-interviewer exchange. For example, the respondent's ability to hear the question can be seen to be determined by the combination of three factors: line quality, his or her aural health, and the level of distraction.

The processes by which survey attributes, respondent attributes and the interview situation affect measurement error can be broadly thought of as being of three types, namely context, satisficing and sensitivity. Many of these processes are common to all modes of survey data collection and are not specific to mobile phone interviews. Nevertheless, they may be amongst the most important sources of measurement error in mobile phone interviews.

14.3.1 Context

The context in which a survey question is answered is determined by the interaction between the survey design, on the one hand, and the knowledge, experiences and predisposition of the respondent (what we refer to in Figure 14.2 as the respondent's item-specific context), on the other. For example, studies have shown that vague quantifiers (very often, quite likely, usually, good, etc.) tend to be interpreted as having a relative rather than absolute meaning (Schaeffer 1991). Unfortunately, different respondents have different frames of reference. For example, some may implicitly compare themselves to others of their own age or of their own social background, while others may compare their current situation with previous periods in their own life. This causes differences between respondents in the way that a question is interpreted and understood (Schaeffer 1991). Similarly, it has been shown that for questions about an inherently quantifiable concept (frequency, lengths of time, money amounts) the response categories offered tend to invoke a perception of the likely true distribution, which in turn influences the response given, as most respondents think in relative terms (Gaskell et al. 1994, Schwarz et al. 1991).

The understanding of the meaning of the question is inferred not only from the wording of the actual question and its scale, but also from the interview context, for example from the nature and content of immediately prior questions. Thus, Schwarz et al. (1991) showed that the correlation between answers to two questions on marital satisfaction and life satisfaction was much stronger when the marital satisfaction question was asked first than when the question order was reversed. This suggests that respondents were more likely to take their experiences of their marriage into account when considering their overall satisfaction with life if the previous question has brought those experiences to mind. Schwarz et al. (1991) additionally showed that the effect of the prior marital satisfaction question was reduced when earlier questions in the interview concerned satisfaction with other life domains. Other examples of this kind can be found in Schuman and Presser (1981).

14.3.2 Satisficing

Satisficing refers to performing a task in a way which is sub-optimal, but adequate. Krosnick (1991) describes why and how this might happen in the context of a survey interview. The idea is that in order to reduce the burden of responding to a question, respondents take shortcuts of one kind or another. These shortcuts can relate to any of the four components of the question-answering process and can therefore introduce errors of comprehension, of recall, of evaluation or of reporting. The shortcutting may be more or less conscious on the part of the respondent, but the extent to which it occurs is hypothesized to result from the interaction between the difficulty of the response task, the ability of the respondent and the motivation of the respondent (Krosnick 1991).

In Figure 14.2, in line with Krosnick's theory, satisficing (the inverse of effort) is influenced by two respondent attributes (motivation and cognitive ability) and by question content and design (specifically question difficulty). In addition, effort is influenced by interviewer skills (Cannell and Fowler 1963). For practitioners it is important to pay attention to the factors which are not under their control (respondents attributes) and those which can be changed (study attributes). Given that two of the four features that interact to determine effort are study attributes, careful thought should be given to ways of increasing the effort of a respondent by manipulating study attributes. We return to this idea in section 14.3.1.

Insufficient effort can affect responses in a number of ways. One is a failure to differentiate between related items that have the same response options, such as a battery of attitude questions. Krosnick et al (1996) and Kaminska et al (2010) both find evidence of such non-differentiation as a result of satisficing. Another form of shortcutting is the tendency to select either one of the first response options presented (primacy effect, Krosnick and Alwin 1987) or one of the last ones (recency effect), depending on which requires less effort. The latter may be the dominant effect in phone interviews (McClendon 1991, Schuman and Presser 1981), though only if steps are taken to ensure that respondents listen to the response options until the end of the list is reached (Holbrook et al. 2007). A common finding is that respondents are more likely to give responses such as "don't know" or "no opinion" if these options are presented explicitly. This has been interpreted as a form of satisficing (Krosnick et al. 1996) and may be a good example of how question design can influence measurement. Another commonly observed response phenomenon is the tendency of respondents to round responses to open questions requiring numeric answers (Huttenlocher et al. 1990, Pudney 2008), resulting in heaping in the data at multiples of 5, 10, 100, etc.

A tendency of respondents to agree to assertions regardless of content is referred to as acquiescence bias. There is considerable empirical evidence of the existence of such a tendency, which can be the result of either of two different underpinning psychological processes, agreeing and conformity (Bentler et al. 1971). The dominance of one or other process is situation-specific.

Asking respondents to report information relating to other household or family members (proxy reporting) has been shown to often lead to greater measurement error than asking for the same information by self-report (Blair et al. 1991). This is likely to be due to lower levels of knowledge, weaker encoding of events in memory, or weaker motivation to respond accurately.

A considerable research literature discusses the effects of memory on respondent's ability to recall information required to answer a survey question. Many studies show that events and details are more likely to be correctly remembered the more recent the event took place, the more salient it is to the respondent and the fewer intervening similar events have been experienced (for summaries of the research evidence see Tourangeau et al. 2000, Sudman and Bradburn 1974). Survey designers can therefore reduce measurement error by asking only about more recent periods (Sudman and Bradburn 1973).

14.3.3 Sensitivity

Whether a respondent is willing to reveal an answer depends on at least three things. The topic of the question may be considered intrusive, regardless of the respondent's true position. The respondent's position on the issue in question may be considered socially undesirable. Or the respondent may be concerned about risks of disclosure of information. Questions with any or all of these characteristics are generally referred to as sensitive questions. Such questions are prone to respondents editing their answers prior to reporting them.

Several studies have demonstrated that, especially in interviewer-administered surveys, survey respondents tend to conceal socially undesirable views and behaviors (DeMaio 1984). For example, illicit drug use is under-reported (Aquilino 1992, 1994, Gfroerer and Hughes 1991), as are alcohol consumption (Aquilino 1994), depression (Aquilino 1998) and abortion (Jobe et al. 1997), while men over-report numbers of sex partners (Tourangeau and Smith 1996). Respondent's willingness to report their answers accurately and honestly is influenced by the perceived privacy of the survey setting, the perceived legitimacy of the survey and rapport between the interviewer and respondent (Holbrook et al. 2003).

Certain survey topics may be felt by some respondents to be intrusive, even though their own responses would not reveal socially undesirable views or behavior. Examples may include sex, income and religion. Reluctance to respond to intrusive questions can introduce measurement error (Willis 1997, Schaeffer and Presser 2003). Similarly, respondents may edit their answers if they fear they could be revealed to some third party (Couper 1998, Couper et al. 2010).

A respondent's perception of the sensitivity of a question will depend on the interaction between the content of the question and their own attitudes and perceived norms. However, whether a respondent is willing to report a socially undesirable answer or to provide sensitive information will depend on interviewer skills, whether others are present and survey transparency. These are the other three factors identified in Figure 14.2 as influencing willingness to reveal answers. Other possible reasons for a respondent to edit their answer before reporting it to the interviewer may include a desire to appear consistent or a desire to not offend the interviewer (Smith 1983).

14.3.4 Specific Features of Mobile Phone Interviews

Mobile phone interviews have some unique features. We discuss these below and the evidence regarding their impact on measurement error. But mobile phone interviews also share many important features with fixed phone interviews. Thus, it is important to bear in mind that the cumulative knowledge about measurement error in phone interviews, on which there is a considerable body of research evidence, applies also to mobile phone interviews.

A number of features of telephone interviewing have distinctive consequences for measurement error. The absence of visual aids such as show cards generally makes the response task more difficult than it might otherwise be, resulting in a heightened risk of satisficing behavior by respondents. The relatively fast pace of telephone interviews may exacerbate this effect. In particular, telephone interview respondents have a tendency to pick the last response option from a list (recency effect) and to pick extreme options (saliency effect) (Schwarz 1996). Additionally, as the interviewer cannot see the respondent, there is a heightened risk that the respondent may be multi-tasking during the interview, a situation which would detract attention from the question-answering task and induce satisficing behavior (Lavrakas et al. 2007). On the other hand, the scope for context effects is limited as the respondent is not aware of questions subsequent to the current one.

Additional features of mobile phone interviews as distinct from fixed line phone interviews can be thought of in three categories:

- Situational features of the interview that are likely to have a negative impact on measurement. Specifically, it has been speculated that mobile phone interviews have a heightened risk of suffering from respondent multitasking, distraction, the presence of others and poor line quality;
- Technical features of the interview situation that are likely to have a positive impact on measurement. Specifically, there is usually no risk of another household member listening in to the conversation on another line and it is often relatively easy for the respondent to move to a location that is quieter and out of earshot of others;

- Psychological features of mobile phones that may have a positive or negative impact on measurement. Specifically, mobile phones are generally perceived as personal devices with which the user has an intimate relationship, unlike shared household or office phones.

However, we would note that these features are not universally present in one mode and universally absent in the other. Rather, there may be a greater tendency for them to be present in one mode. To illustrate this point, one needs only realize that wireless handsets used with fixed lines share some important characteristics with mobile phones (Lynn and Kaminska 2011). We would also note that it is not always possible for the researcher to infer from the number dialed whether a respondent is talking on a mobile phone or a fixed phone. Calls to fixed lines can be automatically forwarded to a mobile number, and vice versa, often without the caller realizing. We now review the evidence regarding each of these three categories of features in turn.

It is often suggested that mobile phone interviews are more likely to be conducted out of the home, in public places, and while respondents are engaged in other tasks (Lavrakas et al. 2007, Brick et al. 2007, Schneiderat and Häder 2007, Witt et al. 2009). However, our framework (see Figure 14.2) suggests that the extent to which this is true should depend on the survey scheduling and rescheduling practices. This insight may explain why studies vary in the extent to which they find that mobile phone respondents are more likely than fixed phone respondents to be out of the home when interviewed (Brick et al. 2007, Kennedy 2010, Lavrakas et al. 2010, Lynn and Kaminska 2011). Related to this, Brick et al. (2007) found that neither contact rate nor completion rate varied much by time of day for calls to mobile phones, whereas calls to fixed phones were more likely to result in contact and completion at weekends and least likely to do so on weekdays between 9 am and 7 pm. Similarly, findings on multitasking rates also differ between studies. Kennedy (2010) found that mobile phone respondents in the US were most likely to be "watching TV, video gaming, or on computer" (45 %) or "eating, drinking, preparing a meal" (16 %), with smaller numbers in a car (8 %), working (7 %) or doing household chores (7 %). Also in the US, Lavrakas et al. (2010) found 7 % watching TV, 6 % driving a car, 6 % using a computer, 6 % doing household chores, 5 % eating or drinking and 3 % watching a child (and concluded that 16 % were engaged in what was judged to be a highly distracting other behavior while being interviewed). In our own study (details in Lynn and Kaminska 2011) using data from Hungary we found 8 % watching TV, 7 % doing housework or cooking and 1 % minding children. Pew Research Center (2006), another US study, reported that in 20 % of cases the interviewer thought that the respondent had been doing another activity during the interview.

Only two studies, to our knowledge, have attempted to measure specifically the extent of distraction in mobile phone interviews, both by asking interviewers to report whether the respondent seemed to be distracted in any way during the interview. In our study (Lynn and Kaminska 2011) we found that 3 % said 'yes' and a further 11 % said 'maybe'. Pew Research Center (2006) found that 1 % of interviewers thought the respondent to be 'very distracted' and 7 % 'somewhat distracted'.

Kennedy (2010) asked respondents to rate the clarity of the phone connection (line quality): 33 % responded ‘perfect, like we were talking face to face’, 51 % ‘very good’, 10 % ‘good’, 5 % ‘fair’ and 0 % ‘poor, like you could barely hear me at times’. These proportions were very similar to those reported by fixed phone respondents.

We are unaware of any study that has measured whether other people are within earshot during mobile phone interviews. Indirect evidence comes from Brick et al (2007), who report that 5 % of mobile phone respondents were in someone else’s home during the interview (we can conjecture that someone else is likely to at least have been in the same home in a majority of those cases) and from our own study (Lynn and Kaminska 2011) in which we found that in 2 % of cases the interviewer felt that someone else had influenced the respondent’s answers (and must therefore have been present). However, these findings tell us very little about the overall proportion of mobile phone interviews likely to be conducted within earshot of others or about whether those others are known to the respondent and how they are related to the respondent.

We are similarly unaware of any studies of the effect of the psychological disposition towards one’s mobile phone.

The net effect of these differences in the interview situation between mobile and fixed phone interviews should result in differences in measurement error. A number of studies have compared indicators of measurement error between the two modes. In general, differences have been small or nonexistent (Dipko et al. 2005, Brick et al. 2007, Kennedy 2010, Lynn and Kaminska 2011, Vicente et al. 2009). On this issue, the AAPOR Cell Phone Task Force (2008) concluded that “currently available evidence … is relatively sparse and mixed” (AAPOR 2008: 40). Since that report appeared, two randomly-assigned experiments have been published (Kennedy 2010, Lynn and Kaminska 2011), but these still do not provide a clear picture of differences in measurement between mobile and fixed phone interviews. These studies have not in any case provided the means to identify the source of any differences in measurement error (for example, whether lower quality responses in one mode were caused by greater distraction, more multi-tasking or lower line quality).

In summary, a lot is known about processes that lead to measurement error in telephone surveys and about design and implementation features of surveys that can be used to minimize the likelihood of such errors arising. However, a lot remains to be learned about the specific features of mobile phone interviews that distinguish them from other phone interviews. Limited evidence exists regarding the extent to which mobile phone interviews are more or less likely than other phone interviews to be subject to the influence of multitasking, distraction, the presence of others and poor line quality. And much of that evidence suffers from a lack of experimental control in the study design: in particular, measurement error is often confounded with nonresponse error. And direct evidence of the impact of those features on measurement error is severely lacking. There is therefore an urgent need for carefully designed methodological studies that will provide better understanding of the measurement error processes identified in our framework (see Figure 14.2), of the circumstances in which the processes take effect, and of the nature and magnitude of the effects. Finally, our framework clearly identifies why researchers should expect measurement error to vary between survey topics, study populations, and even between survey organizations. This happens via the interaction between study attributes and respondent attributes. Consequently, replication of experiments is needed in different contexts.

14.4 Current Best Practice and Future Steps

14.4.1 Survey Design

Figure 14.2 indicates that the means by which a survey researcher can minimize measurement error is via control of the relevant study attributes (the bottom row of factors). However, the effects of these attributes are via interactions with respondent attributes. Thus, to identify the best design the researcher must have some idea of the likely attributes of members of the study population.

A central role is played by the content and design of the individual survey items. This determines the task that each respondent is being asked to perform in order to answer the question. Plenty of guidance is available on how to design 'good' survey questions (Czaja and Blair 2005, Dillman et al. 2009, Fowler 1995, Groves et al. 2009). What may be particularly important to remember in the case of mobile phone interviews is that the attention given to the question-answering task by the respondent may be compromised by distraction and multi-tasking. This will affect the respondent's ability to comprehend the question, to recall required information and to evaluate that information. It may therefore be important to minimize the amount of attention that is required in order to answer the question. This implies keeping lists of response options short, keeping recall reference periods short, using recent rather than distant recall reference periods, keeping question wording and information short and clear, and considering splitting complex questions into multiple simpler questions.

If the topic of the survey questions is sensitive, in the sense described in 14.3.3 above, then it may be advisable to design the questions in a way which minimizes the possible impact of the presence of other people. For example, the respondent could be asked to read out the code letter or number that applies rather than having to verbalize their answer in a way that might reveal its nature to anyone within earshot. For particularly sensitive questions it may be best to consider a self-completion mode of data collection rather than an interviewer-administered mode. One way of doing this just for a subset of the questions might be to switch to a form of self-completion that can be carried out on the mobile phone. This might be an IVR mode, a simple web interview if the phone has the appropriate technology, or via SMS (though switching modes mid-interview can be problematic) (Tourangeau et al. 2002, Couper et al. 2003, Saukshag 2010).

In the absence of strong research evidence about effects of location and multitasking on measurement error, we believe that it is advisable to be cautious. In particular, we suggest that researchers should design rescheduling protocols that are driven by explicit questions regarding the respondent's location and activities. Many surveys already ask, often for ethical or legal rather than scientific reasons, whether a respondent is driving a motor vehicle when they answer the phone to the interviewer. If they are, then the call is terminated and attempted again later (some surveys may only terminate the call if the respondent is not using a hands-off device; we suggest that it should be terminated in all cases). The question may be extended to include other activities such as operating machinery. We think this is wise. It could be argued that engagement in any task likely to divert a considerable amount of the respondent's attention from the interview should be grounds for rescheduling. This would require the initial question about driving to be immediately followed by another question asking the respondent whether he or she is currently engaged in some other activity and, if so, what. An alternative would be simply to ask the respondent if this is a good moment and, if not, to find a better time for an interview, but such an approach would need careful development and testing in order to avoid simply giving sample members an easy way to get out of cooperating. For interviews containing important questions that could be susceptible to social desirability bias if the respondent is in the presence of others, it may be worthwhile asking upfront whether anyone known to the respondent is in earshot and, if so, to reschedule the interview for a time when the respondent can be alone.

Another good practice is for interviewers to be trained to be sensitive to changes in the respondent's circumstances during the course of the interview. Such changes may be detectable through changes in the respondent's tone of voice, interruptions, or background noise, for example. Interviewers may want to adapt their question delivery appropriately (e.g. speaking louder and more slowly if the respondent has just moved to somewhere noisy) or it may be appropriate for them to re-ask the question(s) about multitasking and presence of others, which could therefore lead to a break-off and rescheduling part-way through the interview. Interviewers could also be trained in the case of a sudden change to tell the respondent that they will be happy to wait a moment if the respondent needs to attend to something. This may put the respondent at ease and avoid him or her answering the next questions while less attentive or rushing.

14.4.2 Methodological Research

In this chapter we have identified the paucity of empirical evidence about measurement error processes in mobile phone interviews. The aim of future methodological research should be to contribute to understanding of how and why errors appear, what impacts they have, and how they can be controlled through survey design and implementation features. To this end, we suggest that three kinds of studies are needed:

Between-mode comparisons of features of the interview situation. To identify the extent to which interview situations likely to give rise to increased measurement error are relatively prevalent in mobile phone interviews. Such a design is appropriate when we believe that the prevalence of the feature could differ between modes. For example, we might reasonably suppose that the effect of spousal presence on tendency to edit responses is independent of phone type (fixed or mobile), but that a spouse is less likely to be present with mobile phone interviews. To test this, we would need a two-treatment randomized design in which sample members were randomly allocated to be interviewed either on a mobile phone or on a fixed phone. Outcomes of interest would be a measure, for each treatment, of the proportion of interviews in which a spouse was present, plus knowledge of the other relevant attributes of respondents whose spouse was present.

Within-mode comparisons of features of the interview situation. To identify the effect of a feature on measurement, we need a randomized design in which the treatments correspond to different levels of the feature, but all other relevant aspects of design and implementation, including mode, are held constant. For example, we could experimentally manipulate whether or not the spouse was present during the interview. If we believe that the effect of the feature (spousal presence) on measurement could be mode-specific then the experiment must be carried out entirely using mobile phone interviews. If we do not believe the effect to be mode-specific (but we wish to assess the impact on survey estimates and do not have any prior information about effect size) then the experiment does not to control the mode (e.g. it could be carried out using fixed phone interviews as well). Strong candidates for study in this way are line quality, multi-tasking and the presence of others, though the design challenges are considerable.

Within-mode comparisons of survey design features. Once we have identified features of the interview situation that appear to warrant control we should seek to identify design features that could provide the required control. For example, to minimize the proportion of interviews in which the spouse is present, the call scheduling algorithm could be altered in a particular way (or a rescheduling protocol introduced). To evaluate the impact of the alternative protocols, a randomized design will be needed in which all relevant features other than the feature being studied are held constant between the treatments. Many classic survey methods experiments are of this kind, including the research cited in this chapter into order effects, context effects and the use of vague quantifiers. Strong candidates for study in this way are question and questionnaire design features that aim to simplify the respondent’s task, rescheduling questions and protocols, elements of interviewer training, and ways of communicating the legitimacy and standards of the survey organization (such as advance letters or SMS texts).

14.4.3 Conclusion

In this chapter we have provided a structured framework that identifies the set of features of a mobile phone interview that can influence measurement error. The framework also indicates the ways in which these features interact and the components of the question-answering process that are affected by these interactions. To a large extent, existing knowledge about measurement error in surveys can be mapped onto this framework to provide clear indications of the issues that designers of mobile phone surveys should address. Researchers should therefore not be afraid of mobile phone surveys. They do not represent a complete leap into the dark. Rather, they require us to reassess existing knowledge in a slightly different light. With careful thought, survey procedures and survey instruments can be designed that will yield high quality data from mobile phone interviews.

However, there are some aspects of mobile phone interviews about which we know relatively little. More research is urgently needed in order to help researchers to identify appropriate design and implementation features to combat measurement error. In reality, many of these under-researched issues apply also to fixed phone interviews. Recent interest in mobile phone interviews has provided the catalyst for researchers to pay renewed attention to issues of distraction, multi-tasking and the presence of others, simply because they are amongst the characteristics of the interview situation that are expected to differ between fixed and mobile phone interviews, but that is not to say that they are absent in fixed phone interviews. Thus, the research agenda outlined in this chapter should benefit not only mobile phone surveys but also phone surveys in general. We have little doubt that the methodology of mobile phone interviewing will continue to receive growing attention in the near future and that researchers will soon be far more knowledgeable about how to design and carry out mobile phone surveys.

References

AAPOR (American Association of Public Opinion Research - Cell Phone Task Force) (2008). Guidelines and Considerations for Survey Researchers when Planning and Conducting RDD and Other Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers, `http://www.aapor.org` (accessed December 30, 2009)

AAPOR (American Association of Public Opinion Research - Cell Phone Task Force) (2010). New Considerations for Survey Researchers when Planning and Conducting RDD Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers, `http://www.aapor.org` (accessed January 24, 2011)

Aquilino, W.: Telephone versus face-to-face interviewing for household drug use surveys. International Journal of the Addictions 27, 71–91 (1992)

Aquilino, W.: Interview mode effects in surveys of drug and alcohol use: a field experiment. Public Opinion Quarterly 58, 210–240 (1994)

Aquilino, W.: Effects of interview mode on measuring depression in younger adults. Journal of Official Statistics 14, 15–29 (1998)

Bentler, P.M., Jackson, D.N., Messick, S.: Identification of Content and Style: a Two-Dimantional Interpretation of Aquiescence. Psychological Bulletin 76(3), 186–204 (1971)

Blair, J., Menon, G., Bickart, B.: Measurement Effects in Self vs. Proxy Responses to Survey Questions: An Information Processing Perspective. In: Biemer, P.P., et al. (eds.) Measurement Errors in Surveys, pp. 145–166. Wiley, New York (1991)

Bradburn, N.M., Sudman, S.: The Current Status of Questionnaire Research. In: Biemer, P.P., et al. (eds.) Measurement Errors in Surveys, pp. 29–40. John Wiley & Sons, Hoboken (2004)

Brick, J.M., Brick, P.D., Dipko, S., Presser, S., Tucker, C., Yuan, Y.: Cell Phone Survey Feasibility in the U.S.: Sampling and Calling Cell Numbers versus Landline Numbers. Public Opinion Quarterly 71, 23–29 (2007)

Cannell, C.F., Fowler, F.J.: Comparison of Self-enumerated Procedure and a Personal Interview: a Validity Study. Public Opinion Quarterly 27, 250–264 (1963)
Couper, M.P.: Measuring survey quality in a CASIC environment. In: Proceedings Survey Research Methods Section of American Statistical Association, pp. 41–49 (1998)
Couper, M.P., Singer, E., Tourangeau, R.: Understanding the Effects of Audio-CASI on Self-Reports of Sensitive Behavior. Public Opinion Quarterly 67, 385–395 (2003)
Couper, M.P., Singer, E., Conrad, F.G., Groves, R.M.: Experimental Studies of Disclosure Risk, Disclosure Harm, Topic Sensitivity, and Survey Participation. Journal of Official Statistics 26, 287–300 (2010)
Czaja, R., Blair, J.: Designing surveys: a guide to decisions and procedures, vol. 2. Pine Forge Press, Thousand Oaks (2005)
DeMaio, T.J.: Social Desirability and Survey Measurement: A Review. In: Turner, C.F., Martin, E. (eds.) Surveying Subjective Phenomena, vol. 2, pp. 257–282. Russell sage Foundation, New York (1984)
Dillman, D.A., Smyth, J.D., Christian, L.M.: Internet, Mail and Mixed-Mode Surveys. Wiley, New York (2009)
Dipko, S., Brick, P., Brick, J.M., Presser, S.: An investigation of response difference between cell phone and landline interviews. In: Paper presented at the Annual Conference of the American Association for Public Opinion Research, Miami Beach, Florida, May 12-15 (2005)
Döring, N.: Psychological Aspects of Interviewing by Cellular Telephone. In: Maxl, E., et al. (eds.) Mobile Market Research, vol. 12, Herbert von Halem, Köln (2009)
Fowler, F.J.: Improving Survey Questions: Design and Evaluation. Sage Publication, Inc. (1995)
Gaskell, G., O'Muircheartaigh, C., Wright, D.: Survey questions about the frequency of vaguely defined events: the effect of response alternatives. Public Opinion Quarterly 58, 241–254 (1994)
Gfroerer, J.C., Hughes, A.L.: The Feasibility of Collecting Drug Abuse Data by Telephone. Public Health Reports 106, 384–393 (1991)
Groves, R.M., Fowler Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R.: Survey Methodology, vol. 2. Wiley, Hoboken (2009)
Holbrook, A.L., Green, M.C., Krosnick, J.A.: Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. Public Opinion Quarterly 67, 79–125 (2003)
Holbrook, A.L., Krosnick, J.A., Moore, D., Tourangeau, R.: Response Order Effects in Dichotomous Categorical Questions Presented Orally. Public Opinion Quarterly 71, 325–348 (2007)
Huttenlocher, J., Hedges, L.V., Bradburn, N.M.: Reports of Elapsed Time: Bounding and Rounding Processes in Estimation. Journal of Experimental Psychology: Learning, Memory, and Cognition 16, 196–213 (1990)
Jobe, J.B., Pratt, W.F., Tourangeau, R., Baldwin, A., Rasinski, K.: Effects of Interview Mode on Sensitive Questions in a Fertility Survey. In: Lyberg, L., et al. (eds.) Survey Measurement and Process Quality. Wiley, New York (1997)
Kaminska, O., McCutcheon, A.L., Billiet, J.B.: Satisficing among Reluctant Respondents in a Cross-National Context. Public Opinion Quarterly 74, 956–984 (2010)
Kellogg, R.T.: Fundamentals of Cognitive Psychology. Sage Publications, Los Angeles (2007)

Kennedy, C.: Evaluating the Effects of Screening for Telephone Service in Dual Frame RDD Surveys. Public Opinion Quarterly 71, 750–771 (2007)
Kennedy, C.: Use of cognitive shortcuts in landline and cell phone surveys. In: Paper presented at the Annual Conference of the American Association of Public Opinion Research, Chicago, Illinois (2010)
Krosnick, J.A.: Response strategies for coping with the cognitive demands of attitude measures in surveys. Applied Cognitive Psychology 5, 213–236 (1991)
Krosnick, J.A., Alwin, D.F.: An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opinion Quarterly 51, 201–219 (1987)
Krosnick, J.A., Narayan, S., Smith, W.R.: Satisficing in surveys: initial evidence. In: Braverman, M.T., Slater, J.K. (eds.) Advances in Survey Research, pp. 29–44. Jossey-Bass, San Francisco (1996)
Lavrakas, P.J., Shuttles, C.D., Steeh, C., Fienberg, H.: The State of Surveying Cell Phone Numbers in the United States. Public Opinion Quarterly 71, 840–854 (2007)
Lavrakas, P.J., Tompson, T.N., Benford, R., Fleury, C.: Investigating Data Quality in Cell Phone Surveying. In: Paper presented at the 65th Annual Conference of the American Association for Public Opinion Research, Chicago (2010)
Link, M.W., Mokdad, A.: Advance Letters as a Means of Improving Respondent Cooperation in Random Digit Dial Studies: A Multistate Experiment. Public Opinion Quarterly 69, 572–587 (2005)
Lynn, P., Kaminska, O.: The impact of mobile phones on survey measurement error. ISER Working Paper 2011-07 (2011)
McClendon, M.J.: Acquiescence and Recency Response-Order Effects in Interview Surveys. Sociological Methods and Research 20(1), 60–103 (1991)
McCutcheon, A.L., Callegaro, M., Ludwig, J.: Who's Calling? The Impact of Caller ID on Telephone Survey Response. Field Methods 22, 175–191 (2010)
Pew Research Center. The Cell Phone Challenge to Survey Research (2006), `http://people-press.org/report/276/the-cell-phone-challenge-to-survey-research` (accessed September 8, 2010)
Pudney, S.: Heaping and leaping: Survey response behavior and the dynamics of self-reported consumption expenditure. ISER Working Paper 2008-09 (2008), `http://www.iser.essex.ac.uk/publications/working-papers/iser/2008-09`
Saukshag, J.: An Examination of Mode Switch Nonresponse in a "Recruit-and-Switch" Survey. Presented at International Workshop on Household Survey Nonresponse. Nürnberg (August 2010)
Schaeffer, N.C.: Hardly ever or constantly? Group comparisons using vague quantifiers. Public Opinion Quarterly 55, 395–423 (1991)
Schaeffer, N.C., Presser, S.: The Science of Asking Questions. Annual Review of Sociology 29, 65–88 (2003)
Schneiderat, G., Häder, S.: Anlage und Vorstudien des DFG-Projektes "Telefonbefragungen in der Allgemeinbevölkerung über das Mobilfunknetz" (System and preliminary studies of the DFG project. "telephone surveys in the general population over mobile phone"). In: Gabler, S., Häder, S. (eds.) Mobilfunktelefonie – Eine Herausforderung für die Umfrageforschung, Gesis, Mannheim (2007)
Schwarz, N.: Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation. Lawrence Erlbaum Associates, Inc., Publishers (1996)

Schwarz, N., Strack, F., Hippler, H.-J., Bishop, G.: The impact of administration mode on response effects in survey measurement. Applied Cognitive Psychology 5, 193–212 (1991)

Schuman, H., Presser, S.: Questions and answers in attitude surveys: Experiments on question form, wording, and context. Academic Press, New York (1981)

Smith, T.W.: An Experimental Comparison of Clustered and Scattered Scale Items. Social Psychology Quarterly 46, 163–168 (1983)

Sudman, S., Bradburn, N.: Effect of time and memory factors on response in surveys. Journal of the American Statistical Association 68, 805–815 (1973)

Sudman, S., Bradburn, N.: Response Effects in Surveys. Aldine, Chicago (1974)

Tourangeau, R., Rips, L., Rasinski, K.: The Psychology of Survey Response. Cambridge University Press, Cambridge (2000)

Tourangeau, R., Smith, T.W.: Asking sensitive questions: The impact of data collection mode, question format and question context. Public Opinion Quarterly 60, 275–304 (1996)

Tourangeau, R., Steiger, D.M., Wilson, D.: Self-administered questions by telephone: evaluating interactive voice response. Public Opinion Quarterly 66, 265–278 (2002)

Van der Zouwen, J., Dijkstra, W., Smit, J.H.: Studying Respondent-Interviewer Interaction: the Relationship between Interviewing Style, Interviewer Behavior, and Response Behavior. In: Biemer, P.P., et al. (eds.) Measurement Errors in Surveys, pp. 419–437. John Wiley & Sons, Hoboken (2004)

Vicente, P., Reis, E., Santos, M.: Using mobile phones for survey research: a comparison with fixed phones. International Journal of Market Research 51, 613–633 (2009)

Weeks, M.F., Kulka, R.A., Pierson, S.A.: Optimal call scheduling for a telephone survey. Public Opinion Quarterly 51, 540–549 (1987)

Willis, G.B.: The Use of the Psychological Laboratory to Study Sensitive Survey Topics. In: Harrison, L., Hughes, A. (eds.) The Validity of Self Reported Drug Use: Improving the Accuracy of Survey Estimates. NIDA Research Monograph 167, Washington, DC: Supt. of Docs., U.S. Govt. Printing Office (1997)

Witt, L., ZuWallack, R., Conrey, F.: Out and About: An Evaluation of Data Quality in Cell Phone Surveys. In: Paper presented at the Annual Conference of the American Association of Public Opinion Research, Hollywood, FL (2009)

Chapter 15

Telephone Surveys via Landline and Mobile Phones: Mode Effects and Response Quality

Mike Kühne and Michael Häder

15.1 Introduction

Telephone surveys play a central role in survey-based research. This becomes clear when considering the widespread use and data quality of telephone surveys (Lepkowski et al. 2008). At the same time, the use of telephone surveys is influenced by social and technological changes. Most prominently, mobile phone technology poses a serious challenge to survey research worldwide (AAPOR 2008, Steeh and Piekarski 2008, Häder and Häder 2009). There is increasing concern that a considerable number of households that are currently supposed to be covered by landline sampling frames can, in fact, not be reached by landline telephone surveys because they rely predominantly on mobile phones (Brick et al. 2006, Keeter et al. 2007). This "mobile-mostly" population currently represents 15 % of adults in United States households. Boyle and Lewis (2009), nevertheless, were able to show that households that consider themselves "mobile-mostly" are still likely to answer their landline telephones. A growing number of mobile phone users have substituted their residential landline telephone for a mobile phone. These so called "mobile-phone-only households" may have implications for the representativeness of telephone surveys, since sampling frames for these surveys have traditionally been limited to landline phone numbers. The non-coverage of households without landline telephones may bias the estimates derived from telephone surveys. Recent studies assume that the number of the mobile-only households is growing worldwide. In Europe this share ranges from 3 % in Sweden to 64 % in the Czech Republic (Häder et al. 2009: 23, European Commission 2008: 31). In the United States these households constitute more than 22 % of households overall (Blumberg and Luke 2009).

One of the most important questions regarding mobile phones in telephone surveys concerns the coverage error of the population and the characteristics of the population segments with the highest proportion of non-landline households (Kuusela et al. 2008: 88). The group of "mobile-onlys" can already be described on the basis of several aspects like socio-demographic variables (Mokrzycki et al. 2010), political attitudes and voting behavior (Keeter et al. 2007), media consumption (Fleeman 2006) and health behavior (Blumberg and Luke 2009). At the same time, the results show a high probability of coverage errors. Thus, to ensure the quality of telephone surveys sampling strategies that combine mobile phone and landline phone sample designs (dual frame designs) have become increasingly important. (Brick et al. 2007, Hunsicker and Schroth 2007). Furthermore, current findings indicate that a growing number of persons spend less and less time at home (Tucker and Lepkowski 2008: 11). In addition, a flood of marketing calls has been observed. As a consequence, response rates have been declining for several years (Groves et al. 2002, Singer 2006). In a meta-analysis Kennedy (2010) showed that operations, coverage and sample design are well-studied but few studies have examined measurement errors in telephone surveys using both mobile and landline phones. This chapter provides a brief summary of the theoretical foundations explaining potential differences between landline and mobile phone strategies. Building on a model of survey response processing, several indicators of response quality will be presented and differences between telephone interviews via mobile and landline phones will be examined. We will show that the mode is not the cause but that the potentially associated conditions affect survey response behavior.

15.2 Differences in Response Behavior: Comparing Telephone Interviews via Mobile and Landline Phones

To examine the potential influence of dual frame designs using both mobile and landline phones, we use the findings and theoretical backgrounds of methodological research on mixed-mode data collection (Dillman et al. 2008, de Leeuw 2005). It is known that two surveys covering the same statistical population and use the same questions can lead to differences in response, especially when different modes of data collection are being used (Tourangeau and Smith 1996, Groves et al. 2004, Dillman et al. 2008: 300, Béland and St-Pierre 2008: 298, Tucker and Lepkowski 2008: 6). The mixing of two modes of data collection can reduce data comparability, since different modes provide access to different types of people, attract different types of participants and elicit different responses (mode effects). For this reason, the magnitude and structure of coverage, nonresponse and measurement errors can differ across modes and reduce the comparability of data collected with mixed modes (Jäckle et al. 2008: 1)[1]. The mixing of modes, nevertheless, allows the researcher to compensate the disadvantages of one mode (in our case the non-coverage of mobile-only households via landline) with the advantages of the other (telephone surveys via mobile phone cover the mobile-only households). Furthermore, the combination of different modes of data collection is an effective instrument to reduce both rising implementation costs and increasing nonresponse rates (Dillman et al. 2008, de Leeuw 2005).

[1] In previous research we showed that dual frame designs for national telephone surveys improve the data quality related to nonresponse and coverage error (Häder and Häder 2009). For further information on weighting dual frame designs see Häder et al. (2009). In this study we will focus on measurement errors.

In our research design, we have kept some factors consistent across both modes, but varied others. First, we would like to draw attention to the conditions that remained the same:

- The same interviewers conducted the interviews via landline and mobile phone. This avoided a systematic bias in interviewer behavior.
- Both modes followed the same supervision principles.
- We used almost the same questionnaire (Häder and Häder 2009: 267ff.).
- The same "conversion strategies" (such as call-backs) have been used in order to deal with initial nonresponse.
- The same CATI survey center conducted the survey.
- A dual frame design was used as the sampling procedure. Nevertheless, we were able to ensure an equal probability of selection.
- The field time was almost the same.

The interview as a social interaction is affected by interviewer behavior, the characteristics of the questionnaire, behavior of respondents and the interview situation (Sudman and Bradburn 1974). With the set-up described above, we excluded systematic effects caused by the first two aspects. For that reason, we focus on the behavior of respondents and on the situation in which a respondent finds him- or herself at the time of the interview as a potential source of differences in the response behavior. The hypothesis that the answering behavior between mobile phone and landline surveys will be different arose early. We were especially concerned that people would be more likely to respond inaccurately when interviewed on a mobile phone (Steeh 2004). This assumption is based on several evidence-based considerations; most importantly we assumed that the probability of distraction during a mobile phone interview is higher.

15.2.1 Social Environments

Respondents can be reached in numerous social environments and settings. The range of social environments and settings for mobile phone interviews is clearly greater than for interviews conducted at home due to the freedom of movement. Furthermore, people often multitask while talking on their mobile phone (AAPOR 2010). The same is true for landlines, but the cognitive demand of the competing activities may be greater with mobile phones (Kennedy 2010: 71).

15.2.2 Technical Aspects

Both survey strategies have different technical conditions. Landline interviews should not pose a problem. Mobile phone interviews could be disturbed and interrupted as a result of a lack of network coverage (in the car or train), energy supply and audio quality (AAPOR 2010). While mobile phone audio fidelity has improved, there may still be a difference for some people between the clarity they have on a mobile phone or landline phone device (Kennedy 2010: 72).

15.2.3 Presence of Others

If other people are present at the time of the interview, this may impact on the interview. Respondents may not mind revealing thoughts to interviewers but may still have concerns about disclosures to third parties (Tourangeau et al. 2000: 280). There is empirical evidence of lower data quality in mobile phone interviews when sensitive questions are asked (Lavrakas et al. 2007).

There are several problems with examining the effect of the presence of others. First, the effect of a third party on response behavior is difficult to measure (Smith 1997). Second, it is possible that the effect of a third person's presence influences the question content (Tourangeau et al. 2000: 280). Third, the presence of a third person is also confounded with the degree of privacy.

There are two competing approaches that explain the effect of third parties. One side – the traditional view – emphasizes the negative influences on the data quality, because of response contamination (Corti and Clissold 1992). According to this view, empirical evidence shows that the absence of third parties reduces the risk, that respondents adjust their answers to suit other people. The opposite position focuses the "control function" of other people and the resulting positive effects on data quality (Reuband 1984). It was shown, that familiar third parties prevented response bias towards the supposed expectance of the interviewers (Hartmann 1991). In a situation, where the attending third parties are family or friends who know the opinion and behavior of the respondent, the third person gets a supervisory function. In this situation, which is probably occurs more frequently in the landline phone surveys, the presence of a third person should lead to fewer socially-desirable responses and thus to higher data quality (Van den Bulck 1999).

Research in this field primarily focuses on both the comparability of the described two types of data and the quality of data collected using different modes (Jäckle et al. 2008). Current empirical studies have not found meaningful evidence for the hypothesis that people respond less accurately on mobile phones than on landlines (Steeh 2004, Brick et al. 2007, Lavrakas et al. 2009, Häder and Häder 2009). Little attention, however, has so far been paid to the field of dual frame designs regarding the use of mobile and landline phones. Consequently, there is a need of further investigations in that field (Tucker and Lepkowski 2008: 19f.). This chapter aims to narrow the research gap by providing a theoretical framework and presenting empirical findings with respect to the issues raised above, in particular to the presence of third parties, technical aspects and the social environments.

15.3 Explaining Mode Effects with Psychological Models of Response Processing

Cognitive models of the response process (Sudman et al. 1996, Tourangeau et al. 2000) are helpful in explaining mode effects. In particular, the theories of dual paths to a survey response (Krosnick 1991, Krosnick and Alwin 1987, Strack and

Martin 1987, overview given by Tourangeau et al. 2000: 17) have provided a theoretical framework for explaining numerous phenomena of the response process. Due to the complexity of the response process, it seems questionable whether the respondents actually go through the complete response process, as is described in the cognitive process models (Cannell et al. 1981, Strack 1994: 53, Sudman et al. 1996: 56, overview given by Tourangeau et al. 2000: 5ff.). On the contrary, it is assumed that an interview under unfavorable conditions leads to the use of heuristics in the response generating process (Strack and Martin 1987). Unfavorable conditions means highly cognitively demanding situations and low salience of the relevant object. This indicates that, in case of cognitive strain, the answer is neither based on cognitively demanding judgments nor on the individual's factual knowledge. Instead, the respondent chooses a supposedly simple cognitive solution and responds relatively spontaneously. The typically brief response time[2] supports the assumption that this strategy is fairly common response behavior observed in social scientific surveys. This leads to a decline in the quality of data, since both involve increased nonresponse as well as increased randomness in choosing response category.

Krosnick and Alwin (1987) developed a model of parallel processing paths that can be applied to all phases of the response process. The model assumes that when the cognitive tasks are difficult (difficult question topic, high recall performance, etc.) and the abilities of respondents and their motivation are relatively low (Krosnick 1991), variations from the "optimal" response process are expected. If one or more of these conditions are met, the likelihood of the cognitive processes carried out merely on a superficial level increases (weak satisficing). Some steps may even be skipped completely (strong satisficing). The quality of the responses in a survey is therefore always subject to various influences. One aspect should be emphasized. Based on the assumptions of the cognitive psychological research, one can assume that respondents who go through all four phases of the response process achieve the highest level of elaboration and this leads to higher data quality.

15.3.1 Aims and Hypotheses

From this we can derive our hypotheses:

1. Hypothesis on place of interview: If the interviews are conducted outside the home we assume a higher level of environmental influences and distraction. Therefore, interviews outside home also lead to the use of satisficing strategies and to more random and instable answers.
2. Hypothesis on presence of others: The higher probability of the presence of others during the mobile phone interviews leads to satisficing strategies. This in turn leads to lower data quality.
3. Hypothesis on the degree of privacy: At the same time it is less probable that familiar third parties, like friends and family, are present during mobile phone interviews. This should lead to more socially-desirable responses.

[2] A typical attitude question is on average answered in less than five seconds (Tourangeau et al. 2000: 184).

4. Hypothesis on mode-effects: The impact of the mode (mobile versus landline) gets weaker or may even be lost if the place of the interview and the presence of others are controlled.

15.3.2 Data and Methods

In order to make sure that all telephone numbers could be contacted up to 15 times and in different weeks, the field period lasted from October 2007 to April 2008. In this way sufficient data could be collected to decide on whether the telephone number was working (or not). In total, 2,171 interviews were conducted by the telephone lab of the Technical University of Dresden, 1,124 via mobile phones and 1,047 via landline phones. Table 15.1 provides an overview of the exact disposition codes.

Table 15.1 Disposition codes for the cell phone and the landline survey

	mobile phone	landline
Not eligible		
Fax line	63	667
Non-working number: computer voice	13,585	7,947
Non-working number: non-contact	1,710	2,405
Non-residence	156	704
Other	324	145
Unknown eligibility, non-interview		
Voice mail – don't know if household	5,150	1,135
Eligible, non interview		
Refusal	1,696	2,327
Break off	53	17
Physically or mentally unable	17	14
Other	83	377
Interviews		
Complete	1,124	1,047
Partial	84	67
Total phone numbers used	24,045	16,855

A dual frame approach was used to combine landline and mobile phone samples. Two parallel surveys were carried out: one via landline and one via mobile phone. The "Gabler-Häder-Design" which is a kind of list-assisted random digit dialing (RDD) approach (Gabler and Häder 2002) was used for the selection of landline phone numbers. This design is well established in Germany. The sampling for the mobile phone survey was challenging, as landline samples and mobile phone samples overlap for households with both landline and mobile phones. This issue has already been discussed in detail in the AAPOR mobile phone task force. Two different approaches have been used to handle the overlap: the screening and the overlap approaches.

- The screening approach: Interviewers identify mobile phone respondents that can also be reached via a landline phone at home. These individuals are considered eligible only when sampled from the landline frame.

- The overlap approach: Information about each respondent's phone status is collected during the survey, independent of the sampling frame. The interview is conducted regardless of the frame from which the respondent is selected (landline or cell).

A general consensus on the preferred method has, however, not yet been achieved. Because the "mobile phone only" population makes up only about 8 % of the total population, the second approach was chosen. In addition, the first approach was considered to time- and money-consuming for our purposes. The theoretical foundations for the dual frame approach and the computation of the weights are provided by Callegaro et al. (2010).

15.3.3 Independent Variables

Modus of phone survey

This variable provides information on whether the interview was conducted via mobile phone or landline phone.

Place of Interview

Interviews conducted in the household of the respondents and those conducted outside have to be distinguished. We reached numerous participants on their mobile phones at home and some via landline phones outside the household (with cordless handsets). Because of that the mode of the interview does not equal the place of the interview.

Presence of others

Whether at home or outside, there was a chance of other people being present during the time of interview. This was checked with a direct question. We asked whether others were present during the interview. It was further asked whether these third parties were a) friends, family, acquaintances or b) strangers.

Control variables

For the multivariate analysis three additional variables - gender, age and educational level of the respondents –were included. We expected effects related to the cognitive abilities of respondents. In addition, most mode comparison studies lack validation data and thus cannot separate the effects of differential nonresponse bias from the effects of differences in the measurement errors (Kreuter et al. 2008). Another reason for these three variables is the growing number of "mobile phone onlys" and the known socio-demographic specifics of this group. They are relatively young, male and have a professional training (Graeske and Kunz 2009: 60f.).

15.3.4 Indicators of Data Quality

There are numerous indicators to determine the extent of measurement errors. We have focused on the degree of elaboration in CELLA. Because of that the answering process and the resulting consistency and coherence of the response behavior were examined. The individual indicators of data quality are subsequently briefly introduced.

Recall performance

Differences in memory performance exhibited by the respondents may indicate differences in response quality as well (Noelle-Neumann and Petersen 2000: 183ff.).

The process of reasoning, which precedes the response, is being terminated in the moment the interviewed person thinks he or she has recalled enough knowledge or if heuristics are used to formulate the answer. The latter causes a shortening of the response process. A respondent using heuristics does not need complete access to the relevant knowledge in order to answer the questions, but only a subset. The better the memory performance, the higher the response quality should be. Unfavorable conditions during the interview may lead to poorer memory performance and thus cause an early termination of the information search.

To compare the memory performance of individuals in both modes, the following procedure was used. First, the interviewer asked the respondents which things they consider necessary or dispensable in life. The respondents had to answer the question to what extent they considered for example a two-week holiday or the purchase of branded products dispensable or indispensable. A total of nine questions of this nature were answered. To avoid additional response order effects the order of these nine questions was randomized. The subsequent open question was: "I have just raised different issues. Which do you remember?” It is interesting how many of the nine issues the respondents could still remember correctly and of course which answers were given to the nine questions.

Item nonresponse

Causes for item nonresponse at individual questions are assumed in all four stages of cognitive psychological response model. In particular strong distraction and disruption should lead to a higher probability of satisficing, which in turn may result in item nonresponse (Krosnick,2000). Substantive responses are given in the absence of distractions and disruptions, due to a higher cognitive effort. This leads to less satisficing (Shoemaker et al. 2002: 196).

Inconsistencies in response behavior

If heuristics are used in order to generate answers, inconsistencies in the response behavior are likely to emerge. There are several options to measure response inconsistence. In this study, inconsistencies in response behavior have been measured in terms of response stability.

For this purpose, the same attitude objects have been measured at different times. So far, this procedure has especially been used in trend and panel designs (Converse 1964, Kane et al. 2000, Schuman and Presser 1981: 259ff.). The stability of responses has been little explored in cross-sectional surveys. Slaby (1998) examined the issue of short-term stability of preferences in a telephone survey. He showed that systematically composed attitudes have a higher resistance and persistence compared to heuristically constructed answers (1998: 13) and confirmed findings of longitudinal studies. Kühne and Böhme (2006) also examined the response stability in attitude questions in a cross-sectional design. They found that the retrieval of attitude relevant knowledge increased the stability of responses. To examine changes in attitudes in this cross-sectional survey, an attitude object was measured twice. This approach is not entirely unproblematic but in our view represents a viable way especially since there is already empirical evidence supporting this assumption (Slaby 1998: 2, Kühne and Böhme 2006: 16). In pretests different question versions were tested. We measured the same attitude object twice in the same survey, while making an effort to avoid respondents spotting this duplication[3].

Thus, we created a repeated measurement and due to the short time interval, a high level of consistence could be expected. Due to the slightly different question wording, a constant bias between the two measurement times was to be expected, which represents the semantic difference between the question wordings. The degree to which the responses at both time points differ is an indicator of response stability.

Social desirability

Social desirability is a very important effect in survey research and can be the reason for many phenomena like response sets. To measure social desirability, an SD scale was used and modified according to numerous pretests (originally Edwards 1957, Fordyce 1956). The instrument was taken from the Socio-Economic Panel (SOEP) (Winkler et al. 2006); which was originally developed by Paulhus (1984). Respondents were given six items and asked whether they agreed or disagreed on a answer scale with five categories.

15.4 Results

15.4.1 Descriptive Results and Bivariate Analysis

Altogether we interviewed 1,124 on their mobile phone and 1,047 via landline phone. The vast majority of the landline phone interviews were conducted at home (odds ratio = 42.1/1.7 = 24.8). As expected, the proportion for the mobile phone interviews is different (see Figure 15.1). In contrast to the landline interviews the odds ratio is clearly lower (odds ratio = 35.4/20.8 = 1.7).

[3] The respondents were first asked to assess the impact of the expansion of mobile networks on a scale from negative (1) to positive (5). Later, respondents were asked to assess whether all the advantages outweigh the disadvantages of the expansion of mobile networks from 1 to 5.

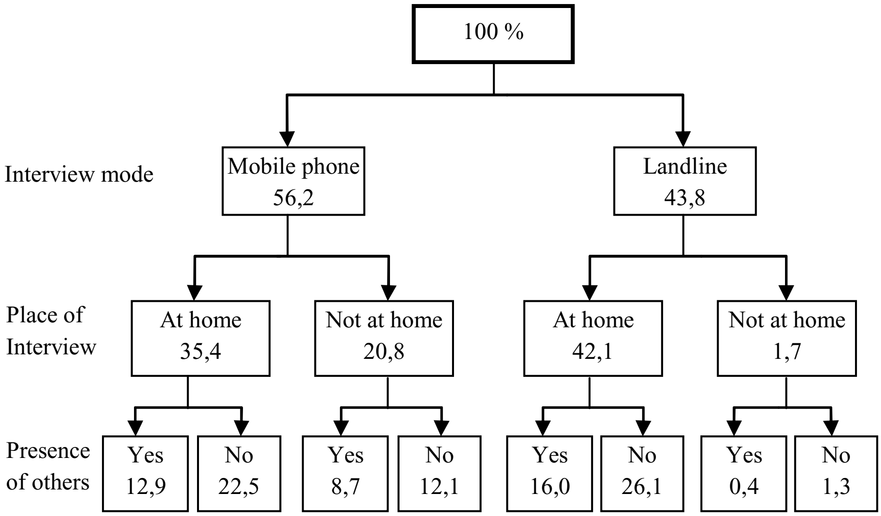


Fig. 15.1 Tree diagram of interview mode, place of the interview and presence of others

Taking into account the presence of others, it can be seen that in all conditions third persons were rarely present. But the highest probability for the presence of others was for mobile phone interviews especially if the interview was not conducted at home (odds ratio = 8.7/12.1 = 0.71). Only in 68 interviews the person present was classified as a stranger. 64 of these interviews were conducted on a mobile phone and most of these (n = 57) outside the home.

The point biserial correlation between the mode (0 = landline, 1 = mobile phone) and memory performance (0 = no recall, 9 = all things recalled) is very low, but points in the opposite direction (r_{pb} = 0.1, $p < 0.05$). Respondents interviewed via mobile phone remembered better. The impact of background variables seems likely. The correlation between the mode and the second indicator of data quality, item nonresponse, shows no significant result. The relationship between mode and response stability (0 = stable, 4 = maximum difference), supports the assumption that persons who are interviewed on a mobile phone answer even less stably (= 0.1, $p < 0.05$). The last indicator shows a significant result too. The point biserial correlation between the mode and social desirability is low but negative (r_{pb} = -0.1, $p < 0.05$). Respondents interviewed via mobile phone showed a less socially desirable response behavior. This result is counter-intuitive, too. To control the influence of other variables, multivariate analyses are needed.

15.4.2 Multivariate Analyses

Some effects can be expected in relation to the cognitive abilities and the response behavior of respondents (Schuman and Presser 1981). Because of that, additional control variables, gender, age and educational level of the respondents, have been included. Because of potential nonlinear relationships between the independent variables, the interaction terms were also included in the models.

Recall performance

The postulated effect of the forth hypothesis has been observed (see Table 15.2). If the place of the interview and the presence of third parties are controlled for, the influence of the bivariate mode effect diminishes. Both the place of the interview and the presence of others have no effect. Not unexpectedly, the empirical evidence also shows that, when it comes to cognitive tasks background variables matter. The counterintuitive results of the bivariate analysis are explained by the socio-demographic differences of the persons interviewed via mobile and landline phones. Respondents with higher education, young people and women recalled information better than other groups.

Table 15.2 Linear regression model predicting the recall performance, unstandardized and [standardized] coefficients

Mode (0 = landline)	- 0.06
At home (0 = yes)	- 0.19
Presence of others (0 = no)	0.03
Years in education	0.09 [0.18]***
Sex of respondent (0 = male)	0.43***
Age of respondent	- 0.03 [-0.24]***
Mode x presence of others	0.04
Mode x at home	0.03
At home x presence of others	- 0.10
Intercept	3.48***
n = 1,924	
R^2 = 0.09	

Note. * = 0.05, ** = 0.01, *** = 0.001; Dependent variable: Recall performance (0 = no recall to 9 = all things recalled), weighted data.

Item nonresponse

Contrary to expectations, the mode influences the item response when controlling for other variables. Interviews via mobile phone lead to higher item nonresponse even when controlling for the other independent variables and interactions. Specific technical aspects may lead to this effect. These are, however, hardly measurable. The significant results of the control variables are comprehensible. Don't know responses are given most frequently by the less educated (Schuman and Presser 1981: 160). With declining cognitive abilities, older respondents are less well able to go through the response process fully and optimally. This leads to higher item nonresponse rates (Slymen et al. 1994, Rodgers and Herzog 1987). Both effects are significant in the model (see Table 15.3).

Table 15.3 Linear regression model predicting item nonresponse, unstandardized and [standardized] coefficients

Mode (0 = landline)	0.17**
At home (0 = yes)	0.17
Presence of others (0 = no)	0.09
Years in education	- 0.01 [-0.05]*
Sex of respondent (0 = male)	- 0.06
Age of respondent	0.01 [0.17]***
Mode x presence of others	- 0.06
Mode x at home	- 0.10
At home x presence of others	- 0.03
Intercept	0.18
n = 1,924	
$R^2 = 0.03$	

Note. * = 0.05, ** = 0.01, *** = 0.001; Dependent variable: Item nonresponse (0 = no item nonresponse to 14 = maximum item nonresponse), weighted data.

Response stability

The hypothesis on mode-effects mentioned above can be confirmed again (see Table 15.4). Neither the independent variables (mode, place of interview and presence of others) nor the interaction terms are significant. There is only one significant influence of a control variable. Less educated respondents answered less stably. Without being able to fully clarify the findings, it is noteworthy that the correlation of mode and response stability in the multivariate model is lost. It seems that particularly the level of education affects the stability of attitudes.

Table 15.4 Linear regression model predicting the strength of response stability, unstandardized and [standardized] coefficients

Mode (0 = landline)	0.09
At home (0 = yes)	0.02
Presence of others (0 = no)	0.08
Years in education	- 0.02 [0.08]**
Sex of respondent (0 = male)	- 0.06
Age of respondent	0.01 [0.01]
Mode x presence of others	- 0.11
Mode x at home	0.04
At home x presence of others	0.04
Intercept	0.94***
n = 1,404	
$R^2 = 0.01$	

Note. * = 0.05, ** = 0.01, *** = 0.001; Dependent variable: Strength of response stability (0 = stability to 4 = instability), weighted data.

Social desirability

In this model, the correlation between mode and social desirability is lost. The mode effect exists no more. At the same time, the model shows (see Table 15.5) that the presence of third parties has an effect. As expected, due to a social desirability bias, respondents gave higher ratings, when a third person was present during the interview. At the same time the model shows no interaction effect. The presence of third parties seems to have an independent influence (regardless of both modes and the interview location) on the response behavior and in that special case on socially desirable responses. The interview location once again has no effect.

In contrast, the control variables have an influence. A higher educational level leads to less social desirability. Older persons and female respondents in turn more often showed socially desirable response behavior.

Table 15.5 Linear regression model predicting the degree of social desirability, unstandardized and [standardized] coefficients

Mode (0 = landline)	0.01
At home (0 = yes)	- 0.08
Presence of others (0 = no)	0.20*
Years in education	- 0.02 [-0.06]*
Sex of respondent (0 = male)	0.33***
Age of respondent	0.02 [0.22]***
Mode x presence of others	0.02
Mode x at home	0.13
At home x presence of others	- 0.04
Intercept	2.94***
n = 1,920	
R^2 = 0.07	

Note. * = 0.05, ** = 0.01, *** = 0.001; Dependent variable: Social desirability (0 = minimum to 6 = maximum), weighted data.

Privacy

The final model only includes cases in which third persons were present. The variable “presence of others” was adapted to this change and is called now “privacy”. The variable is dichotomous, distinguishing whether the third parties during the interview were friends (including family and acquaintances) or strangers. The mode is again without effect. It is interesting that the interaction effect between mode and privacy is significant. This interaction effect supports the third hypothesis. The presence of strangers during a mobile phone interview leads to more socially desirable responses. The effect of control variables remains, as in the previous model. Respondents with higher education, young people and men showed less socially desirable response behavior.

Table 15.6 Linear regression model predicting the degree of social desirability, unstandardized and [standardized] coefficients

Mode (0 = landline)	- 0.11
At Home (0 = yes)	- 0.68
Privacy (0 = friends, family, acquaintances)	- 1.93*
Years in education	- 0.05 [-0.11]**
Sex of respondent (0 = male)	0.40***
Age of respondent	0.02 [0.22]***
Mode x privacy	1.41*
Mode x at home	0.83
At home x privacy	0.41
Intercept	3.35***
n = 769	
$R^2 = 0.08$	

Note. * = 0.05, ** = 0.01, *** = 0.001; Dependent variable: Social desirability (0 = minimum to 6 = maximum), weighted data.

15.5 Conclusion and Discussion

This chapter aimed to explain the influence of mode effects in the proposed mixed mode design on the basis of the theoretical assumptions of the cognitive psychological phase model and its extensions. It was furthermore intended to quantify those effects empirically.

An important assumption of this study was confirmed at the beginning. It was shown that in mobile phone interview situations third parties are present more often than expected and that these persons are often strangers.

We present three main findings, corresponding to the hypotheses we posed initially. First, it was partially confirmed that the mode is not the cause, but instead, the potentially associated conditions affect the survey response behavior. In the further course of our analysis, we examined to what extent these relationships endure if aspects of the interview situation (location, presence of third parties and the interaction terms) and the control variables (age, gender and educational level) explain this effect. For one indicator (item nonresponse) the mode has an effect when all the other variables are controlled for. The assumption of specific technical aspects as a cause seems reasonable, though. The place of the interview had no influence. Second, the presence of others leads to more socially-desirable responses. Especially for mobile phone interviews the presence of strangers raises the probability of social desirability. Third, particularly the socio-demographic variables have an influence on the data quality in the regression models. Especially education has to be stressed. In all models a higher educational level led to higher data quality. The results underline the importance of the socio-demographic variables for explaining differences in data quality and thus differences in persons reached via the two modes.

A possible weakness of the study may be the influence of the presence of a third party. It can be completely neutral on the response quality, if this third person has not been in any interaction with the respondent. In our study we only asked whether another person was present during the interview, but not whether this had any influence on the interview. Furthermore, the question of how strong the effects differ has not fully been elucidated at this point either.

Altogether, the overall influence of the aspects examined is low when considered in detail. These results correspond with results of similar studies (Brick et al. 2006, Steeh 2004). This can be concluded considering, among other things, the very small values for the model quality in the regression models (all coefficient of determination are lower than 0.1). They point to other variables as determinants of the response behavior. Generally, this is very good news for future mobile phone surveys.

References

AAPOR (The American Association for Public Opinion Research) Guidelines and Considerations for Survey Researchers When Planning and Conducting RDD and Other Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers (2008), `http://www.aapor.org/uploads/Final_AAPOR_Cell_Phone_TF_report_041208.pdf` (accessed July 31, 2010)

AAPOR (American Association for Public Opinion Research). New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in the U.S. With Respondents Reached via Cell Phone Numbers (2010), `http://www.aapor.org/Cell_Phone_Task_Force/2819.htm` (accessed January 10, 2011)

Béland, Y., St-Pierre, M.: Mode Effects in the Canadian Community Health Survey: A Comparision of CATI and CAPI. In: Lepkowski, J.M., et al. (eds.) Advances in Telephone Survey Methodology, pp. 297–316. Wiley, New York (2008)

Blumberg, S.J., Luke, J.V.: Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July-December 2008. Division of Health Interview Statistics, National Center for Health Statistics (May 2009), `http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless200905.htm` (accessed July 31, 2010)

Boyle, J.M., Lewis, F.: Cell Phone Mainly Households: Coverage and Reach for Telephone Surveys Using RDD Landline Samples. Survey Practice (2009), `http://www.surveypractice.org` (accessed July 31, 2010)

Brick, J.M., Brick, P.D., Dipko, S., Presser, S., Tucker, C., Yuan, Y.: Cell Phone Survey Feasibility in the U. S.: Sampling and Calling Cell Numbers Versus Landline Numbers. Public Opinion Quarterly 71, 23–39 (2007)

Brick, J.M., Dipko, S., Presser, S., Tucker, C., Yuan, Y.: Nonresponse Bias in a Dual Frame Sample of Cell and Landline Numbers. Public Opinion Quarterly 70, 780–793 (2006)

Callegaro, M., Ayhan, Ö., Häder, S., Gabler, S.: Combining landline and mobile phone samples: a dual frame approach (2010) (submitted)

Cannell, C.F., Miller, P.V., Oksenberg, L.: Research on Interviewing Techniques. In: Leinhardt, S. (ed.) Sociological Methodology, pp. 389–437. Jossey-Bass, San Francisco (1981)

Converse, P.E.: The Nature of Belief Systems in Mass Publics. In: Apter, D.E. (ed.) Ideology and Discontent, pp. 206–261. Free Press, New York (1964)

Corti, L., Clissold, K.M.: Response contamination by third parties in a household interview survey. In: Working Papers of the ESRC Research Centre on Micro-Social Change, Working Paper 13. University of Essex, Colchester (1992)

De Leeuw, E.D.: To mix or not to mix data collection modes in surveys. Journal of Official Statistics 21(5), 233–255 (2005)

Dillman, D.A., Smyth, J., Christian, L.M.: Mail and Internet Surveys. In: The Tailored Design Method. Wiley, New Jersey (2008)

Driscoll, D.L.: Paraphrase: Write it in Your Own Words. Purdue University Online Writing Lab (2007), http://owl.english.purdue.edu/owl/resource/619/01/ (accessed July 31, 2010)

Edwards, A.L.: The Social Desirability Variable in Personality Assessment and Research. Dryden, New York (1957)

European Commission. E-Communications Household Survey. Eurobarometer Spezial 293/Welle 68.2 (2008), http://ec.europa.eu/public_opinion/archives/ebs/ebs_293_full_de.pdf (accessed July 31, 2010)

Fleeman, A.: Merging Cellular and Landline RDD Sample Frames: A Series of Three Cell Phone Studies. In: Paper presented at the Second International Conference on Telephone Survey Methodology, Miami, FL (2006)

Fordyce, W.E.: Social Desirability in the MMPI. Journal of Consulting Psychology 3(20), 171–175 (1956)

Gabler, S., Häder, S.: Die Kombination von Mobilfunk- und Festnetzstichproben in Deutschland (The combination of mobile and landline samples). In: Weichold, M., et al. (eds.) Umfrageforschung. Herausforderung und Grenzen. Österreichische Zeitschrift für Soziologie, Special, vol. 9, pp. 239–252. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Gabler, S., Häder, S.: Idiosyncrasies in telephone sampling. The case of Germany. International Journal of Public Opinion Research 14, 338–345 (2002)

Graeske, J., Kunz, T.: Stichprobenqualität der CELLA-Studie unter besonderer Berücksichtigung der Mobile-onlys (Sampling quality of the CELLA study with special emphasis on Mobile onlys). In: Häder, M., Häder, S. (eds.) Telefonbefragungen über das Mobilfunknetz, pp. 57–70. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Groves, R.M., Dillman, D.A., Eltinge, J.L., Little, R.J.A.: Survey nonresponse. Wiley, New York (2002)

Groves, R.M., Floyd Jr., J.F., Couper, M., Lepkowski, J.M., Singer, E., Tourangeau, R.: Survey Methodology. Wiley, New Jersey (2004)

Häder, M., Häder, S.: Telefonbefragungen über das Mobilfunknetz (telephon surveys via mobile phone). VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Häder, S., Gabler, S., Heckel, C.: Stichprobenziehung. In: Häder, M., Häder, S. (eds.) Telefonbefragungen über das Mobilfunknetz, Konzept, Design und Umsetzung Einer Strategie zur Datenerhebung, pp. 21–45. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Hartmann, P.: Wunsch und Wirklichkeit: Theorie und Empirie sozialer Erwünschtheit (Hope and Reality: Theory and Practice of social desirability). Deutscher Universitäts-Verlag, Wiesbaden (1991)

Hofstätter, P.R.: Bedingungen der Zufriedenheit (Conditions of satisfaction). Edition Interfrom, Zürich (1986)

Hunsicker, S., Schroth, Y.: Combining mobile phone and landline phone samples. A practical application of the dual frame approach. MDA – Methoden, Daten, Analysen. Zeitschrift für Empirische Sozialforschung 2(1), 161–182 (2007)

Hyman, H.H., Sheatsley, P.B.: The Current Status of American Public Opinion. In: Payne, J.C. (ed.) The Teaching of Contemporary Affairs. Twenty-first Yearbook of the National Council of Social Studies, pp. 11–34. National Education Association, New York (1950)

Jäckle, A., Roberts, C., Lynn, P.: Assessing the Effect of Data Collection Mode on Measurement. ISER Working Papers 2008-08 (2008), `http://www.iser.essex.ac.uk/publications/working-papers/iser/2008-08` (accessed July 31, 2010)

Kane, J.G., Craig, S.C., Martinez, M.D.: Ambivalence, Attitude Strength, and Response Instability: A Two-Wave Panel Study of Abortion Attitudes in Florida. Paper presented at the 2000 Annual Meeting of the American Political Science Association, Washington (2000)

Keeter, S., Kennedy, C.K., Clark, A., Tompson, T., Mokrzycki, M.: What's Missing from National Landline RDD Surveys? The Impact of the Growing Cell-only Population. Public Opinion Quarterly 71, 772–792 (2007)

Kennedy, C.K.: Nonresponse and Measurement Error in Mobile Phone Surveys. PhD Thesis, University of Michigan (2010)

Kreuter, F., Presser, S., Tourangeau, R.: Social desirability bias in cati, IVR, and web surveys. Public Opinion Quarterly 5(72), 847–865 (2008)

Krosnick, J.A., Alwin, D.F.: An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. Public Opinion Quarterly 2(51), 201–219 (1987)

Krosnick, J.A.: Response strategies for coping with the cognitive demands of attitude measures in surveys. Journal of Applied Cognitive Psychology 5, 213–236 (1999)

Krosnick, J.A.: The threat of satisficing in surveys: the shortcuts respondents take in answering questions. Survey Methods Newsletter 20, 4–8 (2009)

Kühne, M., Böhme, R.: Effekte von Informationsstand, Wissen und Einstellungsstärke von Befragten auf die Antwortstabilität in Online-Befragungen mit Selbstrekrutierung (Effects of level of information, knowledge and attitudes of respondents on the strength of response stability in online surveys with self-recruitment). ZUMA-Nachrichten 59, 42–71 (2006)

Kuusela, V., Callegaro, M., Vehovar, V.: The Influence of Mobile Telephones on Telephone Surveys. In: Lepkowski, J.M., et al. (eds.) Advances in Telephone Survey Methodology, pp. 87–112. Wiley, New York (2008)

Lavrakas, P.J., Steeh, C., Shuttles, C., Fienberg, H.: The State of Surveying Cell Phone Numbers in the United States: 2007 and Beyond. Public Opinion Quarterly 5(71), 840–854 (2007)

Lavrakas, P.J., Tompson, T.N., Benford, R.: Investigating Data Quality in Cell Phone Surveying. In: Paper Presented at the Annual Conference of the Midwest Association for Public Opinion Research, Chicago (2009)

Lepkowski, J.M., Tucker, C., Brick, J.M., Leeuw, E., de Japec, L., Lavrakas, P.J., Link, M.W., Sangster, R.L.: Advances in Telephone Survey Methodology. Wiley, New York (2008)

Mason, R., Carlson, J.E.: Contrast Effects and Subtraction in Part-Whole Questions. Public Opinion Quarterly 4(58), 569–578 (1994)

Mokrzycki, M., Keeter, S., Kennedy, C.: Cell-Phone-Only Voters in the 2008 Exit Poll and Implications for Future Non-coverage Bias. Public Opinion Quarterly 5(73), 845–865 (2010)

Noelle-Neumann, E., Petersen, T.: Das halbe Instrument, die halbe Reaktion. Zum Vergleich von Telefon- und Face-to-Face Umfragen (Half the instrument, the half reaction. For comparison of telephone and face-to-face surveys). In: Hüfken, V. (ed.) Methoden in Telefonumfragen, pp. 183–200. Westdeutscher Verlag, Opladen (2000)

Paulhus, D.L.: Two-Component Models of Social Desirability. Journal of Personality and Social Psychology 3(46), 598–609 (1984)

Reuband, K.-H.: Dritte Personen beim Interview - Zuhörer, Adressaten oder Katalysatoren der Kommunikation? (Third person being interviewed - listener, addressee or catalysts of communication?). In: Meulemann, H., Reuband, K.-H. (eds.) Soziale Realität im Interview, pp. 117–156. Campus, Frankfurt (1984)

Rodgers, L.W., Herzog, A.R.: Interviewing older adults. The accuracy of factual information. Journal of Gerontology 42, 389–394 (1987)

Schuman, H., Presser, S.: Questions and Answers in Attitude Survey. Sage Publications, Inc. (1981)

Shoemaker, P.J., Eichholz, M., Skewes, E.A.: Item Nonresponse: Distinguishing between don't Know and Refuse. International Journal of Public Opinion Research 14, 193–201 (2002)

Singer, E.: Introduction - Nonresponse bias in household surveys. Public Opinion Quarterly 70(5), 637–645 (2006)

Slaby, M.: Zur Interaktion zwischen Befragten und Erhebungsinstrument (For interaction between respondents and survey instrument). SISS – Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart, No. 3/1998 (1998), http://elib.uni-stuttgart.de/opus/volltexte/1999/415/pdf/415_1.pdf (accessed July 31, 2010)

Slymen, D.J., Drew, J.A., Wright, B.L., Elder, J.P., Williams, S.J.: Item non-response to lifestyle assessment in an elderly kohort. International Journal of Epidemiology 23, 583–591 (1994)

Smith, T.W.: The Impact of the Presence of Others on a Respondent's Answers to Questions. International Journal of Public Opinion Research 1(9), 33–47 (1997)

Steeh, C.G., Piekarski, L.: Accommodating New Technologies: Mobile and VOIP Communication. In: Lepkowski, J.M. (ed.) Advances in Telephone Survey Methodology, pp. 423–448. Wiley, New York (2008)

Strack, F., Martin, L.L.: Thinking, judging, and communicating: A process account of context effects in attitude surveys. In: Hippler, H.-J., et al. (eds.) Social Information Processing and Survey Methodology, pp. 123–148. Springer, Heidelberg (1987)

Strack, F.: Zur Psychologie der standardisierten Befragung: kognitive und kommunikative Prozesse (On the psychology of the standardized interview: cognitive and communicative processes). Springer, Berlin (1994)

Steeh, C: A New Era for Telephone Surveys. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ (2004)

Sudman, S., Bradburn, N.M.: Response Effects in Surveys: A review and synthesis. Aldine publishing company, Chicago (1974)

Sudman, S., Bradburn, N.M., Schwarz, N.: Thinking about Answers. In: The Application of Cognitive Processes to Survey Methodology. Jossey-Bass, San Francisco (1996)

Tourangeau, R., Smith, T.W.: Asking Sensitive Questions: The Impact of Data Collection, Question Format, and Question Context. Public Opinion Quarterly 2(60), 275–304 (1996)

Tourangeau, R., Rips, L.J., Rasinski, K.A.: The Psychology of Survey Response. Cambridge University Press, New York (2000)

Tucker, C., Lepkowski, J.M.: Telephone Survey Methods: Adapting to Change. In: Lepkowski, J., et al. (eds.) Advances in Telephone Survey Methodology, pp. 3–28. Wiley, New York (2008)

Van den Bulck, J.: Does the Presence of a Third Person Affect Estimates of TV Viewing and Other Media Use? Communications 1(24), 105–116 (1999)

Willits, F.K., Saltiel, J.: Question Order Effects on Subjective Measures of Quality of Life. Rural Sociology 4(60), 654–665 (1995)

Winkler, N., Kroh, M., Spiess, M.: Entwicklung einer deutschen Kurzskala zur zweidimensionalen Messung von sozialer Erwünschtheit. Discussion Papers 579. DIW Berlin, Berlin (2006), http://opus.zbw-kiel.de/volltexte/2007/6253/pdf/dp579.pdf (accessed July 31, 2010)

Chapter 16
Data Quality in Telephone Surveys via Mobile and Landline Phone

Michael Häder

16.1 Introduction

In social research, where data are collected by means of surveys, a high quality of the responses and thus of the survey data is pivotal. A vast number of research projects has looked into ways of guaranteeing high data quality and several influences and effects have been identified. Some of the most prominent influences include effects of the question order, question wording, social desirability, sponsorship, effects resulting from the answer categories provided, as well as interviewer effects. These impact on the answers given, such that the true value might not be recorded. This is true for many survey modes, including telephone surveys, which have been frequently subject to methodological research. Especially with telephone surveys researchers tend to be concerned about the quality of the answers provided. Nevertheless, they are currently the most common mode of data collection, not only in Germany but all over Europe. Research so far has delivered important findings that have had a positive impact on the quality of telephone interviews. However, due to the changing nature of telephone surveys new methodological research is needed.

Various new developments the area of telecommunications are relevant for social surveys. For example, the proportion of persons that can be reached via a landline phone – in addition to their mobile phone – is plummeting. This coverage problem necessitates interviews to also be conducted via mobile phones which bring about new methodological concerns. Methodological challenges such as mode effects, the routine of using mobile phones, the situation in which the respondent receives the survey phone call, activities that the respondent needs to suspend, and third parties present during the interview may affect the respondent's behavior and answers.

This sets the scene for the content of this chapter. It continues with setting out the theory drawn on to explain response quality (see section 16.2). Subsequently, we discuss mode effects that might occur in telephone surveys compared to other modes of implementation (see section 16.3). Next, we consider those methodological aspects that arise with the parallel use of landline and mobile phone surveys (see sections 16.4). Finally, our own research findings are presented (see section 16.5).

Psychological personality traits are applied to explain response quality and unit nonresponse. This aspect was first explored within the context of a survey methodological study focusing on combining landline and mobile phone interviewing (CELLA 2).

16.2 Applying Cognitive Theories to Response Quality in Interviews

The foundations for explaining survey responses were laid by a cognitive psychological model (see for example Sudman et al. 1996, Tourangeau et al. 2000) based on the following four components: (1) understanding the survey question, (2) retrieving information from memory, (3) judgment, and (4) editing and formulating a response.

In addition, there are various modifications of this standard model (Strack 1994: 53, Sudman et al. 1996: 56, Tourangeau et al. 2000: 8, Jobe and Herrmann 1996). One set of modifications concerns the provision of a pre-formulated response, where the information retrieval and judgment phases are omitted. This presupposes that the judgment requested in the survey already exists in the respondent's mind. Second, a lack of motivation may lead to a short-cutting of the entire cognitive process, which in turn may lead to a lower response quality. The respondent may for example use heuristics to obtain an acceptable answer (Krosnick and Alwin 1987, Strack and Martin 1987). This is especially important when the survey topic is of little interest to the respondent. Disinterest leads to a lower motivation for finding an accurate answer and to cognitively less demanding and more spontaneous responses. Another situation where this short-cutting might occur is when questions are asked that are particularly difficult, in terms of the comprehensibility of the concepts and in terms of the complexity of cognitive information search. This often causes the use of heuristics and this negatively influences data quality.

In numerous studies researchers have detected various effects caused by these short-cutting strategies leading to biases in the recorded value. Examples of these are answering scale effects (Stadtler 1983, Schwarz et al. 1985, Bishop 1987, Böcker 1988) and context effects (Cantril 1944, cited in Strack 1994, Schuman and Presser 1981: 47, Schuman 1992, Lynn and Kaminska in this book for an overview). The stronger these effects are the lower is the quality of the response. In addition, the probability of order effects increases in such situations (Payne 1949, Hyman and Sheatsley 1950, Dillman et al. 1995).

16.3 Data Quality in Telephone Surveys and Other Survey Modes

The data quality achieved in a specific mode of data collection depends on the efforts made by the respondent to cognitively produce their responses (see section 16.2). In addition, nonresponse processes and under-coverage of the sampling frame play a role (Groves 1989: 502ff.). A specific interview mode might be especially unattractive to certain groups of the population producing much higher levels of nonresponse in these groups than in others. The collected data will thus be biased.

In this paper, we focus on the response quality in telephone interviews applying the cognitive model introduced above. Some authors have voiced general concerns about the potential of telephone interviews (Noelle-Neumann and Petersen 2000), others do not find such clear differences between modes (Brick and Lepkowski 2008). In the following, we pinpoint the most important survey methodological characteristics of telephone surveys.

The role of the interviewers differs in telephone surveys from their role in face-to-face surveys (de Leeuw 1992, 2005, Aquilino 1994, Brick and Lepkowski 2008: 155, Krausch 2007). Furthermore, interviewers are supervised differently across the two survey modes (Reuband 1990, Schnell 1991).

The sampling design of face-to-face surveys is frequently clustered, for example due to the random walk process within a sample point. The efforts undertaken to convince the selected person to participate in the interview differs across survey modes (Goyder 1987: 163ff., Diekmann and Jann 2001: 25, Church 1993, Bosnjak 2001: 90). The length of the fieldwork period also tends to somewhat differ across modes.

In telephone surveys the entire communication between the interviewer and the respondents is solely conducted via a single aural channel. This requires the development of different questionnaires for different modes of interview (Christian et al. 2008: 250ff., Wüst 1998).

Distractions by the respondent's environment during the interview might be especially prevalent for interviews conducted via mobile telephones. The reason is that the researcher has no influence on and prior knowledge of where the survey call will reach the respondent (Redline and Dillman 2002). And generally, the presence of third parties during the interview may influence the respondent's behavior.

Computer assisted interviews (CAI) limit errors when recording responses, facilitate routings, and enable logical checks during the interview. However, where interviewers lack experience in operating the interviewing programs, respondents might be distracted, which in turn might a lower response quality.

In this way each survey mode is associated with an array of methodologically pertinent aspects; each generating effects on response behavior and data quality. All these aspects may cause in the results. We now focus on the telephone surveys conducted via landline or mobile phones. Despite these two modes being quite similar, Brick and Lepkowski's (2008) statement that "both modes may be subject to differential biases" (p. 156) holds.

16.4 Methods for Empirically Determining Response Quality

Our aim is the collection of empirical information about the response quality in telephone interviews. For this we need specially designed and to some extent very complex instruments. Various strategies designed for the CELLA 1 and 2 studies conducted in 2007/08 and 2010 (Häder and Häder 2009, 2011, Häder and Kühne 2010) have proven suitable for detecting such effects. Particularly splits and factorial designs were implemented for this purpose. The instruments of the CELLA studies are briefly introduced below.

High response quality may be expected if the respondent makes an effort to extract information from their memory while searching for an answer to the survey question. Accordingly, if the response is influenced by specific elements of the question wording, this is an indication of a lower response quality. When allocating random parts of the sample (split ballot) to different versions of questions of logically identical content, one can measure the strength of these wording effects. This way we can determine the conditions under which respondents react more or less strongly to the question wording. The extent to which the wording is irrelevant and responses are formulated on the basis of information retrieved from memory is an indicator of the response quality.

CELLA used various scales to record an answer. In this way scale effects were generated, i.e. depending on the implemented scale the answer distributions differ. We can examine in which mode (landline or mobile phone) the scale effects were strongest. This allows us to draw conclusions about the response quality.

Another effect may be generated by presenting indicators in a different order across the splits. This might cause order effects. The strength of these effects also gives an indication of the response quality within a certain survey mode. Stronger effects indicate lower response quality. If respondents ignore the question order, this implies a more intense cognitive processing of the question content and thus a higher response quality.

Another method of eliciting such effects is using logically identical but semantically different expressions in the question wording. In CELLA we replicated a well-known experiment (Rugg 1941), in which the expressions "not allow" and "forbid" are used interchangeably across the splits. These expressions tend to produce differential response behavior. The weaker this effect is the better the response quality; because we can assume that the weaker effect reflects more effort invested in the retrieval of information from memory.

Respondents tend to be hesitant to admit that they do not know the answer to a survey question. If such an admission is considered especially problematic, pseudo-opinions are generated. Respondents then also provide opinions about issues that do not exist. If a person responds to such a fake issue, one can assume a low quality of this answer (Kühne and Böhme 2006). More answers to questions about pseudo-issues point towards lower data quality.

Item nonresponse is another indicator of data quality. More item nonresponse, i.e. more questions not answered by the respondent, indicates a more superficial response behavior and the use of heuristics in the response process (Krosnick 2000, Shoemaker et al. 2002: 196). It might be the case that a telephone survey is perceived as more strenuous via one mode than via the other. Therefore, we might expect differential levels of item nonresponse and data quality across modes.

Persons differ in the extent to which they are influenced by socially desirable behavior. The more likely a respondent is to provide socially desirable answers the more biased their data will be. CELLA implemented special scales (SD scales) to measure social desirability bias in the answering behavior (see for example Krebs 1991, 1993 and first developments already in Edwards 1957 and Fordyce 1956). The data collection mode that generates fewer socially desirable answers produces a higher response quality.

As was pointed out above, more cognitive effort indicates a higher response quality. The more information a person extracts from their memory, the more substantially motivated their answer will be. CELLA directly asked respondents how many aspects of the previous questions the respondent still remembered. The number of (correct) aspects mentioned provides further evidence of the response quality (Häder and Kühne 2009).

In test-retest designs where the same issue is queried repeatedly, consistent answers point towards a higher quality of the cognitive processing and thus of the response than inconsistent answers. When using heuristics for finding the answer to a question the responses are more likely to be inconsistent then responses based upon careful reflection. This is yet another way of assessing the differential response quality in landline and mobile phone surveys.

Such complex instruments for determining the response quality can only be implemented in specially designed studies like CELLA. Another approach would be to implement a survey like the European Social Survey (ESS) in various modes and comparing the results across sub-samples by examining differences in the responses provided. Detected differences suggest the presence of mode effects. However, influences resulting from non-coverage and nonresponse would need to be controlled for. Jäckle et al. (2008) demonstrate such an approach.

16.5 Data Quality in Landline and Mobile Phone Surveys

Mixed-mode designs may lead to differences in data quality across modes. This is especially a concern for the response process, where different mode effects may influence the responses given by the sample unit (see section 16.3). However, particularly landline and mobile phone surveys may also be vulnerable to differential nonresponse and under-coverage. In the following, we summarize the mode effects found in the parallel implementation of landline and mobile phone surveys (see also Häder and Häder 2009, and Häder and Kühne 2009, 2010).

16.5.1 Effects due to a Differential Routine with the Survey Technology

Mobile telephones are now in widespread use across Europe (see country spotlights in this book). Nevertheless, there are certain segments of the population for whom a routine use of this device may not be assumed. Yet as Dillman (2000) showed, routine with using the medium of the survey is important for the quality of the data collected. For example, Dillman and his colleagues found such routine effects in web survey outcomes, which they then compared with outcomes from mail and face-to-face surveys (Christian et al. 2008: 257). A lack of routine has to be compensated for with increased cognitive efforts. These additional cognitive efforts can then not be invested in the response process. Thus, we expect that a lack of routine in using mobile phones will negatively impact on the quality of the responses given in mobile phone surveys. In contrast, for landline surveys we may expect sample units to be sufficiently skilled in using this technology.

At a closer look we notice a likely correlation between the level of routine with mobile phones and the respondents' age. For younger persons having grown up with mobile phone technologies receiving calls on a mobile phone has become common, in both daily life and the survey context. In addition to an age effect we find a gender gap; Men use mobile phones more routinely than women (Häder and Kühne 2010).

The findings from the CELLA study are generally quite positive regarding the data quality of mobile phone surveys. For example, a lack of routine does not lead to large biases in the survey results. Apparently, the technological demands of mobile phones are not so substantial as to negatively impact on surveys. This is somewhat contrary to findings from internet surveys, where such technology effects have been found. At present, the internet seems to be more technologically demanding than mobile phones.

The results show that routine is merely a hygiene factor (rather than a motivator, Herzberg et al. 1959). Routine is only relevant for the quality of survey results, if problems (for example technological problems) with the mobile phone occur. In other words, a lack of routine with mobile phones may negatively impact on response quality; however, an increased routine or perfection in the use of mobile phones does not lead to ever an increasing response quality.

Finally, one should point out coverage errors that may be associated with mobile phone surveys, since persons without mobile phone experience are very unlikely to participate in mobile phone surveys.

16.5.2 Effects Related to the Interview Environment

Any kind of survey constitutes an invasion in the respondent's privacy. Since survey requests via the mobile phone reaches sample units in a diversity of situations, interviews in this mode are likely to be subject to a diversity of environmental influences (Beck et al. 2005). This may well lead to increased cognitive demands on the response process and thus to a lower response quality. Tucker and Lepkowski (2008: 12) thus suspect that survey requests via a mobile phone intrude respondent's privacy even more severely than survey requests via a landline. Further-more, respondents to mobile phone surveys are more likely to be in presence of third parties when the survey request reaches them. We look into this in more detail further below.

In landline surveys, which are typically conducted at the respondents' homes, the survey environment is more familiar. Unexpected events are also less likely to occur during the interview when conducted at home than when conducted at less familiar locations. Furthermore, phone conversations have traditionally taken place within the household. Thus, we may expect fewer distractions by the environment. In addition, at home the respondent is more in control of the interview situation than outside the household.

The findings from the CELLA studies show the following. First, a large proportion of sample units reached via mobile phone were reached at home. This is true for both CELLA 1, where 63 % of the respondents were reached at home (n = 1.176; Häder et al. 2009: 77ff.) as well as CELLA 2, where 77 % of respondents were reached at home (n = 1.427, both data sets weighted), i.e. within the own household. Thus, the probability of an environment effect is limited.

Table 16.1 provides an overview of the locations in which the telephone interviews of the CELLA studies took place. Two issues are worth noting. First, landline interviews were frequently conducted via cordless phones and might thus also take place outside the home. Second, the calling times are important, since they determine the location of the respondent when called for an interview.

Table 16.1 Locations of interview during the CELLA 1 and 2 phone interviews by mode of call (in percent, weighted data sets)

	Mobile phone	Landline
at home, in the apartment	65	92
at work	14	0
on public transport (bus, tram, train)	2	-
in the car	5	-
in a shop, while shopping	2	0
in another location	7	1
in the garden, on the balcony, on the terrace	4	6
in the garage	0	0
in the basement, in the attic	0	0
off the property	2	0
n	2,603	2,240

The empirical findings for the response quality in CELLA 1 and 2 do not support the hypothesis that quality differs across interview locations. Looking at all phone interviews conducted at home and at those conducted outside the home we do not find an influence of the location on response quality. In CELLA 1 there is no association between the location of interview and the respondent’s cognitive efforts, the level of item nonresponse, the amount of socially desirable responses, or the stability of responses within the interview (Häder and Kühne 2009). In CELLA 2 we found that the level of item nonresponse is even a little lower when the interview is conducted outside of the household. Apart from this, no significant effects were found in this study.

16.5.3 Effects due to the Presences of Third Parties

The presence of third parties during the interview can lead to different effects across modes (Tourangeau et al. 2000: 280). This might be due to the different nature of relationships between the respondent and other persons present during the interview. While the persons present during landline interviews at home will usually be known to the respondent, the situation is quite different for mobile phone interviews. In addition to strangers that might listen in on the interview, people somewhat familiar to the respondent, such as colleagues at work, might be in earshot of the respondent.

Table 16.2 shows a comparison of telephone and face-to-face surveys in Germany with respect to third parties present at interview. This table shows that in the majority of interviews respondents are alone. One can also see that in a much smaller proportion of phone surveys outside of the household the respondent is alone. In addition, the composition of people present during the interview is somewhat different for phone interviews conducted outside of the household compared with interviews at home. However, also outside the household only 10 % of interviews are conducted with strangers present.

Table 16.2 Persons present during face-to-face and telephone interviews (in percent)

Mode of interview Location of interview	*Face-to-face* at home	*Telephone* at home	*Telephone* outside
Alone	73	65	58
Partner/Spouse	17		
Children	6	} 33	} 25
Other family	3		
Other persons including	1	2	18
Friends		(1)	(5)
Colleagues		(0)	(3)
Strangers		(1)	(10)
n	20,107	4,239	831

Sources: face-to-face: Summary of ALLBUS studies 1998 – 2008; telephone: CELLA 1 and CELLA 2, weighted data.

The direction of the influence of third party presences, however, is not predictable. One might assume that third party presence leads to socially desirable responses, since the respondent might be influenced by external opinions. Especially, during interviews on landline phones, those present might be family familiar with the respondent's opinions. The presence of a third party can then lead to a (positive) control. In private households the presence of third parties should thus lead to a softening of the tendency to give socially desirable answers (Preiser 2002). The subject of the question is relevant, as well. It is also plausible that especially with phone interviews third parties have no influence of the response behavior.

The CELLA studies showed that, indeed, the presence of other people during the interview has no clear influence on response behavior (Häder and Kühne 2009: 177). Neither the respondents' cognitive efforts, nor the level of item nonresponse seem to be influenced by the presence of other people, and no question order effects, question wording effects and effects of a pseudo-opinion question were found. The only effects found were that respondents needed less time to give an answer and were less likely to admit voting for a right extremist party, when third parties were present.

Our findings thus show that the presence of other people during the interview is only marginally relevant in telephone surveys. This is probably due to the questions not being overheard by others when on the phone. Another reason is that many answers are only given by means of number codes that have no meaning to eavesdroppers who have not heard the respective question text.

16.5.4 Effects due to the Activities That Need to Be Interrupted for the Interview

The types activities the respondent needs to interrupt when interviewed on the phone are likely to differ across modes, especially when the respondent in reached outside the home. This may have consequences for differential participation and response behavior. Table 16.3 lists the activities performed by the respondents when called for an interview in CELLA 2.

Table 16.3 Activities performed directly before the survey call in CELLA 2 by mode (in percent, weighted data)

	Mode	
Activity	Landline	Mobile
on the job – including school and university	13	21
was relaxing or performed a leisure activity, e.g. reading the newspaper, sports, social interactions, going for a walk	30	26
performed a personal activity, e.g. eating, drinking, sleeping	14	13
lost in thought, e.g. worrying, daydreaming	1	1
there was an unexpected external disruption, e.g. a visit, noise	0	0
performed other types of activities, e.g. housework such as cleaning, doing the dishes, running errands	17	13
distracted by physical complaints, e.g. illness, headaches, not feeling well	0	0
watched television	15	13
did something else	10	13
n	1,500	1,507

Table 16.3 shows that in the mobile phone survey more people were interrupted at their job than in the landline survey. For all other activities the differences are minimal, i.e. below four percentage points.

Using the CELLA 2 data we further looked into these results. Our analyses show that there were no differences in the response quality with regards to (a) the respondents' cognitive efforts, (b) social desirability bias, (c) reaction to a pseudo-opinion question, (d) the level of item nonresponse, as well as (e) reaction to differences in the question wording. Apparently, the response quality is insensitive to the types of activities interrupted for the interview.

Comparing telephone interviews with other modes used in social scientific research (see section 16.5) exhibited strong differences across modes. When comparing mobile with landline phone surveys, we need to also consider methodological differences which may impact on the response quality. The CELLA data allowed for more detailed analyzes of these mode effects. The analyses exhibited surprisingly few differences in response behavior. The multitude of methodological experiments in CELLA allow for a differentiated and reliable evaluation. These findings also correspond to those of other authors (Kuusela and Notkola 1999, Kuusela and Vikki 1999, Kuusela and Simpanen 2002, Kuusela 2003).

16.6 Personality Traits as Determinants of Response Quality in Telephone Surveys

Psychologists assume that an individual's personality can be expressed by means of five fundamental personality dimensions (Big Five) which function as predictors for various social scientific variables (Rammstedt 2007a). We thus looked into whether respondents' personality structures explain the quality of the answers given in a phone survey. Using the Big Five operationalization for surveys the fundamental personality structure can be determined empirically (McCrae and Costa 1987, McCrae and John 1992). According to the Big Five Theory there are five bipolar dimensions describing a person.

1. Extro- and introversion; described by terms such as outgoing and energetic versus solitary and reserved.
2. Agreeableness; depicted by adverbs such as interpersonal trust and cooperativeness versus cold and suspicious.
3. Conscientiousness; which concerns dutifulness and aiming for achievement versus careless and easy-going behavior.
4. Neuroticism; described by such expressions as sensitive and nervous versus secure and relaxed.
5. Openness; may be expressed by curiousness and inventiveness versus cautiousness and conservatism.

In the context of this chapter we establish whether these traits explain response behavior in phone surveys. It is also plausible that the methodological characteristics of phone surveys are more predictive of response quality than respondents' personality traits. In this latter case we need to explore, whether there are differences between mobile and landline phone surveys.

For the operationalization of the Big Five in CELLA 2 we implemented an instrument that has been frequently put into practice in the German Socio Economic Panel (GSOEP). In this instrument each dimension is reflected by three items (Gerlitz and Schupp 2005).

In the following we apply the same criteria to evaluating response behavior as before (see section 16.4) as dependent variables and conduct multivariate analyses. These results show that controlling for age and level of education none of the personality dimensions is associated with item nonresponse. Likewise, there is no relationship between the question wording and the personality traits. However, there are some significant findings with respect to a social desirability bias.

- The more extrovert persons are the less likely they are to practice self-deceptive enhancement (SDE), but they are more likely to impression management (IM).

- The more agreeable a person is the less likely their answers are to delude themselves or other.
- The more respondents report to be conscientious, the less likely they are to SDE and IM. Die traits of conscientiousness and endurance are connected to the personality dimension of self-monitoring (Rist et al. 2006: 69). This explains the association of these dimensions with socially desirable response behavior. If a respondent is very consciencious, both dimensions appear less strongly.
- The more neurotic respondents tend to be, the less likely they are to delude others and the more they delude themselves.
- The more open persons are the less self-delusion they practice.

With respect to personality traits we would like to also point towards a concern which might be relevant beyond the scope of telephone surveys. In recent years, researchers increasingly use personality traits as explanatory variables in social scientific research. Examples include attempts to (better) explain health behavior (Caspi et al. 2005, Smith and Spiro 2002) and electoral behavior (Heaven and Bucci 2001, Saucier 2000, van Hiel et al. 2004) by means of people's personality. Both are behaviors that are strongly influenced by norms and thus difficult to study empirically. However, the findings from on CELLA 2 suggest that one should be careful with detailed interpretations of these associations. The reason is the apparently strong association between self-reported personality traits and socially desirable response behavior.

Another criterion for response quality is the respondent's cognitive effort. Strongly extrovert persons tend to make more cognitive effort. Finally, we found that persons who are both more conscientious and more open are also more likely to answer to a pseudo-opinion question.

Other studies have found associations between certain personality traits and a person's punctuality, reliability and thoroughness (Barrick et al. 2001). Accordingly, in the context at hand there should be relationships between personality and item nonresponse (less), cognitive effort (stronger), reaction to a pseudo-opinion question (less), as well as a reaction to the question wording (weaker). However, as shown none of these expectations were fulfilled.

Such a lack of significant effects is generally good news for survey research. It indicates that respondents react to the content of each question, i.e. they provide the true value. Methodological aspects of the question design and the response model seem to play a subordinate role when providing an answer. The following should be emphasized:

For telephone surveys run in parallel via mobile and landline phones the mode has not been found to play a major role in explaining the various dimensions of response quality by means of regression models. We only found that on a landline phone the cognitive efforts made by respondents were somewhat higher than on mobile phones and we found fewer cases of deluding others.

Interestingly, the control variables (age, gender, and education) included in the models quite strongly affect the response quality. This is a concern for survey research, since – independent of the question content – results will be biased if women exhibit a different response behavior than men, if the higher educated react differently to questions than the lower educated, and if younger persons respond differently from older persons.

If we can find reference studies which we can compare the distribution of characteristics in CELLA 2 to, we can also infer associations between the nonresponse process in telephone surveys and personality traits. Comparing demographic characteristics in CELLA 2 with those in the GSOEP provides initial information about differences between population and sample characteristics. In addition, the data from the International Social Survey Programm (ISSP) can provide another reference. However, the ISSP only collected the 10-item Big Five Inventory. Consequently, only summary results may be compared (Rammstedt 2007b).

We can see clear differences between the data sources when comparing CELLA 2 and GSOEP with regards to the agreeableness dimension. In telephone-administered surveys this dimension is more prevalent. In addition, a comparison with the ISSP presents the same picture.

When interpreting these results one needs to be careful again. Both reference studies are conducted face-to-face. It might well be that the decision to participate in a telephone survey is made more spontaneously than the decision to participate in the face-to-face surveys GSOEP or ISSP. In both GSOEP and ISSP the interviewer's visit to the sample unit may be announced with an advance letter and a telephone call. For practical reasons this was not possible in CELLA 2 and is generally difficult or impossible in telephone surveys in Germany. In consequence, participation in telephone surveys might necessitate a higher level of interpersonal trust and willingness to cooperate. These are traits typically found in persons with a high level of agreeableness on the Big Five personality scale. To provide more insight into this issue more research and empirical evidence is needed.

16.7 Outlook

This chapter investigates the data quality achieved in telephone surveys via mobile and landline phones. First, it is worth noting that the difference in modes induces methodological differences across various aspects. One would thus expect the data quality to be quite different across the two modes. The CELLA 1 and 2 studies were specifically designed to research such questions and provide methodological insights which are presented in this chapter. While some questions were answered by these studies, others demand further research.

Our main conclusion is that the data quality in mobile phone surveys does not differ significantly from the data quality in landline surveys. Only few indications of modes effects were found. These results also support other researchers' findings. Nevertheless, there is further need for research into the influence of personality traits on data quality, since we were only able to present some initial findings here. Future research should also consider additional aspects of respondents' personality over and above those measured by the Big Five.

References

Aquilino, W.S.: Effect of Spouse Presence During the Interview on Survey. Responses Concerning Marriage. Public Opinion Quarterly 57(3), 358–376 (1994)

Barrick, M.R., Mount, M.K., Judge, T.A.: Personality and performance at the beginning of the new millennium: What do we know and where do we go next? International Journal of Selection & Assessment 9, 9–30 (2001)

Beck, F.: Legleye, St., Peretti-Watel, P.: Aux abonnés absents: liste rouge et téléphone portable dans les enquêtes en population générale sur les drogues (Been conspicuously absent: unlisted and cell phone in the general population surveys on drugs). Bulletin de Méthodologie Sociologique 86, 5–29 (2005)

Bishop, G.F.: Experiments with the middle response alternative in survey questions. Public Opinion Quarterly 51(2), 220–232 (1987)

Böcker, F.: Scale forms and their impact on rating's reliability and validity. Journal of Business Research 17(1), 15–26 (1988)

Bosnjak, M.: Teilnahmeverhalten bei Web-Befragungen. Nonresponse und Selbstselektion (Participation behavior in web surveys. Nonresponse and self-selection). In: Theobald, A., et al. (eds.) Online-Marktforschung, pp. 79–95. Gabler, Wiesbaden (2001)

Brick, J.M., Lepkowski, J.M.: Multiple Mode and Frame Telephone Surveys. In: Lepkowski, J.M., et al. (eds.) Advances in Telephone Survey Methodology, Wiley, New York (2008)

Caspi, A., Roberts, B.W., Shiner, R.L.: Personality Development: Stability and Change. Annual Review of Psychology 56, 453–484 (2005)

Church, A.H.: Estimating the effect of incentives on mail survey response rates: A meta-analysis. Public Opinion Quarterly 57(1), 62–79 (1993)

Christian, L.M., Dillman, D.A., Smyth, J.D.: The Effects of Mode and Format on Answering to Scalar Questions in Telephone and Web Surveys. In: Lepkowsky, J.M., et al. (eds.) Advances in Telephone Survey Methodology, pp. 250–274. Wiley, New York (2008)

De Leeuw, E.D.: Data quality in mail, telephone, and face-to-face surveys. TT Publications, Amsterdam (1992)

De Leeuw, E.D.: To mix or not to mix data collection modes in surveys. Journal of Official Statistics 21(5), 233–255 (2005)

Diekmann, A., Jann, B.: Anreizformen und Ausschöpfungsquote bei postalischen Befragungen. Eine Prüfung der Reziprozitätshypothese (Incentive forms and response rate for postal surveys. An examination of Reziprozitätshypothese). ZUMA-Nachrichten 25(48), 18–27 (2001)

Dillman, D.A.: Mail and Internet Surveys. The Tailored Design Method. Wiley, New York (2000)

Dillman, D.A., Brown, T.L., Carlson, J., Carpenter, E.H., Lorenz, F.O., Mason, R., Saltiel, J., Sangster, R.L.: Effects of category order on answers to mail and telephone surveys. Rural Sociology 60, 674–687 (1995)

Edwards, A.L.: The Social Desirability Variable in Personality Assessment and Research. Dryden, New York (1957)

Fordyce, W.E.: Social Desirability in the MMPI. Journal of Consulting Psychology 20(3), 171–175 (1956)

Gerlitz, J.-Y., Schupp, J.: Zur Erhebung der Big-Five-basierten Persönlichkeits-merkmale im SOEP. Dokumentation der Instrumentenentwicklung BFI-S auf Basus des SOEP-Pretests (The survey of the Big Five personality traits based on SOEP. Documentation of instrument development BFI-S Basu on the SOEP pretest, DIW- Research Notes 4 (2005), http://www.diw.de/documents/publicationen/73/43490/rn4.pdf (accessed July 4, 2011)

Goyder, J.: The silent minority: Nonrespondents on sample surveys. Polity Press, Cambridge (1987)

Groves, R.M.: Survey errors and survey costs. Wiley, New York (1989)

Häder, M., Häder, S.: Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung (Telephone interviews over the air. Concept, design and implement a strategy for data collection). VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Häder, M., Kühne, M.: Mode-Effekte (mode-effects). In: Häder, M., Häder, S. (eds.) Telefonbefragungen Über Das Mobilfunknetz. Konzept, Design Und Umsetzung Einer Strategie Zur Datenerhebung, pp. 165–242. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Häder, M., Kühne, M.: Mobiltelefonerfahrung und Antwortqualität bei Umfragen (Mobile phone experience and quality of response in surveys). Methoden Daten Analysen mda, 2 (2010)

Häder, S., Häder, M., Graeske, J., Kunz, T., Schneiderat, G.: Realisierung der Stichprobe (Realization of the sample). In: Häder, M., Häder, S. (eds.) Telefonbefragungen über Das Mobilfunknetz. Konzept, Design und Umsetzung Einer Strategie Zur Datenerhebung, pp. 71–82. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)

Heaven, P., Bucci, S.: Right-wing authoritariannism, social dominance orientation and personality: An analysis using the IPIP measure. European Journal of Personality 15, 49–56 (2001)

Herzberg, F., Mausner, B., Snyderman, B., Bloch, B.: The Motivation to Work, vol. 2. Wiley, New York (1959)

Hyman, H.H., Sheatsley, P.B.: The Current Status of American Public Opinion. In: Payne, J.C. (ed.) The Teaching of Contemporary Affairs. Twenty-First Yearbook of the National Council of Social Studies, pp. 11–34. National Education Association, New York (1950)

Jäckle, A., Roberts, C., Lynn, P. (2008), Assessing the Effect of Data Collection Mode on Measurement. ISER Working Papers 2008-08, http://www.iser.essex.ac.uk/publications/working-papers/iser/2008-08 (retrieved March 28, 2010)

Jobe, J.B., Herrmann, D.J.: Implications of models of survey cognition for memory theory. In: Herrmann, D.J., et al. (eds.) Basic and Applied Memory Research, Practical Applications, vol. 2, pp. 193–205. Erlbaum, Hillsdale (1996)

Krosnick, J.A., Alwin, D.F.: An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. Public Opinion Quarterly 51(2), 201–219 (1987)

Krausch, S.: Verzerrungen in Telefonumfragen: eine Analyse zum Einfluss der Einstellung des Interviewers (Distortions in telephone surveys: an analysis of the influence of attitude of the interviewer). VDM Verlag Dr. Müller, Saarbrücken (2007)

Krebs, D.: Was ist sozial erwünscht? Der Grad sozialer Erwünschtheit von Norbert Schwarz. ZUMA-Arbeitsbericht, 18. GESIS-ZUMA, Mannheim (1991)

Krebs, D.: Social Desirability: The Collective Conscience? Judging the Degree of Social Desirability in Attitude Items. ZUMA-Arbeitsbericht,16. GESIS-ZUMA, Mannheim (1993)

Krosnick, J.A.: The threat of satisficing in surveys: The shortcuts respondents take in answering questions. Survey Methods Newsletter 20(1), 4–8 (2000)

Kühne, M., Böhme, R.: Effekte von Informationsstand, Wissen und Einstellungsstärke von Befragten auf die Antwortstabilität (Effects of level of information, knowledge and attitudes of respondents on the strength of response stability). ZUMA-Nachrichten 59, 42–71 (2006)

Kuusela, V., Notkola, V.: Survey quality and mobile phones. Paper presented at the International Conference on Survey Nonresponse, Portland, OR (1999)

Kuusela, V., Vikki, K.: Change of telephone coverage due to mobile phones. Paper presented at the International Conference on Survey Nonresponse, Portland, OR (1999)

Kuusela, V., Simpanen, M.: Effects of mobile phones on phone survey. Practices and results. Paper Presented at the International Conference on Improving Surveys (ICIS), Copenhagen, Denmark (2002)

Kuusela, V.: Mobile phones and telephone survey methods. In: Banks, R., et al. (eds.) Proceedings of the 4th ASC international Conference ASC 2003 - The Impact of New Technology on the Survey Process, pp. 317–327. Association for Survey Computing (ASC), Chesham Bucks (2003)

McCrae, R.R., Costa Jr., P.T.: Validation of the five-factor model of personality across instruments and observers. Journal of Personality and Social Psychology 52, 81–90 (1987)

McCrae, R.R., John, O.P.: Introduction of the five-factor model and its applications. Journal of Personality 60, 175–215 (1992)

Noelle-Neumann, E., Petersen, T.: Das halbe Instrument, die halbe Reaktion. Zum Vergleich von Telefon- und Face-to-Face-Umfragen (Half the instrument, the half reaction. For comparison of telephone and face-to-face surveys). In: Hüfken, V. (ed.) Methoden in Telefonumfragen, pp. 183–200. Westdeutscher Verlag, Wiesbaden (2000)

Payne, S.L.: The art of asking questions. University Press, Princeton (1949)

Preiser, W.: Anwesenheitseffekte als Indikator sozial erwünschten Antwortverhaltens im Interview. Eine Untersuchung der Verzerrung von Angaben zur Mediennutzung (Presence effects socially desirable response behavior as an indicator in the interview. A study of the distortion of information for media use). (2002), `http://www2.dgpuk.de/fg_meth/fg_tag02_abs08.htm` (accessed July 4, 2011)

Rammstedt, B.: Welche Aussagekraft hat die individuelle Persönlichkeit für inhaltliche sozialwissenschaftliche Variablen (What is the significance of individual personality for substantive social science variables?). ZUMA-Arbeitsbericht, 1 (2007a)

Rammstedt, B.: The 10-Item Big Five Inventory. European Journal of Psychological Assessment 23(3), 193–201 (2007b)

Redline, C., Dillman, D.A.: The influence of alternative visual designs on respondents' performance with branching instructions in self-administrated questionnaires. In: Groves, R.M., Dillman, D.A. (eds.) Survey Nonresponse. Wiley, Chichester (2002)

Reuband, K.-H.: Interviews, die keine sind. " und " beim Fälschen von Interviews (Interviews, which are not. "Successes" and "failures" in the forging of interviews). Kölner Zeitschrift für Soziologie und Sozialpsychologie 42/4, 706–733 (1990)

Rist, F., Engbering, M., Patzelt, J., Beißner, J.: Prokrastination als verbreitete Arbeitsstörung. Aber morgen fange ich richtig an (Procrastination as a common labor disturbance. But tomorrow I'll start properly). Personalführung 6, 64–78. (2006), http://wwwpsy.uni-muenster.de/imperia/md/content/psychotherapie_ambulanz/prokrastination_personalfuehrung.pdf (accessed July 4, 2011)

Rugg, D.: Experiments in wording questions. Public Opinion Quarterly 5(1), 91–92 (1941)

Saucier, G.: Isms and the structure of social attitudes. Journal of Personality and Social Psychology 78, 366–385 (2000)

Schnell, R.: Der Einfluss gefälschter Interviews auf Survey-Ergebnisse (The impact of counterfeit interviews on survey results). Zeitschrift für Soziologie 20(1), 25–35 (1991)

Schuman, H. (1992). Context effects: State of the past / State of the art. In: Schwarz, N., Sudman, S. (eds.) Context effects in Social and Psychological Research, pp. 5–20. Springer, New York (1992)

Schuman, H., Presser, S.: Questions and Answers in Attitude Survey. Sage Publications, Inc. (1981)

Schwarz, N., Hippler, H.-J., Deutsch, B., Strack, F.: Response scales: Effects of category range on reported behavior and comparative judgements. Public Opinion Quarterly 49(3), 388–395 (1985)

Shoemaker, P.J., Eichholz, M., Skewes, E.A.: Item Nonresponse: Distinguishing between don't know and refuse. International Journal of Public Opinion Research 14(4), 193–201 (2002)

Smith, T.W., Spiro, A.: Personality, health, and aging: Prolegomenon for the next generation. Journal of Research in Personality 36, 363–394 (2002)

Stadtler, K.: Die Skalierung in der empirischen Forschung (the scaling in the empirical research). Infratest, München (1983)

Strack, F.: Zur Psychologie der standardisierten Befragung. Kognitive und kommunikative Prozesse. On the Psychology of the Standardized Questionnaire Cognitive and Communicative Processes. Springer, Berlin (1994)

Strack, F., Martin, L.L.: Thinking, judging, and communicating: A process account of context effects in attitude surveys. In: Hippler, H.-J., et al. (eds.) Social Information Processing and Survey Methodology, pp. 123–148. Springer, New York (1987)

Sudman, S., Bradburn, N.M., Schwarz, N.: Thinking about Answers. The Application of Cognitive Processes to Survey Methodology. Jossey-Bass, San Francisco (1996)

Tourangeau, R., Rips, L.J., Rasinski, K.A.: The psychology of survey response. University Press, Cambridge (2000)

Tucker, C., Lepkowski, J.M.: Telephone Survey Methods: Adapting to Change. In: Lepkowsky, J.M., et al. (eds.) Advances in Telephone Survey Methodology, pp. 3–26. Wiley, New York (2008)

Van Hiel, A., Pandelaere, M., Duriez, B.: The impact of need for closure on conversative beliefs and racism: Differential mediation by authoritarian submission and authoritarian dominance. Personality and Social Psychology Bulletin 30, S. 824– S. 837 (2004)

Wüst, A.M.: Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften als Telefonumfrage (The General Social Survey as a telephone survey). ZUMA-Arbeitsbericht 4 (1998)

Part V
Recommandations for Telephone Surveys

Chapter 17
The Impact of Introductions in Telephone Surveys

Gerd Meier

17.1 Introduction

Over the last decades, response rates in market and social surveys have been declining continuously. Given that the response rate is one of the most critical methodological issues, a special conference aimed at professionals from the top agencies in market research was devoted to this topic in 2008. The conference title was: "Where have all the respondents gone?" All participants observed a serious loss in respondents' cooperation (AMA and SPSS 2008), although there was no agreement on which response rate level is still acceptable: governments demand 85 %, academics accept about 50 % for their publications, and for an average marketing study 20 % seems to be a good level. At the end of the conference, a discussion took place on how to engage with selected people and make them want to interact with interviewers. This discussion culminated in a call for investigating and working out the basics of how to interact with people in today's world.

Curtin et al. (2005) analysed the refusals in the University of Michigan's Survey of Consumer Attitudes (SCA) over a wide range of years. They reported that nonresponse rates in telephone surveys rose even more steeply in the last decade, mainly due to a rise in refusals. The authors concluded that without better approaches to persuading the selected people to take part in a survey, "the longterm future of telephone survey research does not appear promising" (Curtin et al. 2005: 97).

With this development in mind, it is worthwhile to consider effective techniques that can stop this decline. The quality of surveys is not the only aspect to suffer from low response rates; economic efforts also need to be increased to counterbalance insufficient respondent cooperation.

17.1.1 Techniques for Increasing Respondent Cooperation

The conference's call for more research confirmed findings from a study that experimentally investigated the interviewer-respondent interaction on the telephone and that was conducted by us (Meier et al. 2005). We were convinced that the crucial period in this interaction is the introduction and that there are ways to optimise the beginning of interviews and thus to obtain better response rates. Some years back, Dillman et al. (1976) showed that refusals occur mainly in the introduction period and rarely afterwards.

A promising strategy for interviewers is to build on social reciprocity during the introduction. Reciprocity is a well-known phenomenon in social psychology. It means that someone who receives something from someone else feels obliged to return the favor. Gouldner (1960) was the first author to identify reciprocity as a universal norm: “Generically, the norm of reciprocity may be conceived of as a dimension to be found in all value systems and, in particular, as one among a number of ‘Principal Components’ universally present in moral codes.” (Gouldner 1960: 171). He continued stating that “[t]he conclusion is clear: if you want to be helped by others you must help them. …Insofar as men live under such a rule of reciprocity, when one party benefits another, an obligation is generated. The recipient is now indebted to the donor, and he remains so until he repays.” (Gouldner 1960: 173f.).

This position was elaborated on by Hendrickson and Goei (2009): “Thus, the indebtedness explanation poses an aversive arousal reduction model for human reciprocity in which receiving a favor induces the aversive state of indebtedness that can be eliminated or reduced by subsequently complying with a direct request from a benefactor.” (Hendrickson and Goei 2009: 588). Keysar et al. (2008) compared reciprocity to trading in economic exchange: “In economic exchange, trading is enabled by the shared understanding that a good or service will be provided in exchange for its market value. In social exchange, ‘trading’ is enabled by the universal norm of reciprocity – the shared understanding that resources given or taken will be recouped.” (Keysar et al. 2008: 1280). How is such a norm which affects all people established? Whatley et al. (1999) attributed it to a learning or evolutionary process: “Returning a kindness is simply the right and proper thing to do. One could argue that reciprocity is one of the earliest and most far-reaching of moral precepts taught to children. All major religious and ethical traditions include reciprocity as a prime, if not golden, rule of moral behavior. … Such well-learned and accepted norms may tend to guide our behavior in an almost reflexive way. …It should be to our individual and collective advantage to want both to help others and to reciprocate their help and so any natural tendencies to feel such urges should be selected for in evolutionary terms.” (Whatley et al. 1999: 252). In conclusion the authors noted: “…the norm of reciprocity appears to explain an extraordinary amount of human behavior and is arguably one of the most fundamental and pervasive of all social norms.” (Whatley et al. 1999: 258).

Regan (1971) developed a design for studies on compliance-seeking, arguing that these studies are “of practical importance because there are many situations where we want compliance or assistance from another person, but either do not have or choose not to use the resources which would put a large amount of pressure on him. In such situations, a favor might be a practicable technique for increasing compliance.” (Regan 1971: 627). This is exactly the situation we find in market research, where an interview is requested without the possibility of forcing the respondent to comply. Regan and many other researchers were successful in demonstrating the effect of a favor on compliance. Up to now there is overwhelming evidence for the reciprocity effect (for a review of these studies, see Hendrickson and Goei 2009).

In a recently published paper, Skageby (2010) showed that information can serve as a gift and that reciprocity helps to build up online networks. In our study, we assumed the same. Hence, one treatment in our experiment was to give a piece of information in the hope that it functions as a gift, making people feel obliged to return the favor by being interviewed.

In a second treatment, we mentioned a university, a non-profit organization, and a profit-oriented market research institute as sponsors of the survey to test whether the sponsor has any effect on response rates, but we found no differences between them. This was quite surprising given the evidence in the literature that the sponsor is an important factor for cooperation in surveys (Goyder 1987, Meegama and Blair 1999).

The first treatment, however, was successful. We found that the response rate in telephone surveys is enhanced if the interview is opened with a simple question and some related information on an issue of common interest (see Figure 17.1). The introductory question "Do you enjoy eating in restaurants?" was accompanied with the information "I can tell you that about 25 % of Germans love eating in restaurants."

Furthermore, the respondent characteristics were more similar to the population characteristics if the introductory question and information were given. There

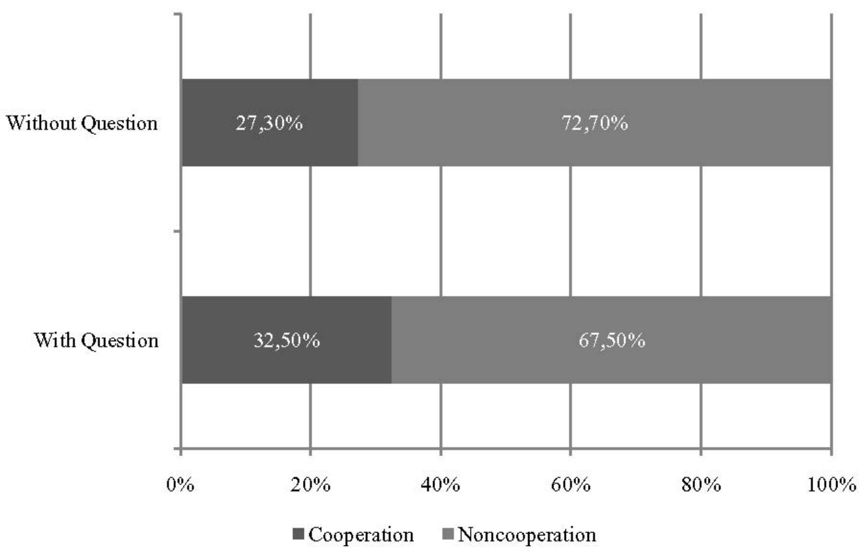


Fig. 17.1 Cooperation and non-cooperation in traditional interviews (without an introductory question; n = 523 = 100 %) and in interviews opened with a question and information (n = 517 = 100 %)

were more male respondents and people with lower education in the sample than usual, both traditionally underrepresented groups in surveys.

It is not yet understood, however, which underlying process fosters interviewees' cooperation. We hypothesized that it is the reciprocity rule, that is, the respondents gain information and feel obliged to return the favor. But it is also possible that an easy-to-answer question may signal that the interview is interesting to participate in and that little effort is needed to meet its requirements.

We therefore designed another experiment to differentiate between these two possible effects.

17.2 Cooperation Experiment with Different Introductions

The second experiment was designed to distinguish between these two possible effects. Apart from a control condition using a traditional introduction with no further question and information, three variations for the experimental variable "introduction" were used:

- information without a question,
- an easy-to-answer question with information, and
- a difficult question with information.

If the reciprocity rule alone is decisive for enhancing cooperation, no differences should occur between the experimental conditions because information was given to participants in all three groups.

If the question and information serve as signals that the interview is easy and interesting, then the group with an easy-to-answer question should show the highest cooperation rate. In that case the interviewer serves as a model by answering an easy question and modelling enhances the cooperation rate. If modelling alone is decisive for increasing the response rate, the other two experimental conditions, information without a question and difficult question, should not differ from the control condition.

Bandura (1977) developed the theory of model learning and stressed that effective modeling leads to an efficacy expectation, the conviction that one can successfully perform as requested. Only if self-efficacy is high enough will humans show the behavior. Otherwise humans will leave the situation: in the case of interviewing, respondents will refuse.

Furthermore we tested whether topical questions lead to higher cooperation than less topical ones and whether mentioning that the interview is conducted by a representative of a university also leads to higher cooperation. The latter treatment is similar to the sponsor treatment in our first experiment. Evidence for the effectiveness of the treatment "topicality" is provided in a study by Stroud and Kenski (2007). They found that media coverage of a specific topic decreases refusals to a survey request.

We created a 3 by 2 by 2 experimental design with the independent variables "introduction", "topicality", and "affiliation of the interviewer". Additionally, we used traditional introduction as a control condition, half of the time with interviewers as representatives of a university. Altogether, the experimental design consisted of fourteen groups (see Table 17.1).

Table 17.1 Study design and sample size

Introduction	Topicality	Interviewer as a member of the university	
		Yes	No
Information without question	Yes	n = 5	n = 5
	No	n = 5	n = 5
No-brainer question with information	Yes	n = 5	n = 5
	No	n = 5	n = 5
Difficult question with information	Yes	n = 5	n = 5
	No	n = 5	n = 5
Traditional interviews		n = 5	n = 5

We conducted telephone interviews with 70 German-speaking subjects, at least 18 years old, randomly and equally assigned to the experimental conditions. There were five interviews in each group. The questionnaire consisted of the manipulated introduction and the collection of demographic data as well as questions on the topicality and the experienced difficulty of the given information, enabling us to check the validity of the experimental manipulation. In the traditional interviews, only the demographic data were collected. A request was regarded successful if the respondent answered positively to the question: "May I ask you some questions?" This was asked after the manipulated introductions or, in the traditional interviews, after the first two sentences. If respondents were undecided, the interviewers tried to convince them to cooperate. The dependent variable was not the successful request or the refusal rate, but rather the number of refusals until a request was successful.

Only after the entire introduction was over we counted the refusals. In our opinion, the overall count of refusals is not a good operational definition of the dependent variable to test our hypotheses. A refusal within the first two sentences of the introduction means that the selected person cannot react to the manipulated conditions. Indeed, most refusals occurred after introducing the name of the interviewer in the first sentence, and these refusals varied randomly between the conditions, as did all other causes for failures, such as engaged lines or nobody at home. Each contact attempt was conducted with a different telephone number.

Thus, the dependent variable varied between zero, which means that the first contact attempt was successful and there was no refusal at all after the introduction, and the maximum value of nine, which means the tenth contact attempt was successful and there were nine refusals after the introduction. We decided upon this kind of dependent variable because counting results in a ratio scale, which enabled us to analyze the data by analysis of variance and t-tests.

The samples of telephone numbers were drawn from the universe of all possible landline numbers in Germany. The frame was constructed using the Gabler-Häder method (Häder and Gabler 1998).

Three students with no experience in interviewing operated as interviewers. All interviewers used all introduction texts. One interviewer was not informed about the hypotheses. There were no differences in performance between the interviewers. Hence, there was no evidence for the Rosenthal effect, that is, that the investigator unconsciously influences the results in favor of the hypotheses.

Although this study can be understood as a replication of the first one, there were differences between the designs: the sample size was small, non-professionals acted as interviewers, and operationalization of the dependent variable differed. But the most important difference was that the study was designed to test two possible behavior tendencies: reciprocity and modelling.

17.2.1 Introduction Texts

In the half of the interviews in which the interviewer was identified as a representative of a university, the interviews began with: “Hello, this is ... from the Leuphana University. I‘d like to ask you some questions for a survey.” In the other half of the interviews, the university was not mentioned. After these first two sentences, the introduction was manipulated with regard to topicality and information given. In the traditional interviews, the interview started after the two sentences with the question: “May I ask you some questions?” If the respondent agreed, the contact attempt was successful and demographic data were collected. Below I list three examples of combinations of the two treatments “topicality” and “information” in the manipulated introductions:

- Only information on a topic currently intensely debated without any further question: “... Bonuses awarded to top managers are a hot topic nowadays. 55 % of the German population don‘t know that for companies with state shareholding the government decides on bonuses.”
- Simple question combined with non-topical information: “Healthy eating has become an important topic in Germany. How often do you eat a warm meal? Few people know that 36 % of Germans eat more than two warm meals a day.”
- A difficult question combined with non-topical information: “Healthy eating has become an important topic in Germany. Do you know the average daily calorie consumption? For your information: 43 % of Germans do not know that the average calorie consumption of middle-aged people ranges between 1200 and 1500 kilocalories a day.”

17.3 Results

17.3.1 Manipulation Check

Respondents were asked on a seven-point scale about the topicality of the issue mentioned in the introduction. An answer of seven meant that topicality was very high. The difference between the groups was significant. The mean in the low-topicality group was 4.33 ($n = 30$; s. d. = 1.52) and in the high-topicality group the mean was 5.43 ($n = 30$; s. d. = 1.31). Because there was no introduction with topical or non-topical issue in the traditional interviews, respondents to these were excluded from the analysis. The results indicate that the manipulation worked.

However, the manipulation of difficulty did not work as expected. The respondents of the simple-question group and the difficult-question group on average rated the difficulty 1.95 and 2.55, respectively. This difference was expected, because a higher value means higher difficulty. But the mean rating of the respondents of the information-only group, 3.85, was the highest one, although we presented the same information as in the easy-to-answer-question group. Starting an interview with information and no question appears to be somewhat irritating and hence difficult for the respondent.

17.3.2 Refusal Rate

We counted 67 refusals after introduction, which is about 18 % of all refusals. The descriptive statistics of the refusals after introduction are given in Table 17.2.

Table 17.2 Mean frequency and standard deviation of refusals after introduction.

Group	n	Mean frequency	Standard deviation
Information without question	20	0.4	0.68
No brainer question with information	20	0.5	0.83
Difficult question with information	20	0.9	1.33
Traditional interviews	10	3.1	4.04

As shown in Figure 17.2 the mean number of refusals after the introduction and hence the cooperation rate varies between the groups ($F(3,66) = 6.12$; $p < 0.01$; eta = 0.47). The lowest frequency of refusals and hence the highest cooperation rate was reached in interviews that opened with information or with a simple question and subsequent information. A priori planned contrasts show that all the experimental conditions, regardless of the variation used, achieved significantly fewer refusals after the introduction and hence higher cooperation rates than the control condition not opened with a question and information ($t(66) > 3.24$; $p < 0.01$). This holds true even when the tests for unequal variances are used.

In contrast to our hypotheses, the topicality of information and the mentioning of the university in the introduction had no significant influence on the frequency of refusals after introduction, although the differences were in the expected direction.

We used a dependent variable on a high measure scale in order to test for interaction effects in an analysis of variance procedure. Yet again there were no significant interaction effects. Perhaps the sample size was too small to reveal the effects of topicality, university affiliation or interaction.

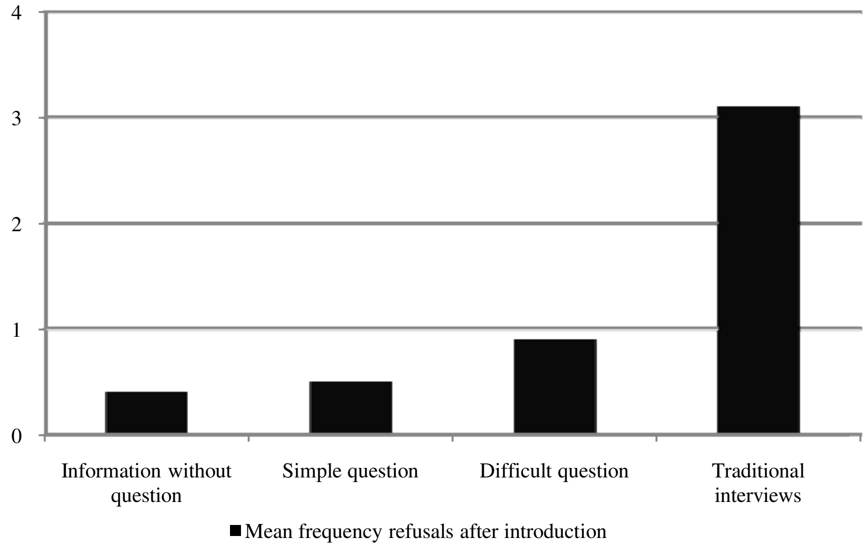


Fig. 17.2 Refusals after introduction in all four conditions

17.4 Discussion

In two experiments with different designs, we showed that a question and information within the introduction of an interview lead to a higher cooperation rate and hence reduced refusals. In the first experiment, professional interviewers conducted the interviews. The sample size was large, with more than three hundred interviews, and the dependent variable was the refusal rate itself. Because only a traditional introduction and an introduction with an easy-to-answer question together with information were tested, it was not clear which process fosters this effect. As hypothesized, the reciprocity rule could explain this effect. However, it is also possible that modelling decreases refusal rate insofar as the respondent learns from the interviewer how easy and interesting it is to answer the questions. Modeling enhances self-efficacy (Bandura 1977).

The second experiment was designed to distinguish between these two possible processes by comparing three different introductions with traditional interviews. In the first introduction, no question but only information was given. The second introduction consisted of an easy-to-answer question and information. The third introduction comprised a difficult question together with the information. All three variations were based on the reciprocity rule; the simple question alone represents the modelling condition because in the other introductions, a respondent cannot learn from the interviewer that the interview will be easy and interesting.

All three introductions based on the reciprocity rule resulted in lower refusal rates than with the traditional introduction. These results confirm the results of the first experiment, although in the second experiment non-professional interviewers were used, the sample size was noticeably smaller and the dependent variable was defined differently (as the number of refusals after introductions). The small sample size indicates that the reciprocity rule results in a very strong effect. Hence, this study contributes another piece of evidence in favor of the reciprocity effect, allowing us to conclude with the words of Hendrickson and Goei (2009) “In the study of compliance-seeking message effects, this reciprocity bent is manifested in several studies that show, among strangers, a target is more likely to comply with a direct request that is preceded by a favor than a direct request alone, even when the favor is unsolicited.” (Hendrickson and Goei 2009: 585).

Because it is not possible to create a modelling condition without giving respondents information, without information the interview would begin traditionally, the condition with the easy-to-answer question together with the information consists of both processes: reciprocity and modelling. This introduction was not more successful than when information alone was given. In the latter condition there was no modelling: the respondent could not learn from the interviewer that the interview is easy and interesting. Therefore, we conclude that reciprocity is crucial to decreasing refusal rate.

At the same time, modelling is not without any effect. This can be seen in the condition with the difficult question. If reciprocity alone were decisive, this condition should not differ from the condition with the simple question. But a difference does exist: if there is a difficult question in the introduction, the refusal rate is higher. We thus conclude that the respondents learnt from the interviewer in the condition with an easy-to-answer question and felt confident enough to answer the subsequent questions. The process, which can be described as model learning (Bandura 1977), increases perceived self-efficacy.

It appears that both hypothesized processes contribute to the higher cooperation rate, although reciprocity seems to have a more pronounced influence.

For practitioners, differentiating between reciprocity and modelling may seem of only academic value. From their perspective, it is more important to know that introductions with an easy-to-answer question together with information provided to the respondent are much better than traditional introductions; why they work better is less interesting. However, from our point of view it is essential that all attempts at improving introductions are grounded on well-founded theories, because this allows for better predictions and for further improvements when theoretical progress is made. The attempts are of only limited worth if they are grounded on empirical data alone.

The topicality of information had no effect on the cooperation rate. This may be due to the fact that the difference in perceived topicality between the groups was not high, albeit significant. To determine which topics best heighten the cooperation rate further research is necessary.

We therefore highly encourage further research to find the optimal introduction. For practical purposes, it is necessary to develop introductions which fit legal demands and requirements of market research associations. Moreover, the introduction should not affect the answers to the subsequent questions. This means that the information given should not be related to the topic of the questionnaire, but should nevertheless be interesting.

As decreasing cooperation rates show, the majority of people nowadays do not see taking part in an interview as a social obligation. Nonetheless, more people agree to an interview request when they are given interesting information. If using the reciprocity rule becomes good practice in the future, people may expect interesting information at the beginning of an interview. Further improvements may be possible, if people know in advance that they will receive this information when selected for an interview. This could be an issue for public relations campaigns. Then people will cooperate not out of a sense of social obligation, but rather because of the interesting information they obtain.

References

AMA & SPSS. Where have all the respondents gone? Marketing Research 20, 22–26 (2008)

Bandura, A.: Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review 84, 191–215 (1977)

Curtin, R., Presser, S., Singer, E.: Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly 69, 87–98 (2005)

Dillman, D.A., Gallegos, J.G., Frey, J.H.: Reducing refusal rates for telephone interviews. Public Opinion Quarterly 40, 66–78 (1976)

Gouldner, A.W.: The norm of reciprocity: A preliminary statement. American Sociological Review 25, 161–178 (1960)

Goyder, J.: The Silent Minority: Non-Respondents on Sample-Surveys. Polity Press, New York (1987)

Häder, S., Gabler, S.: Ein neues Stichprobendesign für telefonische Umfragen in Deutschland (A new sample design for telephone surveys in Germany). In: Gabler, S., et al. (eds.) Telefonstichproben in Deutschland, pp. 69–88. Westdeutscher Verlag, Opladen (1998)

Hendrickson, B., Goei, R.: Reciprocity and dating: Explaining the effects of favor and status on compliance with a date request. Communication Research 36, 585–608 (2009)

Keysar, B., Converse, B.A., Wang, J., Epley, N.: Reciprocity is not give and take: Asymmetric reciprocity to positive and negative acts. Psychological Science 19, 1280–1286 (2008)

Meegama, N., Blair, J.: The effects of telephone introductions on cooperation: An experimental comparison. In: Proceedings of the Survey Research Methods Section, pp. 394–397. American Statistical Association, Alexandria (1999)

Meier, G., Schneid, M., Stegemann, Y., Stiegler, A.: Steigerung der Ausschöpfungsquote von Telefonumfragen durch geschickte Einleitungstexte (Increase the response rate of telephone surveys with clever introductions). ZUMA-Nachrichten 57, 37–55 (2005)

Regan, D.T.: Effects of a favor and liking on compliance. Journal of Experimental Social Psychology 7, 627–639 (1971)

Skageby, J.: Gift-giving as a conceptual framework: framing social behavior in online networks. Journal of Information Technology 25, 170–177 (2010)

Stroud, N.J., Kenski, K.: From agenda setting to refusal setting: Survey nonresponse as a function of media coverage across the 2004 election cycle. Public Opinion Quarterly 71, 539–559 (2007)

Whatley, M.A., Webster, J.M., Smith, R.H., Rhodes, A.: The effect of a favor on public and private compliance: How internalized is the norm of reciprocity? Basic and Applied Social Psychology 2, 251–259 (1999)

Chapter 18
Preference for Mobile Interview Surveys? Interplay of Costs, Errors and Biases

Vasja Vehovar and Ana Slavec

18.1 Introduction

Information and communication technology has an important effect on the development of survey data collection methods. In recent years, computer assisted mobile telephone interviewing (mobile CATI) is becoming an increasingly frequent option, supplementing or replacing face-to-face and fixed telephone interviews (Gallup Europe 2009). This process was sped up by the increasing number of households without a fixed telephone but having at least one mobile telephone (Blumberg and Luke 2009). The decline in coverage – as well as in response rates for the fixed telephone CATI approach – in surveys of the general population led to an intensive search for alternatives.

Various studies have documented the declining fixed telephone coverage, which is usually accompanied with increased mobile telephone coverage, as well as with an increased share of mobile-only households in Europe (European Commission 2008a, Kuusela et al. 2008, Blyth 2008). According to these studies, the median country share of mobile-only households in Europe is around 22 – 24 %, with some countries already exceeding 60 %.

Similar problems are also faced in many other countries, particularly the United States, where considerable research has already been dedicated to this problem, including to the dual frame approach (Brick et al. 2010). Of course, we should be aware that mobile telephone surveys have numerous operational, legal, ethical, cost and statistical difficulties. More specifically, such surveys are more difficult and expensive to conduct, they require new procedures, have safety hazards, have the potential of reduced attention to the survey by the respondent, and very often significant amounts of screening are necessary.

Here we address only the statistical and cost-related problems of the optimal dual frame mixture, where we simultaneously consider data quality (measured by the mean squared error) and corresponding costs. The approach is based on work by Deming (1953), Kish (1965), Groves (1989), and Vehovar et al. (2010), which are among the rare studies that elaborate on survey costs, while most survey-research literature focuses on survey errors only. We may add that Groves (2006) explains this lack of the research on survey costs with the poor understanding of production processes, cost data and organizational culture.

The discussion is further narrowed to analytical observations about the behavior of the estimates under various cost ratio and mixing parameter conditions in dual frame designs. For this purpose we first give an overview of past research in the area of dual frame studies. Next, we develop the cost-error model, which is then used for simulations on real data.

18.2 Dual Frame Designs

Multiple frame surveys refer to two or more sampling frames that can cover the target population. They are useful for sampling rare or hard-to-reach populations and particularly for situations where a single frame does not provide enough coverage. Another reason for dual frame designs can be the cost savings over a single-frame design with comparable accuracy.

There have been three waves of multiple frame research literature. The first wave started in the 1960s and 1970s when researchers – particularly in the U.S. – combined less costly but incomplete frames (e.g. telephone surveys) and expensive but more complete frames (e.g. area-sample frames with personal interviews) to achieve optimal precision. Hartley (1962) was the first to examine multiple frame-sample estimation techniques systematically. He specifically focused on the units in the overlap, where various weighting combinations exist for the estimator. The units in the overlap have the same expected value of the target parameter, so they can be combined with various proportions. Hartley developed an approach to determine the mixing parameter, which optimally combines/weights the units in the overlap between different sampling frames. Hartley's theoretical work on estimators in multiple frame surveys was further elaborated by other researchers (e.g. Bankier 1986) who developed various improvements but also alternatives to the more general approach to this problem (Lepkowski and Groves 1986).

The second wave was in the 1980s and 1990s when scholars further extended Hartley's method and used it specifically to improve telephone survey methods. Groves and Lepkowski (1982) examined the problems of dual frame mixed-mode survey designs and argued that more exhaustive survey error specifications were needed, including other survey elements, such as design effects. Casady et al. (1981 in Groves and Lepkowski 1982) extended the combination of telephone and area frames to stratified multi-stage sample designs. A similar extension was done by Biemer (1983), who included costs and then empirically implemented the results to the Current Population Survey (CPS). Groves and Lepkowski (1982, Lepkowski and Groves 1984) further elaborated the error model to include response errors and biases. In all these studies, the dominant focus was on combining telephone and personal visit surveys to compensate for drawbacks of single-frame survey types, with the central problem of determining the best allocation of sample size between the two frames. Moreover, Lepkowski and Groves (1986) developed an error model for dual frame survey designs that include sampling and interviewer variance, bias, and cost. This model was applied to a large national survey where optimal allocation of the sample (affected by alternative levels in bias in both frames) was examined for two variables. An important issue about error and cost models for dual frame surveys was the fact that usually parameter values are not well known, especially for bias, which is problematic since optimal allocation is very sensitive to bias (at least for some variables).

In this period, Bankier (1986) made an important contribution with an alternative solution providing lower variances (compared to Hartley), which also simplified the computations, especially for the raking-ratio estimator. Traugott et al. (1987) also experimented with how dual frame design could reduce nonresponse in telephone surveys. In particular, they combined random digit dialing (better coverage) and purchased lists of telephone numbers (higher response rates and greater interviewer productivity). An allocation to minimize the total survey budget was suggested.

In the 1990s, Skinner (1991) and Skinner et al. (1996) continued Bankier's (1986) work by studying the efficiency of raking-ratio estimation for multiple frame surveys, based on Hartley's classic multiple frame theory, which he also compared with the Fuller estimator. The Hartley estimator was found to be less efficient but more generally useful than the alternative, which can be problematic when used in certain situations.

In the most recent, third wave of multiple frame studies, the main stimulus came from practical problems related to mobile telephones. This was also a basis for further theoretical discussion on the dual frame approach. Brick et al. (2006) analyzed nonresponse bias (focusing on inaccessibility bias) in dual frame samples of mobile and fixed telephone numbers. It was found out that mobile-only households are more likely to respond to mobile telephone surveys than households with both services. Since frequent mobile telephone users rarely answer their fixed telephones (if they have them), the fixed telephone sample suffers from inaccessibility bias. Different estimation schemes were evaluated, but none of them reduced nonresponse bias substantially. In order to adjust for inaccessibility nonresponse bias, two design approaches are possible. First, the mobile subsample can be weighted appropriately before being merged with the fixed telephone subsample, for example, according to how frequently the mobile telephone is turned on. Second, the mobile telephone sample could be screened before data collection to exclude those with both services and interview (in the mobile telephone sample) only households without a fixed telephone (Brick et al. 2006), which considerably reduces costs. The units in the overlap accessed by mobile telephone might thus be abandoned without the interview (Brick 2010), as they are in principle reachable also by fixed telephone. This may create a small specific additional coverage bias but several studies show that it is not substantial (Brick et al. 2006, Blumberg and Luke 2009) or that it can be eliminated with proper weighting (Keeter 2007).

Keeter (2007) also studied the impact of the growing mobile-only population. They tried to find deficiencies in national fixed telephone RDD surveys by analyzing four dual frame studies with different contents. It was found that despite significant differences in some variables, the additional inclusion of a mobile-only sample did not contribute much to improving the estimates. The non-coverage bias was removed only for young adults who are more likely to be mobile-only users.

Kennedy (2007) evaluated the effects of screening for telephone service in dual frame RDD surveys. Although not affecting coverage properties of a dual frame survey, screening could affect nonresponse and other sources of error. Nevertheless, if data are properly weighted (based on telephone service), the effect of screening on bias and variance is minimal. Further evaluation of mobile-only screening should consider the relative costs of fixed telephone and mobile interviews and screen-outs in the mobile sample. Since the results showed no difference in accuracy between dual frame (with or without screener) and fixed telephone sample-only designs adding a mobile frame did not seem worthwhile (at least when the mobile-only share is below 20 %, as in those studies).

Brick et al. (2010, forthcoming) examined non-sampling errors in dual frame telephone surveys, which were usually ignored in multiple frame literature. It was found that nonresponse and measurement error bias are substantial in specific conditions. A dual frame sampling and weighting method that emphasizes this importance is proposed to reduce this kind of bias. In fact, how the estimates from the two frames are combined is crucial. One of the main findings is that the choice of the compositing factor has only a small effect on variance but quite a significant effect on bias.

Some interesting empirical research was done by the Pew Research Centre analyzing data from eleven dual frame surveys conducted from 2009 to 2010. Various substantial topics were covered in these surveys: public policy issues, economic ratings, political views and attitudes, religious and social values, attitudes toward information-communication technologies, applications and social networking services. Primarily, bias was estimated by comparing the fixed telephone and dual frame samples. Both were weighted to the same standard set of demographic parameters and to telephone usage (fixed only, mobile only, dual users). Although the number of mobile-only households increased, the amount of non-coverage bias is still small for most items. In spite of being small, the differences are all in the predicted direction. This indicates that the estimates are biased, since most of the coverage problem is caused by the mobile-only households. The second comparison was made between fixed telephone samples and fixed telephone samples that additionally include mobile-only households. The differences are a bit lower than in the first comparison but have the same patterns (Leah et al. 2010).

To summarize, in general omitting the mobile-only segments – and also omitting the dual frame – has had only relatively small consequences for the estimates, at least in the situation of 20 % mobile-only units as in the U.S. and in many E.U. countries. However, despite being small, bias estimates can have important substantive implications for some specific segments (i.e., younger, less educated, lower income, non-white) and for some specific variables (political attitudes, technology use patterns, some social issues). In addition, when the share of the mobile-only segment increases – which is an unavoidable trend – the biases may become substantial (Kuusela et al. 2008).

Of course, when determining the overlap status, considerable ambiguity exists with respect to measuring the mobile/fixed status, because mere possession is very different from actual accessibility of a person via the fixed or mobile telephone device. However, as mentioned, in subsequent sections we focus only on the optimization of the mixture parameter, assuming that these segments are measured properly.

18.3 The Mixture Parameter: Analytic Solution

Here we study how to mix the mobile and fixed telephone subsamples in dual frame surveys. We assume that the two frames are independent and that they are sampled accordingly. We are specifically interested in the optimal value of the mixture parameter x which denotes the proportion of the fixed subsample, while (1 - x) is the proportion of the mobile subsample.

At the very beginning, we must explicitly point out that we refer not only to the mixture parameter that figures in the estimation stage of the dual frame, but also to the true mixture of the two samples. One may think that this value is decided based on the relation between the fixed/mobile telephone segments, in particular:

1. The share of the mobile-only population and the share of the functional mobile-only units (who not only posses mobile telephone devices, but are also accessible via them);
2. The relation between fixed and mobile telephone coverage;
3. The relation between mobile-only and fixed-only segments.

The survey practice (European Commission 2009, Leah et al. 2010) shows that when the mobile-only segment is below 50%, the share of the mobile telephone subsample is usually around 20 – 40 %.

Here we use an optimization approach considering accuracy (data quality) and corresponding costs to find the most suitable mixture of the two survey modes. Not many studies deal with survey costs; an example is the Link et al. (2007) study that compared costs of mobile and fixed surveys. There are even fewer studies investigating data quality and survey costs simultaneously. Usually, studies comparing different survey modes dealt only with response rates and differences in responses (Brick et al. 2007, Callegaro et al. 2007, Vehovar et al. 2004).

Our approach is based on optimization as used by Deming (1953), Kish (1965) and Groves (1989) where the product of error and cost is minimized. For simplicity, the discussion is further narrowed to estimating the proportion, which is also a dominant target variable in public polling. With some modification, this can also be easily extended to a general estimator of the mean.

Let the sample value estimate be p for the population parameter; P represents the combination of the two subsample values (p1 for fixed and p2 for mobile). Using the mixture parameter x we have:

$$p = x \cdot p_1 + (1-x) \cdot p_2 = x(p_1 - p_2) + p_2 \tag{1}$$

Survey errors have been variously considered in literature (Biemer and Lyberg 2003, Groves 1989, 2005, Groves et al. 2004, Lessler and Kalsbeek 1992, Weisberg 2005) but here it is restricted to expected differences between sample and true population estimates, namely the accuracy of a survey. These differences can be random or systematic, produced by constant effects over replication of a particular design (Groves et al. 2004). Thus, the total survey error has two components: sampling variance and bias. A standard measure is the mean square error (MSE) (Hansen et al. 1953, Deming 1950, Kish 1965):

$$MSE(p) = Var(p) + Bias(p)^2 \tag{2}$$

There are some practical problems associated with using the mean square error as a total survey error. First, each survey item has a different error that is calculated separately. Second, usually we do not have the true population value and have to find an estimate from auxiliary data, but we cannot be sure of its accuracy. Moreover, it is difficult to separate random error from biases and different bias components from each other. In fact, different designs have different errors: sampling, non-coverage, nonresponse, misspecification, measurement and processing errors. We assume here that the above bias includes all potential components (nonresponse, measurement, non-coverage etc.), despite the fact that these error sources are not easy to distinguish. Nevertheless, we are not interested in these components – although this is an important question and despite the fact that they may behave differently – but only in general bias-variance-cost tradeoffs in dual frame environments. Other approaches that focus on a few error sources only are much less complete (Vehovar et al. 2010).

In case of simple random sampling (SRS) the MSE is:

$$MSE(p) = Var(p) + (p - P)^2 \tag{3}$$

In case of small differences between p1 and p2 – which is usually the case – the variance can be calculated with a simplified formula:

$$Var(p) = \frac{P \cdot (1 - P)}{n} \tag{4}$$

In addition, the above simplification is justified even in case of certain differences among p1 and p2, since the elementary variance of the proportions in large samples, i.e. P(1-P), is robust to moderate changes in the percentage value P.

It should also be explicitly emphasized that the above simplification not only omits the two usual components of the variance for the composed estimate of p (1), but also the entire issue of the specific mixing parameter elaborated by Hartley (1962) and others. Thus, we will implicitly assume that this parameter has a ‘natural’, i.e., a self-weighted value, which reflects the true mixing parameter x, so any further elaboration about its optimization is not needed.

The population value (P) may either be found in official statistics or, usually, can only be modeled, estimated by experts, or simulated. In our work, we perform systematic simulations for P.

We define costs as the sum of the cost of fixed (c1) and mobile (c2) interviews, considering the mixture proportions. To simplify, we use the mobile to fixed cost ratio (C = c2/c1) instead of both cost parameters:

$$COST = nxc_1 + nc_2(1-x) = c_1(n \cdot x + \frac{c_2}{c_1} \cdot n \cdot (1-x)) \propto nx + C \cdot n \cdot (1-x) \tag{5}$$

As mentioned, minimizing the product of costs and errors is a standard approach, particularly in the classic stratification and cluster sampling literature (Cochran 1978, Hansen et al. 1953, Kish 1965). However, these studies are limited only to the product of the costs and sampling variances. A combination of mean square error (as a measure of accuracy) and costs is much more complicated and was theoretically rarely discussed (Lynn and Elliot 2000, Lyberg et al. 1998) and only few empirical applications exist. Besides our evaluations of alternative survey designs (Vehovar et al. 1998, 2001), there is also some work on double sampling nonresponse optimization (Harter et al. 2007, Deming 1953, Lynn et al. 2000) and some evaluations of alternative survey designs (Linacre and Trewin 1993, Westling 2008). In accordance with past studies, what should be minimized here is the cost per unit of accuracy that equals the product of costs and the mean square error:

$$\frac{COST}{accuracy} = \frac{COST}{\frac{1}{MSE}} = COST \cdot MSE \tag{6}$$

Thus, the function to be minimized is:

$$f(x) = MSE(p) \cdot COST = \left(\frac{1}{n}\left(p - p^2\right) + \left(p - P\right)^2\right) \cdot \left(x \cdot n \cdot \left(1 - C\right) + n \cdot C\right) \tag{7}$$

With some algebra, we get the equation of a surprisingly simple form — a polynomial of the second order:

$$\begin{aligned} f(x) = {} & x^3(p_1 - p_2)^2(1-C)(n-1) + \\ & + x^2(C(p_1 - p_2)^2(n-1) + (1-C)(p_1 - p_2)(2p_2(n-1) + 1 - 2nP)) + \\ & + x((1-C)((p_2 - p_2^2) + n(P - p_2)^2) + C(p_1 - p_2)(2p_2(n-1) + 1 - 2nP) + \\ & + C((p_2 - p_2^2) + n(P - p_2)^2) \end{aligned} \tag{8}$$

The optimal mixture parameter x is where the function f(x) has the global minimal value on the parameter's domain of definition which is the interval from

0 to 1. To check if there are any stationary points on this interval, we look for the roots of the first derivative of f(x):

$$\begin{aligned} f'(x) &= 3x^2(p_1 - p_2)^2(1-C)(n-1)'+ \\ &+ 2x(C(p_1 - p_2)^2(n-1) + (1-C)(p_1 - p_2)(2p_2(n-1)+1-2nP)) + \\ &+ ((1-C)((p_2 - p_2^2) + n(P - p_2)^2) + C(p_1 - p_2)(2p_2(n-1)+1-2nP)) = 0 \end{aligned} \tag{9}$$

Roots of this second-order equation are straightforward:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{10}$$

There are two solutions (or one double solution) that we retain only if they are in the domain of the definition of x (i.e. the interval 0 to 1). This occurs when the product of a and c is equal to or higher than -1/4. A special situation occurs when there are no real number solutions, when a is zero, which can happen in three situations: 1) sample size (n) is zero (practically not possible); 2) cost ratio C is 1 (fixed CATI and mobile CATI costs are equal) which is fully reasonable; 3) there is no difference in the subsample values (p1 = p2). In this case, we skip part a and have only a simple linear equation bx + c = 0 with only one root (x = -c/b).

Another special situation occurs when the discriminant is lower than zero, which means two complex solutions. In this case, we do not have solutions and use the limit values of the interval (i.e. 0 or 1). Then we compute the value of the function f(x) in the remaining two or one or none stationary points and in the limit values (0 and 1). The optimal value of parameter x is the lowest of these values, the global minimum of the function f(x) on the interval.

In any case, the final optimal solutions could be 0, 1 or something in between. It denotes the optimal share of interviews to be performed by fixed CATI method. If x = 1, all interviews should be performed with it, while if x = 0, the whole survey should be mobile CATI.

18.4 Empirical Evaluation and Simulation

The above approach was applied to data from a dual frame general attitudes survey (n = 1,000) in Slovenia with the mixture of 60 % fixed CATI and 40 % mobile CATI units. The data is from a 2009 Flash Eurobarometer survey (European Commission 2008b), which required dual frame design in eight European countries, Slovenia among them. We should add that the share of mobile-only units from large survey estimates (ESS, LFS) was 20 %, while the share of units effectively accessible only on mobile was assumed to be much higher. The mobile telephone coverage was 90 %, while fixed telephone coverage was 80 %.

The optimization procedure was carried out for ten different variables (labeled V1 to V10). The question subjects were socio-economic and political issues.

In order to correspond to the population both subsamples were weighted (using the raking method) for gender, age, years of education (three categories) and living area (metropolitan zone, other town/urban centre, rural zone).

On weighted data, the total sample value (p) was computed as a linear combination of the two weighted subsample values (p1, p2) for each variable (see equation 1). Alternatively, p1 and p2 were first calculated as separate stand-alone weighted estimates and then combined with the mixture parameter x.

Since there were no population data for the selected variables, different population values (P) were simulated to be used in the optimization procedure. In fact, this was the essential step in this study. The behavior of the mixture parameter will be observed under different assumptions for the true population value of P, which was not known at all.

We denote the assumed population values with the letter 'P' followed by a number (from 0 to 100) which means the importance (weight) of the fixed subsample in the true population simulation and denotes the true population value in the fixed and mobile telephone segments. There are five main simulations used determined by share – let us denote it with y – of the fixed telephone units:

- y = 0.5, so P50 is the average of estimates
- y = 1.0, so P100 is equal to the fixed CATI
- y = 0.0, so P0 is equal to the mobile CATI estimate
- y = 0.75, so P75 is between P100 and P50
- y = 0.25, so P25 is the average of P0 and P50

Weighted variable values for both subsamples and different estimate proportions (simulations of population value) are presented in Table 18.1. We can observe the usual pattern, where for the majority of variables the differences are small while for some (V7) they are considerable.

Table 18.1 Sample and simulated population values for all studied variables

Variable	*Sample value*		*Simulations of population values*				
	Fixed CATI	Mobile CATI	P100	P75	P50	P25	P0
V1	75.0	73.1	75.0	74.5	74.0	73.6	73.1
V2	65.1	64.8	65.1	65.0	65.0	64.9	64.8
V3	88.1	91.4	88.1	88.9	89.8	90.6	91.4
V4	78.5	75.7	78.5	77.8	77.1	76.4	75.7
V5	89.9	87.0	89.9	89.1	88.4	87.7	87.0
V6	16.9	16.6	16.9	16.8	16.7	16.7	16.6
V7	31.9	43.3	31.9	34.7	37.6	40.4	43.3
V8	62.0	64.2	62.0	62.5	63.1	63.6	64.2
V9	61.5	62.1	61.5	61.7	61.8	62.0	62.1
V10	52.5	51.5	52.5	52.2	52.0	51.7	51.5

Below is a graphic representation of the different sample and population values for V7, which shows the largest difference between the mobile and fixed telephone samples.

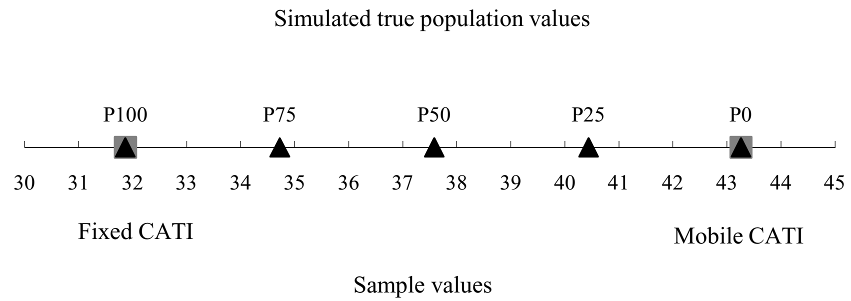


Fig. 18.1 Example of sample and population values (V7)

Different population value simulations and weighted sample values are used to compute the mean square error. For each variable we compute five errors (one for each simulation).

The other component needed in the f(x) function is costs. Costs of mobile CATI surveys are generally higher than for fixed CATI surveys but it is not simple to determine by how much. Thus, we used four different mobile to fixed cost ratio (c2:c1) options: 1.2, 1.6, 3 and 6.

Based on 50 mean square errors and 4 cost-ratio options we get 200 different situations in which to minimize the mean square error and costs product (6).

18.5 Analysis and Results

The results, of course, differ according to the conditions in which they were computed (variable, cost ratio, true population value estimation). In Table 18.2, we compare the optimal mixture parameters for all variables where the cost ratio is 1.2 and the population value is simulated as the average (P50).

Table 18.2 Optimal mixture parameter and product of MSE and costs for all variables (C = 1.2; P50)

Variable	Optimal mixture parameter (X) (optimal design proportion)	MSE*COST for the optimal mix
V1	0.56	0.21
V2	1.00	0.23
V3	0.50	0.10
V4	0.53	0.19
V5	0.52	0.11
V6	1.00	0.14
V7	0.50	0.26
V8	0.54	0.26
V9	1.00	0.24
V10	0.77	0.27

Under these conditions the results for all variables suggest that fixed telephone interviews should prevail in sample design. In some cases (V2, V6, V9), it would even be optimal to have no mobile telephone interviews at all. However, in some cases the suggested proportion is very close to 0.50, but never below. In other words, even if the true mixture of the population value was y = 0.5 (and we would expect x to be the same), the suggested optimal mixture parameter x was rarely x = 0.5 but was uniformly above this value and often even x = 1.0. The latter is of course a logical consequence of the simple fact that there were no differences between estimates, so the costs push the mixture towards the cheaper fixed telephone mode.

In Figure 18.2 the MSE*COST function is presented in relation to the possible design proportions of fixed telephone interviews (interval 0 to 1) for four representative variables. Most variables have a parabolic shape: in addition to V3, V7 and V10 (above) also V1, V4, V5 and V8. The rest have a descending curve: in addition to V6 (above) also V2 and V9. Among the parabolas, V1 is the most gradual and V7 is the steepest curve, while V6 is descending most gradually and V9 most steeply.

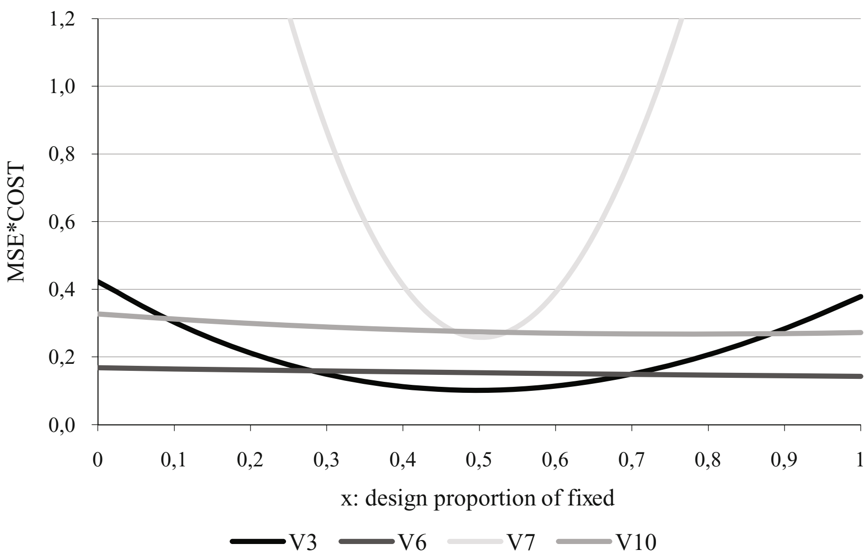


Fig. 18.2 MSE and product cost for mixture x (C = 1.2, P50) for variables V5, V10, V3 and V7

Let us additionally observe the effect of higher cost ratios on the optimal mixture parameter (x). In the last column (see Table 18.3), the value when the cost ratio is 1 is displayed for comparison. With higher costs, even higher proportions of fixed telephone units are found to be optimal.

Table 18.3 Optimal design proportion (X) for different cost ratios (P = P50)

Variable	*C* $=C_2/C_1$				$C_1=C_2$ *(C=1)*
	1.2	1.6	3	6	
V1	0.56	0.65	1.00	1.00	0.51
V2	1.00	1.00	1.00	1.00	0.55
V3	0.50	0.51	0.53	0.56	0.49
V4	0.53	0.57	0.66	1.00	0.51
V5	0.52	0.54	0.58	0.64	0.51
V6	1.00	1.00	1.00	1.00	0.38
V7	0.50	0.51	0.51	0.51	0.50
V8	0.54	0.61	1.00	1.00	0.49
V9	1.00	1.00	1.00	1.00	0.48
V10	0.77	1.00	1.00	1.00	0.50

Figure 18.3 shows the design proportions in the case of a 1:6 cost ratio. As expected, for most variables (where the estimates PF and PM do not differ, e.g. V6 and V10) it would be even more optimal to have only fixed telephone interviews, while for the remaining ones (where differences between PF and PM exist, e.g. V3 and V7) they shift faster and sharper to the population mixture y.

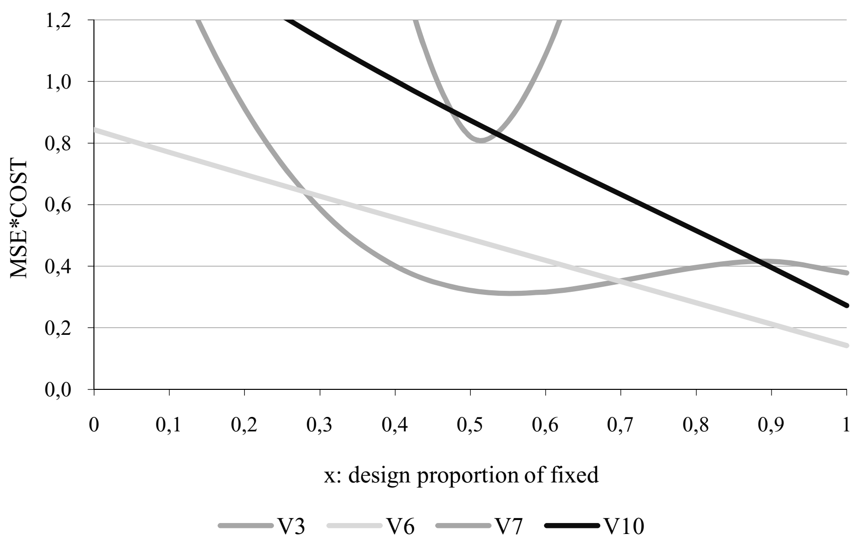


Fig. 18.3 Product of MSE and cost for different mixture proportions (C = 6)

Finally, let us also analyze the problem for other population value combinations in addition to the above value of y = 0.5 (P50).

Table 18.4 Optimal design proportion (x) for different P value simulations for two cost ratios

Variable	*Estimation proportion for population value simulation for C = 1.2*					*Estimation proportion for population value simulation for C = 6*				
	P100	P75	P50	P25	P0	P100	P75	P50	P25	P0
V1	1.00	0.81	0.56	0.31	0.06	1.00	1.00	1.00	1.00	1.00
V2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
V3	1.00	0.75	0.50	0.24	0.00	1.00	1.00	0.56	0.28	0.02
V4	1.00	0.78	0.53	0.28	0.00	1.00	1.00	1.00	0.43	0.13
V5	1.00	0.77	0.52	0.27	0.00	1.00	1.00	0.64	0.34	0.08
V6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
V7	1.00	0.75	0.50	0.25	0.00	1.00	0.77	0.51	0.26	0.01
V8	1.00	0.79	0.54	0.28	0.03	1.00	1.00	1.00	1.00	0.30
V9	1.00	1.00	1.00	1.00	0.75	1.00	1.00	1.00	1.00	1.00
V10	1.00	1.00	0.77	0.51	0.25	1.00	1.00	1.00	1.00	1.00

In general, we can confirm that the higher the "weight" for fixed CATI units in the true population value y, the higher the optimal design proportion x. However, x is never smaller than y – it is at least equal or larger, with many variables with x = 1. For instance, for V2 and V6 it is optimal to have only fixed telephone interviews regardless of the impact of the mobile telephone subsample on the true population value. In addition, V9 and V10 have a very high fixed proportion, which seems unusual but can be explained by the extremely low difference between estimates of the two subsamples (low MSE). Thus, costs determine the optimal solution. When the cost ratio is higher, this is even more radical.

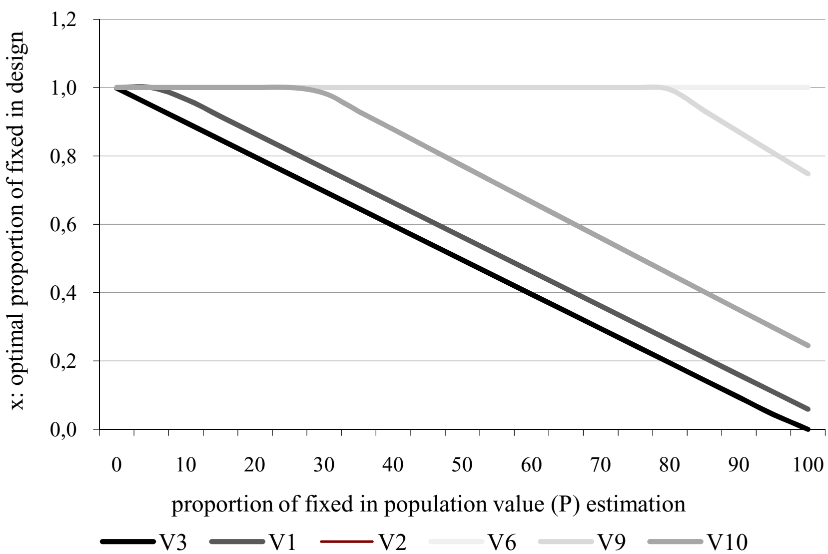


Fig. 18.4 Optimal proportion x for different population values y (C = 1.2)

Since the different variables give different results, a practical question is which variable should be the reference point when designing a survey. Straightforwardly, it should be the key variable in the survey; however, many surveys serve multiple purposes so the average or median value of several variables could be considered.

Here, variable V7, which has the highest difference between the fixed and mobile subsamples, was selected for a detailed analysis. First, we analyzed the relation between the proportion of fixed (design) and the MSE*COST function for a constant cost ratio (1.2) but for different simulations of the true population value, i.e. for different y values.

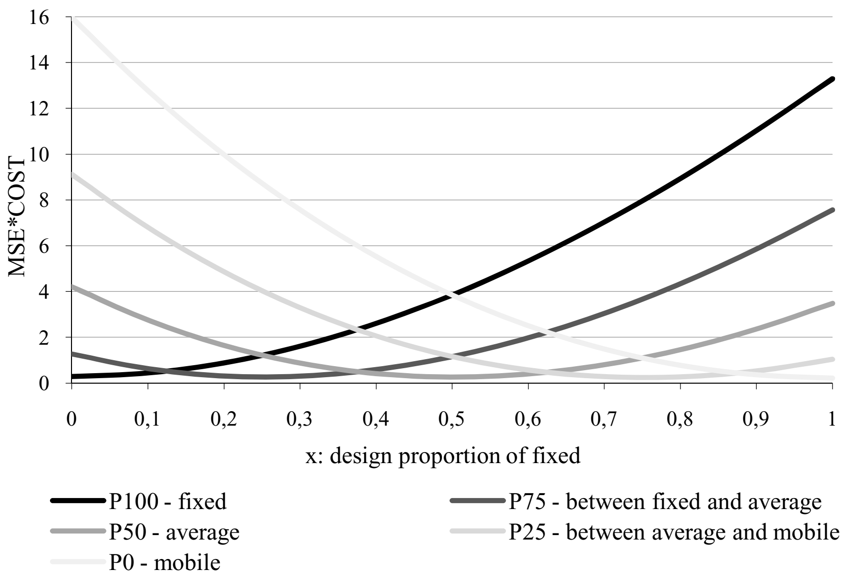


Fig. 18.5 MSE*Cost product in relation to mixture proportion (x) for variable V7 when C = 1.2

The optimal design proportion is where the function MSE*COST has the minimum value. Usually it tends to be closer to 1 (more fixed units) but if the true population estimate favors the mobile subsample, then the optimal design also has some more mobile units. As expected, x tends to have the optimum near the y value; however, the lines are relatively flat, and so deviations from the optimum have no radical consequences. This is also true in the case of higher cost ratios, although to a lesser extent (see Figure 18.6).

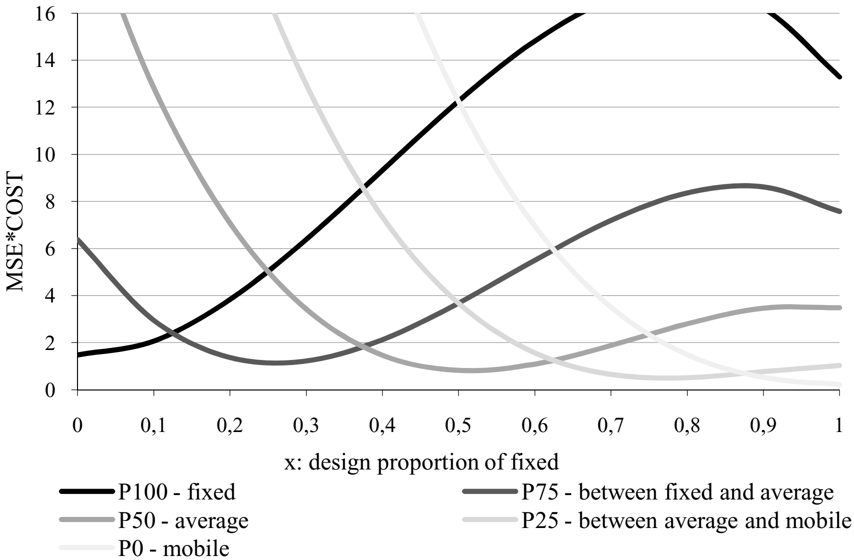


Fig. 18.6 MSE*Cost product in relation to mixture proportion (x) for variable V7 when C = 6

In Table 18.5 the MSE and cost product are shown for different combinations of the true population value simulation (y) and the mixture proportion (x).

Table 18.5 Minimal product of MSE and cost for variable 7 (cost ratio is 6)

		Weight of fixed CATI in estimation of true population value (y)				
		P100	P75	P50	P25	P0
x - proportion of fixed in design	0.0	79.89	45.58	21.08	6.37	1.47
	0.1	59.56	31.71	12.84	2.96	2.06
	0.2	43.03	20.98	7.09	1.37	3.82
	0.3	29.90	12.99	3.43	1.23	6.37
	0.4	19.77	7.35	1.47	2.13	9.31
	0.5	12.26	3.68	0.82	3.68	12.26
	0.6	6.97	1.58	1.09	5.50	14.81
	0.7	3.51	0.65	1.88	7.19	16.58
	0.8	1.50	0.52	2.80	8.36	17.18
	0.9	0.53	0.77	3.47	8.61	16.21
	1.0	0.22	1.03	3.48	7.57	13.29

We see that there is a connection between the weight of the fixed CATI subsample and the design proportion of fixed CATI. The product of MSE and costs is minimal where the two are similar. Different cost ratios do not seem to have a substantial effect on this relation (see Figure 18.7).

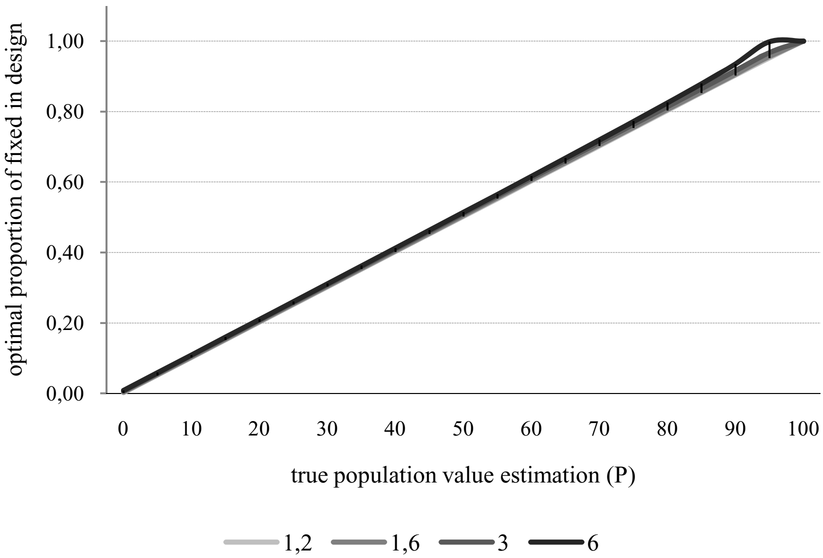


Fig. 18.7 Optimal proportion of fixed in design in relation to different P value estimations

However, this holds true only for the point that estimates x at minimum MSE*COST, while from previous figures we can see that the optimization curves are in fact relatively flat and that in case of no or little differences between PF and PM estimates the mixture parameter moves to x = 1, i.e. entirely to the cheaper mode.

18.6 Conclusion

We analyzed the problem of the mixture parameter for dual frame designs that combine fixed and mobile telephone subsamples. The intention was to simulate the unknown population value P as a combination of the mobile and fixed subsample estimates (with y being the share/weight of the fixed subsample component). It was found that if the true population value is in the middle (y = 0.5), it is optimal to take 50 % of each subsample to form the population estimate. Of course, if there were no differences among them, the fixed telephone should be used fully, simply because it is cheaper.

The parameters PF, PM and P are population values in the fixed telephone subsample, in the mobile telephone subsample, and in the total population, respectively. The relations among them are fixed and defined by an unknown parameter y. We just have to adopt the research design (within a given cost structure) to these relations by determining the mixture, i.e. the share of the fixed telephone subsample (x) in the total sample size.

First, the optimal mixture parameter was developed analytically, while in the empirical example we took real data from the Flash Eurobarometer 2008 survey for Slovenia to estimate PF and PM. Then, various true values of P were simulated by varying the combination of PF and PM, i.e. by varying the proportion of the fixed telephone estimate (y). For each given y the optimal value of the mixture parameter x of the dual frame was looked for by optimizing the accuracy per unit of cost (MSE*Costs).

The analysis shows that we might be calling too many mobile telephone numbers. The reasons for this are as follows:

- The majority of real variables showed almost no differences between PF and PM, so only costs matter for them, which means that only the fixed telephone sample is needed;
- When larger differences were observed between PF and PM (e.g. V7), the simulations showed that the optimality curves are relatively flat, so considerable deviations are tolerable. If the cost ratios were moderate (which is usually the case according to empirical evidence, when mobile telephone costs are less than 2 – 3 times higher) the additional gap might be 10 %. This means that instead of the optimal value x = 0.7 the share x = 0.8 would be also acceptable without much loss in accuracy per unit of cost;
- Very often the optimal value of the mixture parameter x was larger than y and thus closer to the fixed telephone extreme, while it was of course never below y;
- There are additional advantages of higher shares of mobile telephone samples, which are related to the larger samples that could be afforded in such an option. This may have other advantages not included in the optimization quantity MSE*Costs.

In addition, it should be emphasized that the above approach is very specific and relatively tricky, as it does not take into account the population proportions of the fixed/mobile telephone segments. These data were not incorporated into calculations, as we assumed that they are already incorporated in the parameter y, i.e. in the relations among PF, PM and P. It is true however, that with a 15 % mobile-only population and with almost full coverage of both frames (as is the situation in Slovenia) the value P would be much closer to PF than to PM.

It should also be added that, in practice, for calculating p (the estimate of P) some population values or at least their estimates for various fixed/mobile segments (possession as well as accessibility) would be definitely needed. We did not discuss this problem at all. If we calculate the posterior value of y (from actual estimates for P, PF, PM) we can see that it is always above y > 0.85, so the estimates have the 0.85 weight on PF (or its estimate).

Based on all aspects discussed above, the actual share we observe in the survey (x = 0.4; i.e. 400 units in mobile telephone sample and 600 units in fixed telephone sample) seems to be too high.

Thus, it is usually optimal to have many fewer mobile telephone interviews in the sample than the share of mobile-only users. However, in some cases some target variables are very sensitive to the bias component and the higher cost ratio reinforces this impact, so we need to be careful when lowering the parameter x. Let us repeat that parameter x is the true share of the fixed telephone subsample and not the weight received in the estimation procedure. Therefore, we used specific assumptions and self-weighted samples, so that the issue of determining the mixture parameter, which combines the two estimates, was avoided. It is possible, although somewhat unlikely (because the assumed differences for the two estimates are small or negligible), that the optimization of this estimation parameter would further improve the estimate. However, this exceeds the focus of our research.

In future, many of the barriers preventing wider usage of mobile telephone technology in survey data collection are expected to be eliminated, or at least minimized. With lower per minute charges, simpler pricing models, fewer legal restrictions and technological problems, more standardization, and more congruent coverage rates it will be less complex and expensive to use mobile telephones for data collection. Therefore, the dilemmas about mixture parameters will become much less important.

References

Bankier, M.D.: Estimators based on several stratified samples with applications to multiple frame surveys. Journal of the American Statistical Association (1986), http://www.jstor.org/stable/2289086 (accessed October 26, 2009)

Biemer, P., Lyberg, L.E.: Introduction to survey quality. Wiley, Hoboken (2003)

Biemer, P.P.: Optimal Design of Dual Frame Survey Results of a Simulation Study. In: Proceedings of the American Statistical Association, Section of Survey Methods Research, pp. 630–635 (1983)

Blumberg, S.J., Luke, J.V.: Wireless substitution: Early release of estimates from the National Health Interview Survey, July-December 2008. National Center for Health Statistics (2009), http://www.cdc.gov/nchs/nhis.htm (accessed July 29, 2010)

Blyth, B.: The implications of variation in national data-collection mode access and rates of access change: a European overview (2008), http://www.moaweb.nl/bibliotheek/materiaal-bijeenkomsten-1/2009/de-aanstaande-revolutie-binnen-het-marktonderzoek-22-jan-2009 (accessed August 26, 2010)

Brick, J.M., Dipko, S., Presser, S., Tucker, C., Yangyang, Y.: Nonresponse bias in a Dual Frame Sample of Cell and Landline Numbers. Public Opinion Quarterly 70(5) (2006), http://poq.oxfordjournals.org/cgi/content/abstract/70/5/780 (accessed October 26, 2009)

Brick, J.M., Sherman, W.E., Lee, S.: Sampling telephone numbers and adults, interview length, and weighting in the California health interview survey cell phone pilot study. Public Opinion Quarterly 71(5), 793–813 (2007)

Brick, M.J., Ismael, F., Cervantes, I.F., Lee, S., Norman, G.: Nonsampling Errors in Dual Frame Telephone Surveys. Unpublished Paper (2010)

Callegaro, M., Steeh, C., Buskirk, T.D., Vehovar, V., Kuusela, V., Piekarski, L.: Fitting disposition codes to mobile phone surveys: experiences from studies in Finland, Slovenia and the USA. Journal of the Royal Statistical Society: Series A (Statistics in Society) 170(3), 647–670 (2007)

Cochran, W.G.: Sampling techniques, vol. 3. Wiley, New York (1978)
Deming, W.E.: Some theory of sampling. John Wiley & Sons, New York (1950)
Deming, W.E.: On a probability mechanism to attain an economic balance between the resultant error of response and the bias of nonresponse. Journal of the American Statistical Association 48(264), 743–772 (1953)
European Commission. Special Eurobarometer 293. E-communication Household Survey. TNS Opinion and social (2008a), http://ec.europa.eu/public_opinion/archives/ebs/ebs_293_full_en.pdf. (accessed July 29, 2010)
European Commission. Flash Eurobarometer 251. Public attitudes and perceptions in the Euro area. Gallup Europe, Brussels [Producer]. GESIS, Cologne [Publisher]. ZA4743, data set version 2.0.0. Description (2008b), http://info1.gesis.org/dbksearch13/sdesc2.asp?no=4743&db=e&doi=10.4232/1.10007 (accessed July 29, 2010)
European Commission. Flash Eurobarometer 25. Public attitudes and perceptions in the Euro area. Analytical Report. Gallup Europe, Brussels (2009), http://ec.europa.eu/public_opinion/flash/fl_251_en.pdf (accessed July 29, 2010)
Gallup Europe: Alec M. Gallup Future of Survey Research Forum: Incorporating Mobile Phones in Social and Policy-oriented Surveys, Brussels (October 14, 2009)
Groves, R.M., Lepkowski, J.M.: Alternative Dual Frame Mixed Mode Survey Designs. In: Proceedings of the Survey Research Methods Section. American Statistical Association (1982), https://www.amstat.org/sections/SRMS/Proceedings/papers/1982_028.pdf (accessed October 26, 2009)
Groves, R.M.: Survey errors and survey costs. Wiley, Hoboken (1989)
Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R.: Survey methodology. Wiley, Hoboken (2004)
Groves, R.M.: Total survey error: Past, present, and future. Paper Presented at the Total Survey Error Workshop, Washington, DC (2005)
Groves, R.M.: Survey budgets, cost models, and responsive survey designs. Paper presented at the Survey Cost Workshop, Washington, DC (2006)
Hansen, M.H., Hurwitz, W.N., Madow, W.G.: Sample survey methods and theory. Wiley, New York (1953)
Harter, R., Mach, T., Wolken, J., Chapline, J.: Determining subsampling rates for nonrespondents. Paper presented at the Third International Conference on Establishment Surveys, Montréal, Canada (2007)
Hartley, H.O.: Multiple Frame Surveys. In: Proceedings of the Social Statistics Section. American Statistical Association (1962)
Keeter, S.: What's Missing from National Landline RDD Surveys? The Impact of Growing Cell-Only Population. Public Opinion Quarterly 71(5) (2007), http://poq.oxfordjournals.org/cgi/content/abstract/71/5/772 (accessed October 26, 2009)
Kennedy, C.: Evaluating the Effects of Screening for Telephone Service in Dual Frame RDD Surveys. Public Opinion Quarterly 71(5) (2007), http://poq.oxfordjournals.org/cgi/content/abstract/71/5/750 (accessed October 26, 2009)
Kish, L.: Survey sampling. Wiley, New York (1965)
Kuusela, V., Callegaro, M., Vehovar, V.: The Influence of Mobile Telephones on Telephone Surveys. In: Advances in Survey Methodology. Wiley (2008)

Leah, C., Keeter, S., Purcell, K., Smith, A.: Assessing the Cell Phone Challenge to Survey Research. The Pew Research Center for the People & the Press (2010), http://pewresearch.org/pubs/1601/assessing-cell-phone-challenge-in-public-opinion-surveys (accessed May 24, 2010)

Lepkowski, J.M., Groves, R.M.: The Impact of Bias on Dual Frame Survey Design. In: AMSE model - Proceedings of the Section on Survey Research. American Statistical Association (1984), http://www.amstat.org/sections/SRMS/Proceedings/papers/1984_051.pdf (accessed October 26, 2009)

Lepkowski, J.M., Groves, R.M.: A Mean Squarred Error Model for Dual Frame, Mixed Mode Survey Design. Journal of the American Statistical Association 396(81) (1986), http://www.jstor.org/stable/2289062 (accessed October 26, 2009)

Lessler, J.T., Kalsbeek, W.D.: Nonsampling error in surveys. Wiley, New York (1992)

Linacre, S.J., Trewin, D.J.: Total survey design: Application to a collection of the construction industry. Journal of Official Statistics 9(3), 611–621 (1993)

Link, M.W., Battaglia, M.P., Frankel, M.R., Osborn, L., Mokdad, A.H.: Reaching the U.S. cell phone generation: Comparison of cell phone survey results with an ongoing landline telephone survey. Public Opinion Quarterly 71(5), 814–839 (2007)

Lynn, P., Elliot, D.: The British Crime Survey: A review of methodology. National Centre for Social Research, London (2000)

Lyberg, L.E., Japec, L., Biemer, P.: Quality improvement in surveys: A process perspective. Paper presented at the Survey Research Methods Section (1998)

Skinner, C.J.: On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys. Journal of the American Statistical Association 86(415) (1991), http://www.jstor.org/stable/2290412 (accessed October 26, 2009)

Skinner, C.J., Rao, J.N.K.: Estimation in Dual Frame Surveys with Complex Designs. Journal of the American Statistical Association 91(433) (1996), http://www.jstor.org/pss/2291414 (accessed October 26, 2009)

Traugott, M.W., Groves, R.M., Lepkowski, J.M.: Dual Frame Designs to Reduce Nonresponse. Public Opinion Quarterly 51, 522–539 (1987)

Vehovar, V., Lozar Manfreda, K.: How many mailings are enough? In: Koch, A., Porst, R. (eds.) Nonresponse in Survey Research, pp. 139–149. Zentrum für Umfragen, Methoden und Analysen, Mannheim (1998)

Vehovar, V., Lozar Manfreda, K., Batagelj, Z.: Sensitivity of electronic commerce measurement to the survey instrument. International Journal of Electronic Commerce 6(1), 31–52 (2001)

Vehovar, V., Belak, E., Batagelj, Z., Čikić, S.: Mobile phone surveys: The Slovenian case study. Metodološki Zvezki (Advances in Methodology and Statistics) 1(1), 1–19 (2004)

Vehovar, V., Berzelak, N., Lozar Manfreda, K.: Mobile Phones in an Environment of Competing Survey Modes: Applying Metric for Evaluation of Costs and Errors. Social Sciences Computer Review 28, 303 (2010)

Weisberg, H.F.: The total survey error approach: A guide to the new science of survey research. University of Chicago Press (2005)

Westling, S.: A simulation approach to evaluate the cost efficiency of nonresponse follow-ups. Working Paper series, Handelshögskolan vid Örebro universitet 2008(11) (2008), http://www.oru.se/oruupload/Institutioner/Ekonomi%20statistik%20och%20informatik/Forskning/wp2008-11.pdf (accessed October 26, 2009)

Chapter 19
Fieldwork Monitoring in Telephone Surveys

Patrick Schmich and Franziska Jentsch

19.1 Introduction

In Germany, the telephone as a mode is still most commonly used for collecting quantitative data in empirical market and social research (e.g. www.adm-ev.de). Although a number of problems – such as declining response rates (e.g. Curtin et al. 2005) and a constantly changing telecommunications market – make it necessary to try out new access routes (e.g. online surveys), and although these are increasingly being used, there is a lack of methods for generating samples for representative online surveys (Faas 2003, Couper and Coutts 2006, AAPOR 2010). While online surveys are a quick and inexpensive instrument (for example, for questioning the members of an access panel), telephone and randomly generated samples for telephone surveys still play an important role for low-cost, population-based representative surveys.

CATI management comprises many different areas, such as drawing a sample that represents the respective population, constant monitoring of fieldwork and the recruitment and supervision of interviewers. Good CATI management thus forms the foundation for valid data.

There are currently few analyses based on methodological data collected during fieldwork in telephone studies. This is partly due to the fact that only a limited amount of data is made available to the users. For example, practice-oriented analyses of large samples are often relatively old (Glemser 2002), since those who commission such studies are primarily interested in the outcome rather than the method used to obtain it. In addition, it is necessary to constantly evaluate the knowledge gained to date because of the highly dynamic developments in the field of telecommunications.

The aim of this study is to, on the one hand, present data and analyses and, on the other hand, to update and review existing knowledge. For this purpose, we will describe selected methodological parameters from the survey "German Health Update" (GEDA) 2010. GEDA is conducted in regular "waves" by the Robert Koch Institute (RKI) on behalf of the German Federal Ministry of Health (BMG).

This chapter presents the framework conditions and broad parameters of fieldwork management on the basis of the methods developed by Gabler and Häder (Häder and Gabler 1998) and compares these with respective results from the literature. In addition, the study aims to put forward ideas for further analyses and practical instructions for conducting similarly designed studies.

19.2 Key Data on the Study

The GEDA 2010 study complements the examination surveys – “German Health Interview and Examination Survey for Adults (DEGS)” and the “National Health Interview and Examination Survey for Children and Adolescents (KiGGS)” – which are conducted by the RKI within the context of health monitoring. Data from the regularly repeated GEDA study are used to continuously observe developments related to certain diseases as well as the population's health- and risk-related behavior. This information forms the basis for quickly identifying health trends in the population or in population groups for both health reporting and government health policy. The size of the sample also allows regionalized or deeply structured context analyses (Kurth et al. 2009).

The study ”German Health Update” (GEDA 2010) is a representative survey of the resident German-speaking adult population living in private households who have a fixed telephone line. The survey is conducted as a computer-assisted telephone interview (CATI). The telephone-number sample is created using the Gabler-Häder Design (Häder and Gabler 1999); it is based on phone numbers taken from public telephone books. In order to include in the generated number pool also people whose numbers cannot be found in any of these directories, the Waksberg method (Waksberg 1978), which is primarily used for telephone interviews in the USA, was adapted to the situation in Germany. This makes it possible to obtain a randomly generated number pool consisting of both published and non-published phone numbers. In order to give each element of the population the same theoretical likelihood of being interviewed, an additional selection of target persons is carried out at the household level, using the “last-birthday method”. Attempts are also made to avoid distortions caused, for example, by the different accessibility of individual household members.

During GEDA 2010's 10-month field period, 138 interviewers conducted a total of 22,076 interviews in 353 shifts (each shift lasting about four hours) on 223 days based on a total of 215,513 gross telephone numbers (see Table 19.1).

The data were collected in the Robert Koch Institute's telephone laboratory, which has a total of 40 interviewer and three supervisor workstations. Calls were made six days a week (Monday to Saturday) in a total of ten shifts (see section 19.3.4). The methodology used in GEDA 2010 builds on experience gathered in previous telephone health surveys (Kohler et al. 2005, RKI 2011).

Table 19.1 Key data from the GEDA 2010 field

GEDA 2010	
Interviewer	
Number	138 (82 women and 56 men)
Average age	41 (between 20 and 70)
Subjects (un-weighted figures)	
Number	22,076 (56.6 % women and 43.4 % men)
Average age	48 (between 18 and 99)
Questionnaire	
Adjusted mean duration of interview[1]	31 minutes
Quantitative parameters	
Willingness to take part in telephone survey again	82.8 %
Willingness to take part in survey again in writing	36.3 %
Stated email address	19.5 %
Number of refusals at the target-person level	12,130
Number of nonresponder questionnaires[2]	1,809
Number of phone numbers called	215,513
Total number of calls made	908,883
Mean no. of calls per interview (overall sample)	41.2

19.3 Sample Management

All efforts at sample management are based on the need to reflect the population under study as precisely as possible. To achieve this, the aim was to ensure that the likelihood of being selected was identical for all individuals in the sample. At the same time - since this representative study is carried out for several months – it is necessary that each month a certain number of interviews are conducted in order to offset seasonal effects.

Essentially, two management mechanisms can be named for fieldwork management. One is the call and call-back management, the other is concerned with when generated phone numbers are imported into the system.

19.3.1 Call and Call-Back Management

When designing a system of call and call-back management, it is important to ensure that potential subjects are not unnecessarily burdened with too many calls, especially bearing in mind the general decline in willingness to participate in interviews (De Leeuw and De Heer 2002). On the other hand, the chances of reaching subjects at a favorable moment and encouraging them to take part can be increased by repeatedly trying to call at different times. The great challenge is finding the right balance between these two restrictions.

1 After deduction of time spent on introductory remarks and goodbyes.

2 The "nonresponder" questionnaire contains eleven questions which aim to shed some light on people who are unwilling to take part in the survey. Analyses have revealed that this specific group differs only marginally (if at all) in terms of their socio-demographic characteristics from those who are willing to take part.

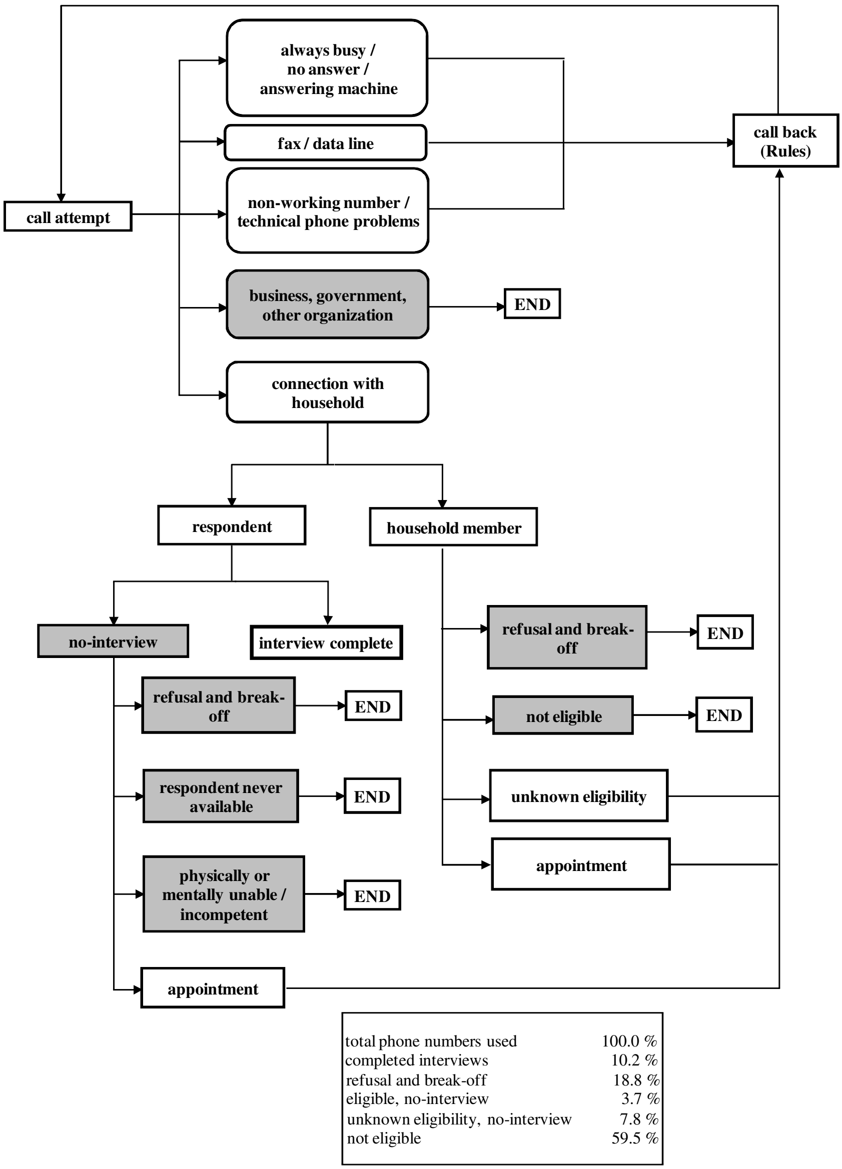


Fig. 19.1 Call management and outcome rates in GEDA 2010 based on the standard definitions of AAPOR (2010)[3]

---

[3] Total phone numbers used 100.0%; completed interviews 10.2%; refusal and break-off 18.8%; eligible, no-interview 3.7% unknown eligibility, no-interview 7.8%; not eligible 59.5 %.

Call Outcomes and Call Codes

Every call and call-back management system is based on a method for tagging the outcome of calls with corresponding call codes (see Figure 19.1). These call codes and the rules associated with them govern, when a phone number comes up again for redialing. When designing the call-back rules, the fieldwork duration of about ten months needs to be taken into account, ensuring that the numbers appear for call-back on different days of the week and at different times of day, in order to increase the likelihood of making contact. In addition, attempts were made to define these rules in a way that each data record had an approximately equal chance of being called often enough in the course of the field phase.

The rules in GEDA were defined in advance in such a way that every call always remains in the call-back pool for about a month[4]. Another task of call-back management is to lay down the priorities governing how the individual data records are to be re-submitted depending on the outcome of the call. Thereby, call-backs are given precedence over 'fresh' data records, i.e. telephone numbers that have not been called yet.

The call-back rules are laid down before the fieldwork begins and cannot be changed in the course of the fieldwork. They are subject to automated management.

Figure 19.1 shows the field management system and summarized call-up results. In GEDA 2010, a total of approximately 60 % of all the telephone numbers that were used belong to the "not eligible" category. In other words about 21,000 interviews had to be obtained from the remaining 40 %.

Number of Calls

In addition to the call-back rules, the number of contact attempts to yet unproductive each telephone number is also defined prior to fieldwork. Each call that has not been assigned a final code (e.g. "interview broken off", "interview refused", "business line", etc.) is re-contacted several times. The maximum number of call attempts was set at 15 (Kohler et al. 2005), recent sources report an optimal number of 12 attempts (Hansen 2008). Ultimately, this upper limit always represents a compromise between the aim of maximizing the potential of the sample and external parameters such as time, the financial framework and acceptance by the respondents.

Table 19.2 shows a breakdown of contact attempts and the resulting interviews. It becomes clear that about a third of the interviews are completed after the second call, and that nearly 90 % of the subjects agree to an interview after another eight calls.

Comparing age and education distributions by the number of calls shows whether or not it is worthwhile to continue calling after the tenth attempt.

The key factor influencing a decision on the optimum number of calls is whether calling more than ten times leads to interviews with subjects who

[4] Appointments with subjects can cause deviations as a result of which a data record may remain in the active sample for a longer period.

according to the literature are underrepresented in telephone surveys (Hartmann and Schimpl-Neimanns 1992). For example, phone interviews with people with a lower level of education are generally regarded as difficult. An analysis of the age distributions in Table 19.2 shows one thing in particular: approximately 15 % of respondents below the age of 30 are reached and motivated for an interview after more than 10 call attempts. By contrast, people aged 65 years and over are easier to reach; almost 30 % are already interviewed at the first call and only about 5 % after more than 10 call attempts.

An analysis of the distribution of educational groups as a function of the number of call attempts reveals that there are hardly any differences between educational groups. According to this, the age distribution should be the main criterion when deciding on whether more than 10 calls are worthwhile or efficient. Depending on a survey's focus and its target group, investing in further contact attempts can have a noticeable effect. However, it also becomes clear that making more than 16 calls cannot be justified; it increases neither the overall number of subjects nor the coverage of certain age or educational groups.

Table 19.2 Interviews by contact attempts – broken down by education and age in percent

	Interviews		*CASMIN Educational groups*[5]			*Age*			
calls	n	Cum. [%]	Low	Inter-mediate	High	- 29	30-44	45-64	65+
1	4,151	18.8	21.5	18.0	17.8	14.2	15.2	18.9	28.2
2	3,849	36.2	18.4	17.1	17.1	15.7	16.2	16.7	22.2
3	2,920	49.5	13.1	13.2	13.4	12.1	13.3	13.5	13.6
4	2,178	59.3	9.6	9.7	10.5	9.4	10.2	10.6	8.5
5	1,661	66.9	7.3	7.6	7.5	7.7	7.8	7.9	6.2
6 to 10	5,013	89.6	20.6	23.5	23.2	26.3	25.3	22.6	16.0
11 to 15	1,827	97.8	7.3	8.8	8.1	11.4	9.4	7.9	4.5
16 +	477	100.0	2.1	2.1	2.3	3.0	2.7	2.0	0.9
Total [%]			100	100	100	100	100	100	100

Soft Refusals

Any measure aimed at maximizing the response rate should always be repeatedly challenged and checked. A particularly critical and sensitive area is how to handle so-called "soft refusals", which Sudman (1998) defines as follows: "A soft refusal is one that says 'I'm too busy right now', or 'This isn't a good time'" (p. 69). If, as in this case, no real contact can be made with the person and they cannot be told the purpose of the call because, for example, the conversation is already ended during the interviewer's introductory remarks, a renewed attempt to contact them is made[6] – hoping on reaching a different person or the situation being different. This approach follows the guidelines of the Council of German Market and Social Research (Rat der Deutschen Markt- und Sozialforschung e. V. 2011).

[5] E.g. Braun et al. (1999).
[6] In the GEDA study a repeat call attempt was scheduled for a week later.

By contrast, further attempts to contact a person who has clearly expressed the desire not to be called again (hard refusal) are prevented by the rules of call-back management.

The interesting question now is the composition of the group of people[7] who are willing to take part in the survey at a later date after a call has been coded as "soft refusal". The question here is the extent to which a subsequent call is justified in this sensitive situation.

Table 19.3 Interviews after a soft refusal compared to the adjusted total number of interviews

	Interviews without soft refusal		*Interviews after soft refusal*	
	n	[%]	n	[%]
Gender distribution				
Male	8,835	43.3	741	44.8
Female	11,586	56.7	914	55.2
Age groups				
up to 29	3,565	17.5	269	16.2
30-44	5,725	28.0	370	22.3
45-64	7,358	36.0	622	37.5
65 +	3,754	18.4	393	23.7
CASMIN educational groups				
lower education	4,562	22.3	507	30.6
Intermediate education	10,196	49.9	800	48.3
Higher education	5,591	27.4	339	20.5

It is a fact that about eight percent of all interviews are conducted after a "soft refusal". Although previous analyses (see Table 19.3) have shown that older people often already agree to an interview during the first call, several respondents within this group are evidently quite skeptical about telephone interviews and tend to be more difficult to persuade to take part in such surveys (Schnauber and Daschmann 2008). In these cases, renewed contact after a soft refusal may well have the desired result. A binary-logistic regression analysis incorporating age and education shows that particularly people above the age of 64 or with a background of low education can be successfully motivated to participate by a repeat call.

19.3.2 When to Import Numbers

There are basically two ways of managing a sample: gross management and net management (von der Heyde 2002). In the case of gross management, every telephone number brought into the system is called according to the previously defined rules until there is a final result. In net management, the cell distribution is predetermined quantitatively and takes its orientation from the marginal distributions of the population under study. Gross management was chosen for

[7] It is impossible to check at this point whether the person interviewed is ultimately the same person whose previous reaction caused the "soft refusal" result.

GEDA 2010, since there is no prior information on the subjects to be included, and no prior screening was carried out. In order to maximize the potential, the amount of telephone numbers (gross sample) needed to reach the target sample size (about 21,000 interviews) was determined before the study. A factor of one interview to ten telephone numbers was calculated (1:10), based on the experience in previous studies. This factor was confirmed almost exactly after the end of the study at 9.8 (215,513 phone numbers to 22,076 interviews).

The timing of importing numbers is important for maximizing both the results from the sample and the inclusion of people who are difficult to reach. The principle is that a number is most likely to pass through the pre-determined number of calls if it is imported at an early stage of fieldwork. But if a telephone number is imported very late, i.e. shortly before the end of the fieldwork period, this also reduces the likelihood of it being dialed as often as required by the call-back rules. This would mean that not every target person would have the same chance of getting into the sample. Several possibilities are conceivable for determining the best time for importing numbers.

1. Importing a quantity of telephone numbers at the beginning of fieldwork that has previously been calculated to be necessary to achieve the sample; this quantity must be determined as accurately as possible. In their diagram (see Figure 19.2), Buckley et al. (1998) show the composition of the sample and how it changes when all the numbers are imported at the beginning of fieldwork. Here, the disadvantage is the large number of "completed" interviews at the beginning of the fieldwork and an ever-diminishing margin of possible interviews towards the end. This can cause fatigue among the interviewers and also produce seasonal effects. At the same time, the active group of interviewers would have to be continuously reduced, making it impossible in practice to make full and effective use of the workstations' capacity.

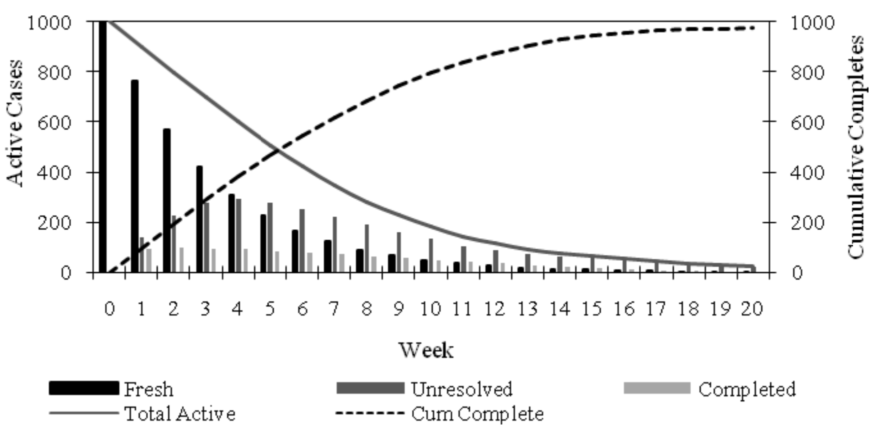


Fig. 19.2 Case mix obtained by releasing all cases at the start (according to Buckley et al 1998)

2. Another possibility is to import the phone numbers according to pre-defined intervals[8]. The size of the individual batches can be based on the actual number of completed interviews required. One possibility would be 'close meshed' monthly import. The individual batches would have to be slightly larger at the beginning of the fieldwork in order to be able to use up all the remaining numbers in the same way towards the end of the ten-month field period. Buckley et al. (1998) describe this approach in an exemplary way as follows:

- Initial phase - high proportion of “fresh numbers”
- Middle phase - mixture of non-contacted numbers and call-backs
- Final phase - the proportion of call-backs increases, the proportion of fresh numbers decreases

The advantage of the interval system lies in better management possibilities. If one batch cannot be called as expected, the following one might be correspondingly smaller; this could minimize the above-mentioned problems. Interval-based management is also regarded as practicable because of the relatively long fieldwork period. Its decisive advantage is its greater flexibility and the constant utilization of the available interviewer workplace capacity.

Figure 19.3 demonstrates an idealized example of this approach; its advantage lies in the constancy of the number of interviews completed over time, so that it comes very close to the preferred approach.

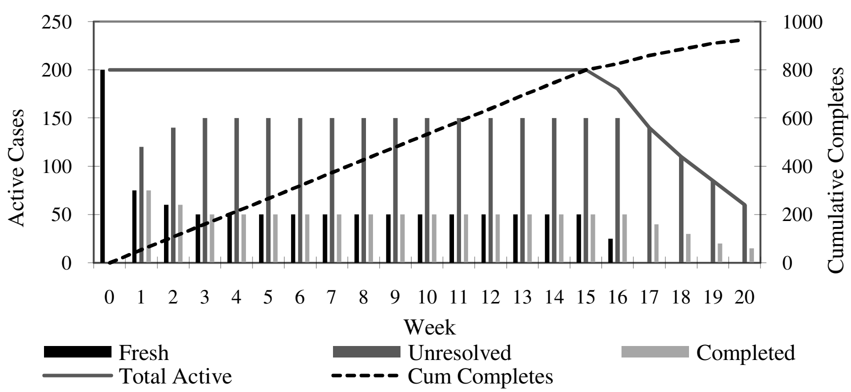


Fig. 19.3 Case mix obtained by release of fresh cases over time (according to Buckley et al 1998)

[8] For an example of a very dense interval, see the “Rolling Cross-Section Survey”(Schmitt-Beck et al. 2006). Here, new numbers are imported into the system every day. From a purely pragmatic point of view this is not feasible for a project of the size of GEDA.

The fieldwork procedures used in GEDA 2010 are closest to the second approach and can be described as interval management (see Figure 19.4). The figure shows the month in which the number was imported ("fresh") and the development of the total active cases, which is consists of the amount of fresh numbers, the unused numbers[9], and the unresolved numbers (call-backs). In the illustration, the amount of "unused numbers" is always shown at the moment before the numbers are imported, because it determines the timing of the number import. It is important for effective field management to keep the proportion of available telephone numbers constant and to find a balance between fresh sample and call-backs. In this context, the rules and priorities described in section 19.2 play a central role in field management.

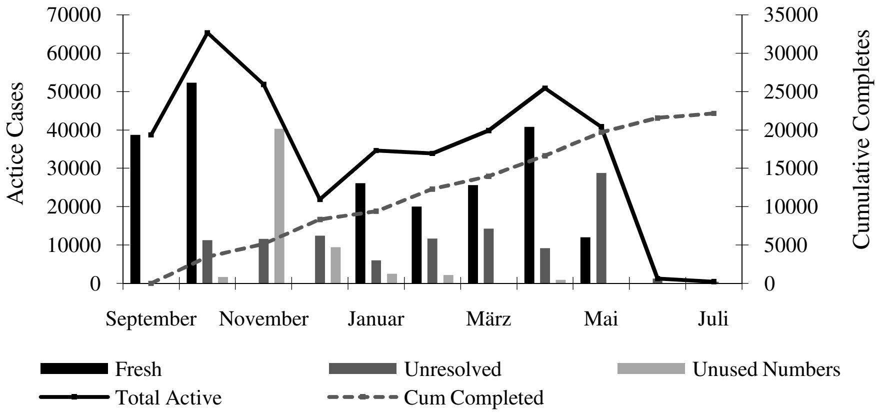


Fig. 19.4 Fieldwork during GEDA 2010

The advantage of this interval-based gross management lies in monitoring the fieldwork. An appropriate "fresh sample" can be introduced into the fieldwork at any time and is oriented towards the remaining call-backs and the ratio between completed interviews and the total target number of interviews. By the last field day, the sample was almost completely used up in the course of GEDA 2010[10]. The final batch of (12,000) fresh numbers was imported seven weeks before the end of the fieldwork, since at this point, it was possible to estimate how many new numbers would needed in order to reach the planned number of interviews.

[9] Deviating from Buckley's figure, Figure 4 presents the "unused numbers". These are identical to the "fresh" sample, but are already in the "field". The focus is on the timing of the number import, which is characterized here by the word "fresh".

[10] Number of call-backs on the last field day: 441 (0.2 % of total sample). The majority of these call-backs were the phone numbers of people who could not be reached during the field period.

19.3.3 Density of the Entered Numbers

The so-called mean entry density and the hit rate (ratio of neutral losses[11] as a percentage of the gross sample) were used to evaluate the quality and therefore the efficiency of a number sample. The entry density is often also referred to as the 'block density' and is calculated from the arithmetic mean of the number of telephone numbers listed in the phone book in all blocks of one hundred (Deutschmann and Häder 2002). The higher the entry density, the greater the likelihood of finding an actually assigned number. The average entry density fell from 47.33 in 1999 to 18.80 in 2008 (Häder et al. 2009). The mean entry density (block density) in the gross sample delivered for GEDA 2010 was 17.87. This is doubtlessly also one of the reasons why more than 40 call attempts were needed to complete an interview (see Table 19.1).

Another feature that is included in the sample delivered by GESIS[12] is information on whether a number is listed in the phone book or not. This will be studied in greater detail later on. As Häder found out as early as 1996 (Häder 1996), the percentage of people who are listed in a phone book has been declining since 1992. It is also known that newly assigned phone lines have been more rarely listed after 1991, when listing one's phone number in the directory became voluntary, which was related to a liberalization of the telecommunications market. As a result, phone books tend to list primarily older lines. This is confirmed by the findings of GEDA 2010.

It is difficult to determine how many up-to-date valid private landlines are listed in the phone book. In the case of the sample delivered for GEDA 2010, an entry was found for 18 % of telephone numbers.

About 86 % of the listed phone numbers and about 28 % of unlisted numbers led to an assigned line (see Figure 19.5). No distinction is made at this point between private households and business lines.

---

[11] In addition to the phone numbers clearly categorized as "not assigned" there is an unknown number of phone numbers which were for example constantly busy or where there was never an answer. According to AAPOR, a certain percentage of these numbers is rated as "not assigned". GEDA 2010's estimated "hit rate" is about 50 %.

[12] GESIS – Leibniz-Institut für Sozialwissenschaften.

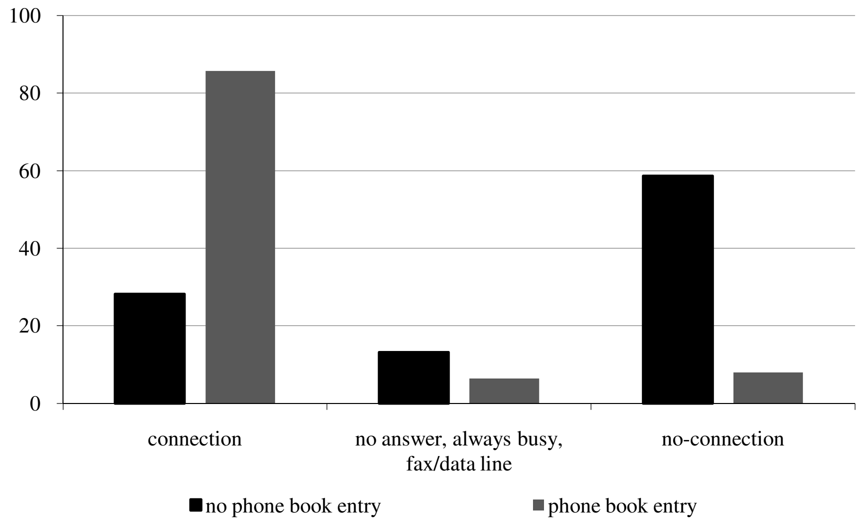


Fig. 19.5 Distribution of samples by whether a telephone number is listed or not

Since one common motivation for an unlisted phone line is to avoid unsolicited calls, it is worthwhile looking at the ratio between interviews and refusals by target persons as distinguished by the feature "phone book entry". An analysis of Table 19.4 reveals an interesting finding: the willingness to participate is greater among subjects with no entry (63.5 %) than among subjects who are listed in a phone book (58.3 %). This difference is also statistically significant, so that a person with no record in the phone book is nearly 25 % more likely to take part in an interview than someone who is listed (odds ratio = 1.25, $p < 0.001$). People listed in public phone books probably receive more (advertising) calls than the comparison group, and their willingness to participate decreases as a result.

Table 19.4 Interviews vs. refusals by whether a telephone number is listed or not

		Main questionnaire complete	Final refusal	Total
No phone book entry	n [%]	13,228 [63.5]	7,605 [36.5]	20,833 [100]
Phone book entry	n [%]	8,847 [58.3]	6,331 [41.7]	15,178 [100]

Table 19.5 gives an overview of the distribution of interviews according to education and age, comparing phone-book-listed with unlisted people.

It can be clearly seen that the proportion of people with a phone book entry increases with age. The figure ranges between only about 29 % for those aged 29 years to 63 % for people interviewed for GEDA aged 65 and older. Another striking feature is that across all age groups, people in the higher educational group are more likely to have an anonymous line.

Table 19.5 Phone book entry and education by age

Age groups	*Phone book entry*		*Lower education*		*Inter-mediate education*		*Higher education*	
	n	[%]	n	[%]	n	[%]	n	[%]
up to 29								
No entry	2,521	71.4	413	67.6	1806	72.0	302	74.0
Phone book entry	1,008	28.6	198	32.4	704	28.0	106	26.0
30 - 44								
No entry.	3,962	69.9	473	65.9	2159	69.0	1330	73.0
Phone book entry	1,704	30.1	245	34.1	968	31.0	491	27.0
45 - 64								
No entry	4,238	56.8	931	51.8	1946	57.4	1361	59.8
Phone book entry	3,229	43.2	868	48.2	1446	42.6	915	40.2
65 +								
No entry	1,437	37.1	526	32.2	458	37.4	453	44.6
Phone book entry	2,437	62.9	1109	67.8	765	62.6	563	55.4

19.3.4 Contact Times: Shift Times and Weekdays

Another aspect that has a regulating effect on the process of data collection concerns the best possible time to call. Optimal planning of these times increases the likelihood of contacting people and therefore successfully completing interviews. According to studies, promising times to reach people are in the evening after 5 pm and on weekends (Weeks 1987, Swires-Hennesy and Drake 1992). In addition, the likelihood of reaching people who seldom answer their phone can be boosted best by varying the days of the week and times of the day when calls are attempted (Hüfken 2000).

In the course of the ten-month GEDA 2010 study, the calls were made on weekdays between 10 am and 2 pm[13] and between 4:30 and 8:30 pm[14], as well as on Saturdays between 2 and 6 pm. The calling periods – between 10 am at the earliest and 8:30 pm at the latest – thus matched the recommendations issued by the Working Group of German Market and Social Research Institute (ADM 2008) in its Guidelines for Telephone Surveys.

Performance by Shift Times

In total, there are 40 available interviewer workstations per shift in the CATI laboratory. Because experience shows that people are harder to reach in the morning, no more than 20 stations were occupied in the early shifts.

[13] Except for Thursdays; after early May, the morning shifts were limited to Mondays and Wednesdays.

[14] From mid-March the calling time was extended to about 4.5 hours (from 4 to 8:30 pm).

Table 19.6 Performance by days of the week and shifts

	Weekdays						Weekend
	Mo	Tue	We	Thu	Fr	*Total*	Sa
Early shifts							
Average no. of interviews / interviewers	2.22	1.94	2.10		2.28	*2.13*	
Late shifts							
Average no. of interviews / interviewers	2.99	2.80	2.83	2.59	2.77	*2.79*	2.54

As expected, a direct comparison between early, late and Saturday shifts shows a better performance[15] in the late shifts (average 2.79) compared to the general performance of the day shift (average 2.13). Looking at the weekdays, it becomes clear that the best performance during the morning shifts is achieved on Mondays and Fridays. The patterns for the late shifts are a little different. Here, the high level of performance of the Monday shifts is maintained until the middle of the week; starting on Thursday, there is then a slight decrease until Saturday. It should be noted here that a different time interval was set for the Saturday shift, which concentrates especially on the afternoon (see Table 19.6). The main reason for this time frame was a consideration for the respondents. If, however, the potential for conducting interviews were the main criterion, it would be worth considering moving the Saturday shift to the mornings and evenings (Weeks et al. 1987).

Who is Reached when?

Section 19.3.4 has shown that the accessibility of respondents varies depending on the time of day and the day of the week. Furthermore, the question has to be raised, whether the choice of contact times influences the composition of the sample.

Looking exclusively at the gender distribution in relation to the shifts reveals a striking difference: Almost two thirds of all interviews achieved during the morning shifts are conducted with women (see Table 19.7). An analysis of the distributions of the late and Saturday shifts shows that they cannot compensate for this imbalance.

Table 19.7 Gender distributions by shifts, in percent

	Early shifts	Late shifts	Saturday shifts
Male	37.4	44.3	49.9
Female	62.6	55.7	50.1
Total	100	100	100

[15] In this context the performance value describes the ratio of interviews per shift for each interviewer.

A comparison of shift times and gender in terms of the selected socio-demographic variables of age and education also reveals major differences in distribution (see Table 19.8).

For example, more people of both sexes with a lower level of education are interviewed on the early shifts, and the difference is more pronounced and statistically significant among women. The chance of interviewing women with a lower level of education is 25 % higher on the morning shifts than it is on the evening shifts (odds ratio = 1.25, $p < 0.001$), even after controlling for age, household size and employment. Furthermore, the percentage of people aged 65 or older in the morning shifts is higher and statistically significant (odds ratio = 1.39, $p < 0.001$) for both men and women. The difficult-to-reach group of people below the age of 30 can best be persuaded for an interview in the course of the afternoon or on Saturdays.

Therefore, it can be stated that the composition of the sample in terms of socio-demographic characteristics differs depending on the time of day (early or late shift) and on the day of week (weekdays vs. Saturday).

Because the last-birthday method is used to select the respondents at the household level, the differences resulting from the different shift times should not lead to a distortion of the overall sample. After all, there is a good chance of being able to conduct the interview with the selected person, thanks to the broad range of calling times and the high number of contact attempts.

Table 19.8 Distributions of education and age by gender and shifts, in percent

	CASMIN educational groups in percent (n=19,664)			*Age groups in percent (n = 19,736)*			
	Lower education	Inter-mediate education	Higher education	- 29	30-44	45-64	65 +
Men							
Early shift	25.8	42.6	32.0	15.2	22.2	32.1	30.6
Late shift	23.6	44.4	32.0	19.0	27.6	35.8	17.5
Saturday shift	20.7	46.8	32.4	24.7	27.7	35.0	12.6
Total	23.6	44.4	32.1	19.0	26.6	35.0	19.4
Women							
Early shift	26.9	53.0	20.1	13.2	28.7	34.3	23.8
Late shift	21.3	54.8	23.9	16.8	28.4	38.1	16.8
Saturday shift	21.3	53.6	24.9	18.9	27.8	37.6	15.8
Total	22.7	54.3	23.1	16.1	28.4	37.1	18.4

This raises the question of whether the contact times have an influence on the sample distribution, despite a selection key and repeated contact attempts.

With their findings relating to the so-called “double hurdle”, Asef and Riede (2006) show that the distribution of contact times has a significant influence on the measurement of employment, although an attempt was made to minimize sample distortions by both using the Kish-Selection-Grid in the selection of respondents and spreading contact attempts over different days and times. According to the authors, the likelihood of interview participation is higher if the first contact is made directly with the target person specified by the selection mode. In cases where the person who first answered the phone was not the target person and a further hurdle had to be taken via a contact person in addition to the selection procedure, a third of the target persons refused to participate or could not be reached. These people are then under-represented in the survey. It is problematic if there are systematic differences between the target persons who are reached with the first contact and those who are only reached via at least one contact person. This was proven for the period before 5 pm with reference to employment (Asef and Riede 2006).

The "double hurdle" can also be demonstrated on the basis of the data from GEDA 2010. The loss of people who were selected but could not be reached at the first contact was as high as 75 %[16]. Figure 19.6 clearly illustrates the better chances of an interview if the first contact is made with the target person and therefore no further obstacles need to be overcome.

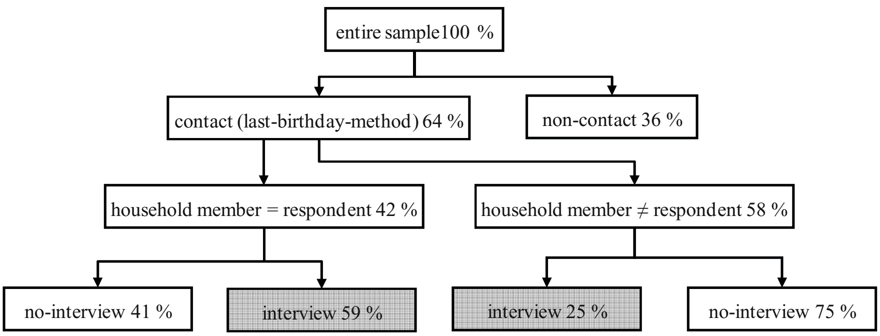


Fig. 19.6 Illustration of the double hurdle in GEDA 2010 (according to Asef and Riede 2006)[17]

If this knowledge is transferred to the sample distribution in GEDA 2010, the higher proportion of women can be explained by said double hurdle relative to the overall sample, because many interviews were made with women during the morning shifts. On the other hand, the shares of both interviewees with a lower education and people over 65, who are underrepresented in the overall distribution, are largest in the morning shifts.

It becomes clear that shift times do have effects on selection. They are offset neither by the selection mode nor by a broad distribution of contact attempts. This should be taken into account when planning studies. In this context, the extent to

[16] All contacts with people who could not unequivocally be identified as the target person during the first call were incorporated into the analysis (for example if the contact person had no time to determine the target person, or if the person who answered the phone was unable to provide information). This probably somewhat overestimates the number of failed calls.

[17] All numbers called minus the ones that are not eligible.

which a (perhaps) smaller response may be acceptable in the interests of obtaining a more representative sample must also be considered.

19.4 Summary and Outlook

The aim of this study was to discuss selected parameters for the fieldwork management of a study conducted on the basis of the Gabler-Häder method. An attempt was made to present factors that influence the composition of the completed sample. It should be noted at this point that this chapter concentrates on the factors that are decisive in daily fieldwork. As a result, factors such as the number of calls, the effect of the calls after a "soft refusal" and the choice of contact times were analyzed and described.

1. As regards the number of call attempts it can be stated that, especially in the context of a cost-benefit analysis, more than ten contact attempts might seem inefficient if the focus were only directed at maximizing potential. After all, 90 % of the subjects agree to participate within ten call-up attempts. However, an analysis of the sample composition as a function of the number of call attempts shows that there are differences between the people who were interviewed after the first-to-tenth call and those who were not reached for the interview until after the tenth call attempt. Above all, younger people who were difficult to reach could be persuaded to take part by calling more than ten times. The results clearly show that the maximum number of calls should depend on the purpose and prerequisites of the study. When deciding on this number, a compromise needs to be found between the aims of potential maximization, sample representativeness, cost considerations and, not least, the acceptance by the subjects. The result will differ depending on one's perspective and the relative weight given to each aspect.
2. Calling back after a "soft refusal" seems justified, since a renewed contact can persuade a certain percentage (8 %) of people to take part in an interview. Since this group of people differs from the other respondents in terms of certain socio-demographic characteristics (e.g. people over the age of 65 and people with a lower education), not making an additional attempt to contact them would lead to under-coverage of these groups in the sample.
3. Also, when planning call times, it is essentially a matter of finding the balance between two objectives that are not always congruent. One is to conduct interviews as efficiently as possible, i.e. to match the contact times with the time when the respondents are easiest to reach in order to contact the target person with the least effort (few call attempts). On the other hand, the primary purpose of a scientific survey is always to achieve a sample composition that is as representative as possible. In this context, matching the days and times when most people are accessible must not be the sole criterion for deciding on the shift times. Furthermore, the effects of the selection of contact times on the composition of the sample should be intensely analyzed and shift times chosen to suit the population group of which the survey group aims to be representative.

By providing this overview of selected methodological parameters of a large-scale, health-related study conducted by telephone, this study aims to stimulate ideas for conducting similarly designed studies and for more in-depth analyses. In future, the need for information and therefore surveys of very different populations can be expected to increase rather than decline. A combination of scientifically oriented research and knowledge from opinion research therefore seems to be a highly useful way of trying to achieve the best possible results in accordance with uniform standards.

References

AAPOR (American Association for Public Opinion Research), Standard definitions. Final disposition codes and outcome rates for surveys. Lenexa, Kansas (2010)

ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.) (Working Group of German Market and Social Research Institutes), Richtlinie für telefonische Befragungen (Guidelines for Telephone Surveys) (2008)

Asef, D., Riede, T.: Kontaktzeiten in einer Telefonerhebung – wie beeinflussen sie die Messung der Erwerbstätigkeit ? (Contact times in telephone surveys – how do they influence the measurement of employment?). In: Bundesamt, S. (ed.) Wirtschaft und Statistik, Wiesbaden, vol. 6, pp. 581–586 (2006)

Brauns, H., Steinmann, S.: Educational reform in France, West Germany and the United Kingdom: updating the CASMIN educational classification. ZUMA-Nachrichten 23, 7–44 (1999)

Buckley, P., Dennis, J.M., Saulsberry, C., Coronado, V.G., Ezzati-Rice, T., Maes, E., Rodin, A., Wright, R.A.: Managing 78 simultaneous RDD samples. In: Proceedings of the Section on Survey Research Methods, pp. 957–961. American Statistical Association, Alexandria (1998)

Couper, M.P., Coutts, E.: Online-Befragung. Probleme und Chancen verschiedener Arten von Online-Erhebungen (Online surveys. Problems and opportunities of different kinds of online survey). In: Dieckmann, A. (ed.) Methoden der Sozialforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, vol. 44, pp. 217–243. VS Verlag für Sozialwissenschaften, GWV Fachverlage GmbH, Wiesbaden (2006)

Curtin, R., Presser, S., Singer, E.: Changes in telephone survey nonresponse over the past quarter century. Public Opinion Quarterly 69, 87–98 (2005)

De Leeuw, E.D., De Heer, W.: Trends in household survey nonresponse: A longitudinal and international comparison. In: Groves, R.M., et al. (eds.) Survey Nonresponse, pp. 41–54. Wiley, New York (2002)

Deutschmann, M., Häder, S.: Nicht-Eingetragene in CATI-Surveys (Unlisted numbers in CATI surveys). In: Gabler, S., et al. (eds.) Telefonstichproben: Methodische Innovationen und Anwendungen in Deutschland, Waxmann, pp. 68–84 (2002)

Faas, T.: Umfragen im Umfeld der Bundestagswahl 2002: Offline und Online im Vergleich (Surveys in the context of the 2002 federal parliamentary elections: offline and online, a comparison). ZA-Informationen 52, 120–135 (2003)

Glemser, A.: ADM-Telefonstichproben in der Praxis (ADM telephone samples in practice). In: Gabler, S., et al. (eds.) Telefonstichproben Methodische Innovationen und Anwendungen in Deutschland, Waxmann, pp. 46–58 (2002)

Häder, S.: Wer sind die »Nonpubs«? Zum Problem anonymer Anschlüsse bei Telefonumfragen (Who are the 'nonpubs'? On the problem of anonymous phone lines in telephone surveys). ZUMA-Nachrichten 39, 45–68 (1996)

Häder, S., Gabler, S.: Ein neues Stichprobendesign für telefonische Umfragen in Deutschland (A new sample design for telefone surveys in Germany). In: Gabler, S., et al. (eds.) Telefonstichproben in Deutschland, pp. 69–88. Westdeutscher Verlag, Opladen (1998)
Häder, S., Gabler, S.: Generierung von Telefonstichproben mit TelSuSa (Generating telephone samples with TelSuSa). ZUMA-Nachrichten 44, 138–143 (1999)
Häder, S., Gabler, S., Heckel, C., et al.: Stichprobenziehung, Gewichtung und Realisierung (Drawing samples, weighting, realization). In: Häder, S., Häder, M. (eds.) Telefonbefragungen über das Mobilfunknetz: Konzept, Design und Umsetzung einer Strategie zur Datenerhebung, pp. 21–82. VS Verlag für Sozialwissenschaften, Wiesbaden (2009)
Hansen, S.E.: CATI Sample Management Systems. In: Lepkowski, J.M., et al. (eds.) Advances in Telephone Survey Methodology, pp. 340–385. Wiley, New Jersey (2008)
Hartmann, P., Schimpl-Neimanns, B.: The representativity of sociodemographic features of Allbus multivariant analyses regarding the middle-class bias in survey research. ZUMA-Arbeitsbericht 1 (1992)
Hüfken, V.: Kontaktierung bei Telefonumfragen. Auswirkungen auf das Kooperations- und Antwortverhalten (Contacting in telephone surveys. Effects on cooperation and response behavior). In: Hüfken, V. (ed.) Methoden in Telefonumfragen, pp. 11–31. Westdeutscher Verlag, Opladen (2000)
Kohler, M., Rieck, A., Borch, S., Ziese, T.: Erster telefonischer Gesundheitssurvey des Robert Koch-Instituts – methodische Beiträge (The Robert Koch Institute's first telephone health survey – methodological contributions). In: Beiträge zur Gesundheitsberichterstattung des Bundes. Robert Koch-Inst., Berlin (2005)
Kurth, B.M., Lange, C., Kamtsiuris, P., Hölling, H.: Gesundheitsmonitoring am Robert Koch-Institut – Sachstand und Perspektiven (Health monitoring at the Robert Koch Institute – current position and prospects). Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz 52(5), 557–570 (2009)
Rat der Deutschen Markt- und Sozialforschung e.V (Council of German Market and Social Research). Richtlinie für telefonische Befragungen (Guidelines for Telephone Surveys) (2008), `http://ratmarktforschung.de/fileadmin/user_upload/pdf/R04_RDMS.pdf` (accessed February 01, 2011)
RKI (Robert Koch-Institut) Daten und Fakten: Ergebnisse der Studie "in Deutschland aktuell 2009" (Data and facts: results of the "German Health Update" 2009). In: Beiträge zur Gesundheitsberichterstattung des Bundes. Robert Koch-Inst., Berlin (2011)
Schmitt-Beck, R., Faas, T., Holst, C.: Der Rolling Cross-Section Survey – ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung (The Rolling Cross-Section Survey – an instrument for analysing dynamic opinion-development processes). Bericht zur ersten deutschen RCS-Studie anlässlich der Bundestagswahl 2005 (Report of the first RCS study on the occasion of the federal parliamentary elections in 2005). ZUMA-Nachrichten 58, 13–49 (2006)
Schnauber, A., Daschmann, G.: States or traits? Factors influencing the willingness to participate in telephone surveys. Methoden — Daten — Analysen 2(2), 97–123 (2008)
Sudman, S.: Survey research and ethics. Advances in Consumer Research 25, 69–71 (1998)
Swires-Hennessy, E., Drake, M.: The optimum time at which to conduct survey interviews. International Journal of Market Research 34(1), 61 (1992)
Von der Heyde, C.: Das ADM-Telefonstichproben-Modell (The ADM telephone sample model). In: Gabler, S., et al. (eds.) Telefonstichproben: Methodische Innovationen und Anwendungen in Deutschland, Waxmann, pp. 32–45 (2002)
Waksberg, J.: Sampling methods for random digit dialing. Journal of the American Statistical Association 73, 40–46 (1978)
Weeks, M.F., Kulka, R.A., Pierson, S.A.: Optimal call scheduling for telephone survey. Public Opinion Quarterly 51, 540–549 (1987)

The Authors

Anna Andreenkova

She is Deputy Director and Senior Researcher in CESSI (Institute for Comparative Social Research), Russia as well as Director of the division of surveys in countries of the former Soviet Union in CESSI. Anna Andreenkova is also National coordinator of the European Social Survey in Russia. She holds a Dr. in political sciences.

Current research areas: methodology of cross-national comparative survey research; translation issues and questionnaire design for comparative surveys; political values, political cleavages and participation in new democracies.

Jelke Bethlehem

He is Professor of Statistical Information Processing at the Amsterdam School of Economics (University of Amsterdam) and Senior Project Manager, Department for Statistical Methods, Statistics Netherlands. He holds an MA in Mathematical Statistics, University of Amsterdam and a PhD in Social Sciences, University of Amsterdam.

His main research interests cover survey research, non-sampling errors, data editing, computer-assisted interviewing and electronic dissemination of statistical information.

Dirkjan Beukenhorst

He studied anthropology and sociology at the University of Amsterdam. Dirkjan Beukenhorst conducted both qualitative and quantitative research on diverse subjects, such as agricultural modernization, small-scale fisheries, migration and ethnic minorities. After joining Statistics Netherlands he continued doing both quantitative and qualitative research. He now works at the methodology department, working for the Questionnaire laboratory, pre-testing questionnaires, and contributing to mixed-mode designs of data collection.

Mario Callegaro

He is Survey Research Scientist at Google, London. He focuses on measuring customer satisfaction. Mario holds a PhD in Survey Research and Methodology from the University of Nebraska, Lincoln. Prior to joining Google, Mario Callegaro was working as survey research scientist for Knowledge Networks. He is associate editor of Survey Research Methods.

Current research areas: web survey design, smart-phone surveys, telephone and cell phone surveys, and questionnaire design.

Fannie Cobben

She studied Econometrics at the Free University of Amsterdam. Her main subject was Statistical Econometrics. After obtaining her pre-doctoral degree, Fannie Cobben joined the Department of Methodology and Quality of Statistics Netherlands in 2004. In 2009, she finished her Ph.D.-thesis on nonresponse adjustment in household surveys. She now works as a survey methodologist and a project manager at Statistics Netherlands.

Her research interests: methodology behind conducting surveys concerning nonresponse, mixed-mode surveys and indicators for the quality of survey response.

Michèle Ernst Stähli

She is Senior Researcher at FORS, the Swiss Centre of Expertise in the Social Sciences at the University of Lausanne, Switzerland. After 15 years of academic research in the field of sociology of work and family and 3 years of managing scientific and administrative surveys in a fieldwork agency, she is now the National Coordinator for the European Social Survey and in charge of the international survey unit at FORS. Michèle Ernst Stähli is Dr ès sociology and social sciences.

Current research areas: translations, educational measure, mixed mode design.

Siegfried Gabler

He is the chief of the Statistics team at GESIS – Leibniz-Institute for the Social Sciences, Mannheim. Member of the International Statistical Institute and member of the Sampling Expert Panel of the European Social Survey. Siegfried Gabler teaches at the University of Mannheim. He holds a Dr in Mathematics and a PD in Economics.

His research area covers sampling designs, especially for telephone surveys and for cross-cultural surveys, weighting for nonresponse, design effects, and decision theoretic justification of sampling strategies. He is involved in several projects in the context of telephone surveys and in the Census 2011 project for Germany.

Michael Häder

He holds the chair for Empirical Research Methods at the Institute for Sociology at the Dresden University of Technology since 2001. He studied sociology at the Humboldt-University in Berlin. Thereafter Michael Häder was research assistant at the University in Leipzig, at the Academy of Sciences in Berlin and at the Centre for Survey Research and Methodology (ZUMA) in Mannheim.

His current working fields are: the Delphi-Method and Survey Research.

Sabine Häder

She is Senior Statistician at GESIS – Leibniz-Institute for the Social Sciences, Mannheim. She is also a member of the Sampling Expert Panels of the European Social Survey and of SHARE; and National Sampling Manager of the Programme for the International Assessment of Adult Competencies (PIAAC). Sabine Häder holds a Dr in Economics.

Current research areas: sampling designs, especially for telephone surveys and for cross-cultural surveys, weighting for nonresponse, design effects.

Christiane Heckel

She leads the department of sampling at the BIK ■ ASCHPURWIS + BEHRENS GmbH. She holds a degree as Diplom-Sozialwirtin.

Christiane Heckel is responsible for the built-up and development of the ADM-CATI-Sampling Frame and also for the F2F-Sampling-Frame. Since 2006 CATI-Sampling-Frames for various European countries are originated. On the basis of sampling frames, she is also involved in the process of modeling Dual-Frame-Approaches in different European countries.

Franziska Jentsch

She studied Psychology at the Humboldt University, Berlin (Diplom-Psychologist) and behavior-trainer, scientific researcher at the Robert Koch Institute Berlin.

Key working areas: quality management, conception and evaluation of interviewer training, field control and supervision in telephone surveys.

Dominique Joye

From 1999 to 2006, he was director of the SIDOS, the Swiss data archive, and associate professor of methodology for the social sciences at the University of Neuchâtel. Since then Dominique Joye is professor of Sociology at the University of Lausanne, and associated to the Swiss part of the main social sciences international surveys (ESS, EVS, ISSP) realized by FORS. He holds a PhD in political sciences from the University of Geneva.

Main research area: survey methodology, but also on stratification analysis and social change, with a life course perspective.

Olena Kaminska

She is survey statistician at the University of Essex. Olena Kaminska has a PhD in Survey Research and Methodology from the University of Nebraska-Lincoln.

Her research interests encompass sampling, non-response and measurement errors and the interaction among them in cross-sectional, longitudinal and comparative surveys.

Kathrin Kissau

She is senior researcher at FORS - Swiss Centre of Expertise in the Social Sciences - Lausanne. In 2008 she received a Dr degree from the University of Muenster.

Her main research topics: new media developments and their political and societal effects, political participation, and representation; migrants´ media culture as well as online research methods, appointed coordinator of the European COST Action "WEBDATANET: web-based data-collection – methodological challenges, solutions and implementation" 2010-2014.

Mike Kühne

He is Postdoctoral Research Associate at the Dresden University of Technology and Lecturer on applied statistics and data analysis at the Chair of Empirical Research.

Current research areas: educational sociology, response behavior and mode effects.

Vesa Kuusela

He is Head of the Development Social Survey Unit, Statistics Finland. His expertise: advisor in survey methodology, especially computer assisted data collection. He holds a Dr degree.

Special research interests: telephone coverage and history of survey research.

Iris Lehnhoff

She is student assistant at GESIS – Leibniz-Institute for the Social Sciences, Mannheim and Masters student of Sociology at the Heidelberg University. Previously: Bachelor of Arts in Sociology at Mannheim University (2006-2009).

Oliver Lipps

He is head of the methodological research programme at FORS – Swiss Centre of Expertise in the Social Sciences - Lausanne and member of the Swiss Household Panel (SHP) team. In addition, Oliver Lipps is lecturer in survey methodology and survey research at the Institute of Sociology at the University of Basel (Switzerland). He holds a degree as Dr. ing.

Research areas: unit and item nonresponse in cross-sectional, and longitudinal designs; incentive effects, interviewer effects, and income imputation methods in cross-national and in panel surveys, using data from, e.g., SHARE, ESS and SHP.

Peter Lynn

He is Professor of Survey Methodology at the University of Essex. His research interests and publications encompass all aspects of survey design and data collection. Peter Lynn is editor of Methodology of Longitudinal Surveys, published by Wiley (2009), co-founder of the European Survey Research Association, and was previously vice president of the International Association of Survey Statisticians.

Elisabeth Mardian

She is student assistant at GESIS – Leibniz-Institute for the Social Sciences, Mannheim and Masters student of Sociology at the Heidelberg University. Previously: Bachelor of Arts in Sociology at Mannheim University (2006-2009).

Gerd Meier

He is Professor of Research Methodology and Market Psychology at the Leuphana University of Lueneburg since 2000. There Gerd Meier is teaching Market Psychology, Methodology, General Psychology, and Developmental Psychology. From 1993 - 2000 he worked at GfK Nuremberg Market Research Institute as Senior Consultant.

His current research areas cover: Methods in Market Research, Advertising, and Consumer Behaviour.

Elizabeth Nelson

She is Executive chairman of Fly Research Ltd, London, specialized in online and mobile internet research. Previously Elizabeth Nelson (Dr., OBE) was founder and chairman of Taylor Nelson (TNS); president of the World Association for Public and Opinion Survey, secretary-treasurer of the Market Research Society, member of MRS Disciplinary Tribunal; Gold Medal MRS.

Teresio Poggio

He is executive coordinator of the Research Laboratory (survey unit and data archive) of the Department of Sociology and Social Research at the University of Trento, Italy. His main substantive research interests are on housing and welfare systems and in social stratification. He holds a PhD in Sociology and Social Research from the University of Trento.

Teresio Poggio's main methodological research interests include the use of new technologies in the collection of survey data.

Alexandre Pollien

He has a Graduate in sociology from University of Geneva.

Alexandre Pollien works on survey methodology at the Swiss Centre of Expertise in the Social Sciences (FORS). His interests include participation in surveys and experience of interviewers.

Elizabeth Reis

She is Professor at the Quantitative Methods Department, Lisbon University Institute (ISCTE-IUL), Portugal.

Her current research areas: multivariate statistics, segmentation methods, survey, and census methodologies.

Marlène Sapin

She is researcher at the Swiss Centre of Expertise in the Social Sciences (FORS), at the University of Lausanne. Marlène Sapin works on the monitoring of International surveys in Switzerland~~,~~ and on methodological research.

Her main research interests include survey methodology, measurement equivalence, life course research, and social network analysis.

Tino Schlinzig

He studied sociology, law and communication research at Dresden the University of Technology in Dresden (TUD) and Victoria University of Wellington. Since August 2008 Tino Schlinzig works as research associate at the Chair of Micro-Sociology and since July 2009 at the Collaborative Research Centre 804; Transcendence and Common Sense' at TUD. From 2006 to 2008 he was research associate at the Chair of Social Empirical Research at TUD.

His research focuses on Sociology of the Family, Sociology of Knowledge, Qualitative and Quantitative Empirical Methods.

Patrick Schmich

He studied Sociology at the Freie Universität Berlin (Diplom-Sociologist); presently Patrick Schmich is Researcher at the Robert Koch Institute Berlin.

Key working areas: Project Management for telephone surveys, quality management, and survey methodology.

Götz Schneiderat

He studied sociology, history, and philosophy at the University of Technology in Dresden (TUD). Since September 2006 Götz Schneiderat has been working as research associate at the Chair of Social Empirical Research in the project funded by the German Research Foundation (DFG); 'General Social Telephone Surveys via Mobile Phone (CELLA 2)' at TUD.

His research focuses on Qualitative and Quantitative Empirical Methods, especially Survey Methodology, and Telephone Surveys.

Barry Schouten

He is Senior Methodologist in the Division of Methodology and Quality at Statistics Netherlands. Barry Schouten holds a PhD.

His areas of research interest include nonresponse and response bias in mixed-mode surveys, indicators for representativeness of response, and adaptive and responsive survey designs.

Matti Simpanen

He is the Head of Research in the CATI Centre of the Social Survey Unit at Satistics Finland.

Matti Simpanen's research area covers all aspects of methodology for telephone surveys.

Ana Slavec

She (BSc.) is Research Assistant at Faculty of Social Sciences, University of Ljubljana, Slovenia.

Ana Slavec's current research areas: post-survey adjustments, dual frame surveys, nonresponse, internet surveys, mixed modes, and cross-national research.

Vasja Vehovar

He is Professor of Statistics at Faculty of Social Sciences, University of Ljubljana, Slovenia. He is a member of the European Social Survey methods group; developer of the WebSM portal, Principal investigator of the national project Internet in Slovenia (RIS) and also involved in various EU information society research projects.

Current research areas: survey methodology - especially web surveys, mixed modes, nonresponse, sampling, complex designs, post-survey adjustments, cross-national research, and substantive Internet research.

Paula Vicente

She is Assistant Professor at the Quantitative Methods Department, Lisbon University Institute (ISCTE-IUL), Portugal.

Her current research areas: telephone surveys (including mobile phones), mixed-mode effects, use, and analysis of paradata.

Kathrin Wiese

She is working since 2006 at the BIK ■ ASCHPURWIS + BEHRENS GmbH. Kathrin Wiese supervises beside the collection and maintenance of official statistics and statistics based on secondary data also the sampling of different market research studies. The emphasis is on the development of sampling frames for other European countries, especially on sampling of mobile numbers.

Index

O

P

Q

R

S

T

W