Deliverable Number: 2.17
Deliverable Title: Report on what persons live in institutions and the most relevant institution types they live in, the availability of data to select them and their ability to be reached in practice
Work Package: 2 – Representing the population
Deliverable type: Report
Dissemination status: Public

Submitted by: ESS ERIC

Authors: Jan-Lucas Schanze (GESIS – Leibniz Institute for the Social Sciences), Maayan Levinson (The Hebrew University of Jerusalem)
Date: May 31, 2019

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 654221.
SERISS (Synergies for Europe’s Research Infrastructures in the Social Sciences) aims to exploit synergies, foster collaboration and develop shared standards between Europe's social science infrastructures in order to better equip these infrastructures to play a major role in addressing Europe’s grand societal challenges and ensure that European policy making is built on a solid base of the highest-quality socio-economic evidence.

The four year project (2015-19) is a collaboration between the three leading European Research Infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey for Health Aging and Retirement in Europe (SHARE ERIC) and the Consortium of European Social Science Data Archives (CESSDA AS) – and organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and the WageIndicator Survey.

Work focuses on three key areas: Addressing key challenges for cross-national data collection, breaking down barriers between social science infrastructures and embracing the future of the social sciences.

Please cite this deliverable as: Schanze, J. and M. Levinson (2019). Report on what persons live in institutions and the most relevant institution types they live in, the availability of data to select them and their ability to be reached in practice. Deliverable 2.17 of the SERISS project funded under the European Union’s Horizon 2020 research and innovation programme GA No: 654221. Available at: www.seriss.eu/resources/deliverables
7 Survey interviews with institutionalized respondents

7.1 Cognitive screening .................................................. 51
7.2 Proxy interviews ...................................................... 52
7.3 Potential adaptations of the questionnaire and survey procedures ............ 55
    7.3.1 Survey mode ................................................... 55
    7.3.2 Informed consent ............................................. 57
    7.3.3 Wording and simpler language ............................ 57
    7.3.4 Timing and multiple meetings ............................. 58
    7.3.5 Translation .................................................... 59
7.4 Incentives ............................................................. 60
7.5 Ethical considerations of surveying the institutionalized population .......... 60

8 Discussion .................................................................. 63

8.1 Is it necessary to include institutionalized populations? ......................... 63
8.2 Is it feasible to include institutionalized populations? ............................. 64

9 Conclusion ................................................................... 66

References ...................................................................... 68

Appendix ......................................................................... 75

Additional Figures ........................................................ 75
List of Figures

1  Definition of institutions and the institutionalized populations based on SERISS Deliverable 2.16 (Schanze 2017) ................................................. 10
2  Share of the population living in collective households in European countries (Eurostat 2016) ................................................................. 14
3  Share of the institutionalized population older than 65 years within the institutionalized population of European countries (Eurostat 2016) ......................... 17
4  Number of prisoners per 100,000 inhabitants in European countries in 2010 and 2016 ............................................................... 18
5  Gender distribution of European prisoners in 2016 .......................................................... 18
6  Types of surveys conducted by the survey managers responding to the SERISS expert survey .............................................................................. 26
7  Types of institutions covered by the survey managers responding to the SERISS expert survey .............................................................................. 26
8  SERISS expert survey: Who was contacted in the process of gaining access to the institutionalized respondent? (Multiple selection allowed, percentage of all experts) .... 46
9  SERISS expert survey: In general terms, how would you best describe the process of gaining access or approval of the operators of institution? (N = 44 in the entire sample) 47
10 SERISS expert survey: How did you decide whether a proxy interview was used instead of a direct interview? (multiple selection allowed) ........................................ 53
11 SERISS expert survey: Who could answer as proxy on behalf of the respondent? ................ 54
12 SERISS expert survey: Which survey mode would you recommend for surveying the institutionalized population? ................................................................. 56
13 Recruitment protocol for people residing in care homes as part of the Newcastle 85+ study. 75
14 Part of the summary the tips and advice by Trulson and Marquart (2004). .......................... 76
15 An excerpt out of the full report of Survey of adults with learning difficulties in England 2003/4 ................................................................. 77

List of Tables

1  List of socio-demographic variables that have a positive statistical impact on the likelihood of living in institutions for the elderly (Schanze and Zins 2019) ..................... 20
List of health-related variables that have a positive statistical impact on the likelihood of living in institutions for the elderly (Schanze and Zins 2019) ........................................... 21

Sampling frames used by the survey programs (Source: SERISS Expert Survey 2018) ........................................... 30

To what extent does the sampling frame used suffer from coverage issues? (Source: SERISS Expert Survey 2018, stars indicating contradicting categories used for a single register) ........................................... 33

Does the sampling frame allow the identification of institutionalized residents? (Source: SERISS Expert Survey 2018) ........................................... 39
1 Introduction

Many surveys exclude the institutionalized populations. Social surveys like the German ALLBUS, the European Social Survey (ESS), the European Value Study (EVS), the European Quality of Life Surveys (EQLS), the Gender and Generations Survey (GGS), or the American General Social Survey (GSS) only cover the population living in private households. Even health surveys like the French and the Spanish Health Barometers, the Swiss Health Survey (SGB), the Health Survey for England (HSE), or the Welfare and Services in Finland Study (HYPA) exclude institutionalized residents. The same restrictive definition of the target population is applied by some ageing surveys such as the Irish Longitudinal Study on Ageing (TILDA), the Berlin Aging Study (BASE), or the Belgian Ageing Studies (BAS).

Obviously, survey managers working on those surveys made a deliberate decision to exclude a certain part of the population on the basis of a cost-benefit analysis (Schanze and Zins 2019). In line with a classification of hard-to-survey populations (Tourangeau 2014), persons living in institutions can be labeled as hard-to-reach and sometimes hard-to-interview (Feskens 2009; Schanze 2017). Protective gatekeepers working in the institution or relatives of institutionalized residents create additional barriers for interviewers. Moreover, depending on the specific type of institution, survey interviews might be more demanding due to language or cultural barriers, potential functional and cognitive impairments of respondents, and questions that do not really apply to the living situation of institutionalized residents.

To quote another SERISS Deliverable, "SERISS Work Package 2 'Representing the population' is focused on ensuring that European surveys continue to remain state of the art when it comes to accurately describing phenomena in the population" (Scherpenzeel et al. 2016, 4). As part of this Work Package, the institutionalized population was identified as a relevant and potentially critical group and a distinct task was concerned with analyzing this group. Researchers working on this task compiled survey practices in the SERISS Survey Inventory. This inventory contains more than 150 examples of international and national survey programs that covered institutionalized residents. Technical reports published by those survey programs are a first source of information for this Deliverable. The SERISS expert survey, which will be described in detail in Chapter 4, is another rich source of information. In addition, scientific literature has been reviewed and used in this report.

The present SERISS Deliverable aims to cover the following aspects:

1. Advancing a generic definition of institutions and the institutionalized population that helps survey researchers to define this group for their interviewers
2. Offering a quantitative description of the institutionalized population in Europe and presenting variables that could be biased if the elderly institutionalized population is excluded from survey samples
3. Providing information on specific sampling frames in various (European) countries that could be used to sample institutionalized residents

---

1 Baromètres santé and the Barómetro Sanitario
4. Providing guidance on gatekeepers, their motives and ways to ensure their collaboration to gain access to a hard-to-reach population

5. Presenting potential survey tools that help to conduct fieldwork with hard-to-interview respondents living in institutions (and in private households)

6. Summarizing whether it is necessary (in the light of potential bias) and feasible (since tagged as hard-to-survey) for social survey to cover institutionalized residents in Europe

To achieve those purposes, Chapter 2 elaborates on the definition of institutions and the institutionalized population, which has been laid down in the SERISS Deliverable 2.16 (Schanze 2017). Chapter 3 contains a quantitative description of the institutionalized population and discusses the (self-)selection mechanisms that lead to the institutionalization of old-aged persons. In Chapter 4, we briefly describe the methods and contents of the SERISS expert survey, which relied on the SERISS Survey Inventory to identify relevant experts of surveys covering institutionalized respondents. The Chapters 5, 6, and 7 are the core chapters of this report and present survey tools that can help to sample, access and interview institutionalized residents. In Chapter 5, we extend the analysis of sampling practices of Deliverable 2.16. This chapter introduces country-specific sampling frames that were used to sample the institutionalized population, mentions potential coverage issues, and advances ways to avoid bias in gross samples. Chapter 6 discusses the issue of access to respondents, since surveys of the institutionalized population face gatekeepers more frequently than surveys of the population living in private households. The chapter describes reasons for certain behaviors of gatekeepers and delves into strategies of how to get access to the respondents. The survey cycle does not come to an end when a sample is drawn and respondents are recruited. Therefore, Chapter 7 deals with various aspects of fieldwork, such as cognitive tests, proxy interviews, and potential adaptations of fieldwork procedures to allow a collection of (hopefully valid) data in institutions. The Deliverable concludes with a discussion of the main questions whether it is necessary and feasible for social surveys to cover institutionalized populations in Europe.

---

2The underlying methodology of the SERISS Survey Inventory has been introduced in greater detail in Deliverable 2.16 (Schanze 2017)
2 A Top-Down Definition of Institutionalized Populations

A concise definition of the target population is crucial for surveys for several reasons. The definition might be guided by the research interest of the survey or by practical considerations. During the fieldwork, interviewers need to be aware of the definition to know which respondents they should interview and which residents should not be interviewed because they were defined as ineligible according to the survey guidelines. Primary and secondary users of survey data want to know whether they can answer a research question with the given survey data or whether they have to prepare the data in a certain way, e.g. by excluding certain (groups of) respondents. In this chapter, we offer a definition of the institutionalized populations summing up the definition given in SERISS Deliverable 2.16 and extending it in some ways. At the end of the chapter we will advance a number of practical key recommendations.

The institutionalized population is mainly characterized by its specific housing situation, since the residents have their permanent place of residence within a retirement home, prison, refugee accommodation, or any other type of institution. As a consequence, institutionalized residents are neither hidden, nor hard-to-identify like other hard-to-survey groups (see Tourangeau 2014). In SERISS Deliverable 2.16, we recommended a "top-down approach to derive a clear-cut and detailed definition of the institutionalized population and distinguish this subgroup from the population living in private households. In a first step, institutions are defined as centers of aggregation for the institutionalized population. In a second step, all persons who live in an institution, or, more precisely, have their usual place of residence within an institution, are described and classified as belonging to the institutionalized population." (Schanze 2017, 6). Figure 1 illustrates this top-down definition of the institutionalized population in two steps.

To start with the upper level of our definition, institutions, collective living quarters, group quarters, or communal establishments have certain features that clearly differ from private households. They are permanent structures owned and operated by a public agency or a private company. In contrast to private households, institutions "are designed for habitation by large groups of individuals or several households" (European Commission 2009, 63). Facilities, such as bathrooms, a kitchen and sometimes dormitories are usually shared by all the residents (United Nations 2008). The operators of institutions employ staff to run the institution, supervise it and offer the necessary services to the residents. With their services, the operators address a clearly defined target group (see OECD 2006), which also helps to identify different types of institutions. By running and organizing the institution, operators accomplish a specific common purpose of the society (education, care, organizing immigration, detention of criminals) or they serve a concrete private interest of their residents (Goffman 1957; Schanze 2017; United Nations 2008).

As second layer of the top-down definition, all residents living in the institutions permanently, or at least for a certain substantial period of time, belong to the institutionalized population. Surveys and the census need to define the period of time that is necessary to be classified as usual resident at a certain address. In the case of institutions, the temporal qualification aims to exclude short-term visitors, who will return to a private household rather sooner than later (European Commission 2009; Schanze 2017; United Nations 2008). For instance, this applies to most inpatients in hospitals, clients in hotels, refugees in initial receptions centers, or inhabitants of convalescent homes.

---

3E.g., 3 months, half a year, or 1 year
As mentioned above, institutions are suitable for fulfilling a very concrete demand by society or individuals. Some residents “choose to live in an institution because they want to be educated, nourished, receive medical attention, participate in a religious community, [whereas] another group of residents is forced to live in an institution, for instance prisoners, military staff, orphans, or refugees” (Schanze 2017, 7).\(^4\) Institutions “take over ‘responsibility’” for their inhabitants (Goffman 1957, 316). As a consequence, institutions exert a strong influence on the lives of their residents. Erving Goffman coined the term of “total institutions” to describe how a “single authority” prescribes all aspects of life of institutionalized populations.

\(^4\)Some of the factors that explain the selection mechanism for specific residents are described in the following section 3.
residents (1961, 314). Throughout this Deliverable, we will use the term "institutionalized populations" while referring to the entire "institutional population". The term underlines two characteristics of this group: We use the passive version of the adjective "institutional" to emphasize a certain loss of autonomy residents of institutions are facing after moving to an institution. The plural in the term "institutionalized populations" highlights the heterogeneity of the institutionalized population (see Chapter 3).

Comparing various types of institutions, the residents display a varying degree of dependence on the services offered by the institution. This continuum is not necessarily required to distinguish different types of institutions, since this can be already achieved in the upper part of Figure 1. Placing groups of institutionalized residents on the continuum might help researchers to distinguish related types of institutions (e.g., nursing homes, retirement homes, assisted living facilities). It might be also relevant to figure out the extent of adaptations they need to make in the course of fieldwork (see Chapter 7).

Some residents rely on the medical and custodial services provided by the institution (typically in nursing homes, hospitals and to a smaller extent in retirement homes), whereas other institutions do not offer those kinds of services (e.g., educational institutions, refugee accommodations, or workers dormitories).

In terms of life management and autonomy, some residents are not able (nursing homes, hospitals) or not allowed (prisoners or refugees) to move freely and leave the institution to change their usual place of residence. Many institutionalized residents get their meals prepared by the staff of the institution and are provided with further essentials for living (e.g., nursing homes, retirement homes, hospitals, prisons), while other institutionalized groups share a communal kitchen, prepare their own meals and buy all necessary essentials for living (e.g., residents in educational institutions, worker dormitories, or refugee accommodations).

In addition to the life management, some institutionalized residents are active members of the labor force. In the 2010 U.S. census, the distinction between "institutional group quarters" and "noninstitutional group quarters" was made by drawing on the labor force participation of the residents in noninstitutional group quarters, who are "primarily eligible, able, or likely to participate in the labor force while residents". (U.S. Census Bureau 2012, B-16). While residents of retirement and nursing homes are usually economically inactive, residents of student dorms, refugees and prisoners might earn money in a side job (labor force participation possible). Residents living in monasteries, military institutions and worker dormitories are definitely part of the labor force.

Obviously, survey guidelines cannot be this long. The instructions used by the Survey of Health, Ageing and Retirement in Europe (SHARE) are a good example of how to provide interviewers with guidance on how to identify institutions, specifically retirement and nursing homes in the case of SHARE. A nursing home "provides all of the following services for its residents: dispensing of medication, available 24-hour personal assistance and supervision (not necessarily a nurse), and room & meals". This brief definition contains elements from both levels of Figure 1 and clearly separates institutionalized residents from community-dwelling residents.

In Figure 1 the arrow of the dependency continuum indicates that true autonomy can only be achieved outside the institutions.

Recommendations:

- Institutionalized residents are part of their group because of their housing situation. Hence, institutionalized residents are usually not hidden or hard-to-identify. Using the suggested top-down definition helps to distinguish various types of institutions.

- Survey instructions should include brief checklists for interviewers that enable them to identify specific types of institutions. These checklists can be based on extracts from Figure 1.

- Even if a survey aims to exclude institutionalized residents from their target population, a precise definition of institutions helps interviewers to identify ineligible sample units. Instructions also help to secure the standardization of individual decisions made by interviewers.

- If your survey program will continue to survey only the community-dwelling population, you should state very clearly that the survey sample only allows conclusions for the population living in private households. Users should know and avoid to draw conclusions for the „entire population“ or the „French population older than 60 years“.
3 Which Groups Belong to the Institutionalized Population?

The following chapter provides a quantitative description of the institutionalized population. It mainly relies on the data collected during the 2011 European census and extends the brief quantitative description provided in the first Deliverable of SERISS Workpackage 2.5 (see Schanze 2017). In the light of different methods and preconditions in the census (UNECE 2009; Valente 2010), Eurostat applied an output-harmonization and asked the countries to provide various variables and 60 combinations of variables, so-called hypercubes (Eurostat 2011).

Eurostat published two variables that identify the institutionalized populations: As part of the housing arrangements, the residents in collective living quarters are reported, and a second variable on the household status captures when a person lives in an institutional household (Eurostat 2016). In 22 out of 32 European countries, the overall number of residents in the two categories is exactly the same or at least very close with respect to the absolute number and the relative share of residents.7 In Estonia, Germany, Greece, Iceland, Lithuania, the Netherlands, and Slovenia the numbers of persons living in collective living quarters and institutional households differ to a greater or lesser extent.8 The largest absolute difference occurs in Germany, where the two number deviate by nearly 600,000 residents.9 The largest relative difference can be observed in Slovenia, where the share of residents living in collective living quarters and the share of institutionalized residents differs by more than 1 percentage point relating to the entire population.

Apart from general problems of census data, such as restrictions linked to the timeliness of information and the amount of information available, Eurostat does not require the countries to provide information on the specific types of institutions the institutionalized population lives in. Hence, the census does not allow us to draw a proper cross-national comparison in terms of specific types of institutions. As proxies, we make use of age and gender to distinguish within the group of institutionalized residents.

3.1 Quantitative description of the institutionalized populations

According to the 2011 European census, about 1.33% of the European population does not live in a private household. Hence, the non-private population amounts to 6.75 million people. Within the non-private population the people living in institutions constitute by far the largest group.10 In the EU-28, the European census counted 6.6 million people living in collective living quarters and 5.7 million people living in institutions. Figure 2 was already presented in the first Deliverable and just works as a reminder, since it is important to be aware of the share of institutionalized residents in specific European countries.

---

7 No deviation at all, or deviation of shares smaller than 0.1 percentage points.
8 In addition, Ireland, Finland, and Switzerland did not report any numbers regarding the institutionalized residents.
9 Most probably, this deviation is due to the underlying definition the German Federal Statistical Office applies to the institutionalized populations. It defines some types of institutions as sensitive and other institutions as non-sensitive (Geiger and Styhler 2012). Following this definition, the amount of information collected during the census differs and consequently some hypercubes cannot be filled for all institutionalized residents.
10 In addition, some countries provide numbers on a third form of residence, neither private nor institutional (e.g., homeless shelters).
and cross-national differences. Two thirds of all European countries reported a share of collective households higher than 1%. A few countries (Bulgaria, Cyprus, Finland, Italy, Norway, Poland, and Spain) reported a very small share of collective households with less than 0.6% of the entire population. SERISS Deliverable 2.16 concluded that a geographical pattern was absent in Europe, since there was no clear "North-South or West-East divide" of countries (Schanze 2017, 9).

![Figure 2: Share of the population living in collective households in European countries (Eurostat 2016)](image)

A more distinct geographical pattern can be discovered when further variables, such as age and gender, are taken into account. Figure 3 shows the relative share of institutionalized residents in different age and gender cohorts. A majority of European countries follow the two patterns shown in Figure 3. A first group of countries depicts only one peak of the share of institutionalized residents within age and gender cohorts, while a second group of countries shows two peaks. Germany and Italy are two examples of a group of 15 countries with only one significant peak. In this group of countries, the institutionalized population is a minor group with a share of 2% at the maximum in all the younger age cohorts. The relative size starts to grow in the age cohort of people aged between 70 to 79 years and reaches its maximum in the oldest age cohort. Apart from Germany and Italy, Austria, Belgium, Croatia, Cyprus, Finland, Iceland, Ireland, Luxembourg, Malta, Norway, Portugal, Slovakia, and Spain fall also into this first group of countries with a similar pattern of institutionalized residents within different age and gender cohorts. France and the United Kingdom display a second significant peak in the younger age cohorts between 10 to 29 years. The Czech Republic, Denmark, Estonia, Hungary, Netherlands, Slovenia, and Sweden also belong to this group. It is safe enough to assume that the younger institutionalized age cohorts live predominantly in educational institutions like student dorms or boarding schools. Indeed, French census data confirm this assumption. In 2009, the mean age of residents in boarding schools was 18 years and in student dorms it was 22 years (Pirou and Poullain 2013). The average age in prisons was 34 years, while

---

11 Graph on EU-26: Bulgaria and Lithuania missing.
soldiers living in military barracks were 26 years old on average (ibid.). On the other end of the age scale, residents living in health-related institutions (50 years in average) religious institutions (66 years), or retirement homes (84 years) were significantly older in average than the general population (40 years, ibid.).

As the remaining group in Figure 3, Poland, Greece, Latvia, and Romania do not follow either of the two patterns but show a distinct pattern. In Greece and Romania, the share of institutionalized residents is the largest in the young age cohorts. More than 6% of the Greek men aged between 20 to 29 are institutionalized according to the census data. In Romania, nearly 4% of both genders are institutionalized in this age cohort. In contrast to all the other European countries, the elderly population lives predominantly...
in private households. In Latvia and Poland, a very small fraction of people live in institutions. The share of institutionalized residents does not differ across various age cohorts and the relative share of institutionalized residents grows only marginally in the oldest age cohort. Even more than for age, the gender distribution bears a great resemblance across the European countries. As it can be seen in Figure 3, a higher share of women than men lives in institutions in the oldest age cohort in all European countries. Also, in the population aged between 70 and 79 years, a higher fraction of women is institutionalized compared to the respective share among men. With a very small number of exceptions, this observation holds for all countries. The gender pattern is reversed for the younger age cohorts. In nearly all these age cohorts in most countries, a higher proportion among men is institutionalized compared to the equivalent share among women. However, the difference is not as distinct as for the older age cohorts and for the group of residents younger than 20 years the gender-related differences do not exist in many countries. Indeed, the French census in 2009 shows the differences in gender distributions across different types of institutions: Very few women lived in worker dormitories (15% of residents were female), military institutions (12% were female), and prisons (4% were female), whereas many residents in retirement homes (74%) and religious institutions (71%) were female (Pirou and Poulain 2013). In boarding schools, health-related institutions, and student dorms the gender distribution is more or less balanced (ibid.).

Following from the country plots in Figure 3 and further sources, age and gender are important variables in explaining institutionalization and especially the types of institutions the population lives in. These variables can serve as proxies for the distribution of institutionalized residents across various types of institutions, even if this kind of data were not collected and published by Eurostat at the European level in the last census. In some countries, the data can be retrieved from national sources. For the 2011 census in England and Wales, the British Office for National Statistics (ONS) published additional statistics on the population living in so-called communal establishments (ONS 2015). According to those numbers, 40.8% of the institutionalized population (382,000 English and Welsh residents) lived in educational institutions, 37.6% of all institutionalized residents lived in care homes with or without nursing (352,000 residents), 6.9% lived in prisons (64,000 residents), and 43,000 residents lived in military institutions (4.6%). In Sweden, the Statistical Office counted 178,000 residents living in "special housing" in 2015, adding up to 1.8% of the entire population. The numbers come from the population register and show an equal number or residents living in student dorms (47.9%) and institutions for elderly and disabled persons (47.7%). In France, the 2009 census estimated a share of 30.2% of institutionalized residents living in retirement homes (489,000 residents), 26.4% living in boarding schools, 18.8% living in health-related institutions, and 8.5% living in worker dormitories (Pirou and Poullain 2013). France, Sweden, and the UK all belong to the group of countries with two peaks according to Figure 3. As a consequence of the significant share of adolescents and students living in educational institutions, the institutionalized population in those countries is more heterogeneous than in other countries. Figure 4 shows the degree of diversity of the institutionalized population in all European countries. The graph

12 Exceptions can be observed in Estonia, France, Latvia, and Romania, where a higher share of men than women is institutionalized in this age cohort.
13 In the population younger than 10 years a very small fraction of residents is institutionalized, usually less than 0.5%.
14 Further substantial groups are residents in hotels (3.7%) and hospitals (1.9%).
SERISS Deliverable 2.17 - Institutionalized populations

provides the share of all institutionalized residents older than 65 years within the entire institutionalized population of the respective country. If we assume that these age cohorts live predominantly in retirement and nursing homes like in France, a higher share implies fewer residents living in other types of institutions, such as prisons or educational institutions. Figure 4 underlines the gender split in institutions that can be also observed in Figure 3. However, as Section 3.2 shows, a number of multivariate analyses revealed that a positive relation of female gender and institutionalization in old-aged residents is due to further underlying explanatory variables. Gender as such is not a significant explanatory variable in some of the multivariate models. Summing it up, Figures 3 and 4 allow the conclusion that institutionalization among men is spread across a larger number of different types of institutions. Men make up a large part of prisoners, workers in dormitories and military staff. In contrast, long-term care settings are more often inhabited by women.

At the European level, further information on the number of prisoners are available at the Eurostat homepage (Eurostat 2018). In the 28 EU member states, 0.12% (115 prisoners among 100,000 residents) of the total population lived in a prison in 2016 (see Figure 5). The three Baltic countries have the highest share of prisoners in Europe, a little more than 0.2% of the entire population lives in prisons in these countries (215 of 100,000 residents). In contrast, the neighbouring Scandinavian countries have the lowest share or prisoners with 60 prisoners among 100,000 residents. Comparing these numbers to global

15Please note, the bars cannot be compared in absolute terms between the countries, since they are all based on different denominators.
figures (see Walmsley 2016), they are far away from the number of residents in penal institutions in the USA (nearly 700 of 100,000 residents), Russia (455), or Turkey (255).

Comparing the numbers of 2016 with the numbers 6 years before, shows the development of the institutionalized population in prison. In 21 out of the 28 EU member countries, the share of prisoners has decreased, at the European aggregate the share of prisoners in 2016 amounts to 90.5% of the count in 2010. In Belgium, the Czech Republic, Romania, and Slovakia the share of prisoners remained the same.
whereas the share of prisoners increased in France (by 9.5%), Hungary (10.2%), and Portugal (20.2%). Regarding the gender distribution in prisons, Figure 6 confirms the above-mentioned observations regarding the clear prevalence of male prisoners. Again, gender is a highly relevant explanatory variable, as nearly all prisoners in Europe were male in 2016.

### 3.2 Selection mechanisms and the peril of bias

Whenever survey managers decide to leave out specific parts of the population, they assume that the excluded parts of the population would not change the survey estimates to a significant extent if they were included. Bias is the product of two elements: 1) the size of the excluded population and 2) its statistical distinctiveness (Groves et al. 2009). In the first part of this section, we described the relative size of the institutionalized population in the entire population (see Figure 2) and within specific age and gender cohorts (see Figure 3 and Figure 4).

In the following paragraphs, we will elaborate on the selection mechanisms that influence the institutionalization of old-aged community-dwelling residents. We focus on a number of studies that investigated the probabilities of living in institutions for the elderly or moving to these types of institutions. This focus is justified with the quantitative importance of the elderly population living in private households and institutions in Europe. The studies cited in Table 1 and Table 2 identified a number of variables that help explain institutionalization. Many of the cited surveys relied on survey data, analyzing data from SHARE (Angelini and Laferrère 2012; Geerts and Bosch 2012; Laferrère et al. 2012; Rodríguez-Sánchez et al. 2017), the U.S. HRS (Castora-Binkley et al. 2014; Noël-Miller 2010; Thomeer et al. 2015), the Canadian NPHS (Asakawa et al. 2009), the Swedish H70 survey (Bravell et al. 2009), the French HID (Désesquelles and Brouard 2003), U.S. surveys EPESE (Hays et al. 2003) and AHEAD (Kasper et al. 2010), the Dutch LASA (Pot et al. 2009), or the German LEILA75+ survey (Riedel-Heller et al. 2000). In some countries, the authors analyzed the population registers (Einiö et al. 2012; Martikainen et al. 2009) or large-scale administrative surveys (Grundy and Jitlal 2007; McCann et al. 2012; Nihtilä et al. 2008). Other authors conducted a meta-analysis (Gaugler et al. 2007; Toot et al. 2017) or a systematic review (Luppa et al. 2010).

Andersen developed a model that structures different explanatory factors for institutionalization (1995). The model comprises predisposing factors (demography and economic factors), need factors (health and demanding care), and enabling resources (availability of informal and public care). Table 1 and Table 2 list variables that have been identified as significant explanatory variables in multivariate models in the above-mentioned studies. A coincidence of several socio-demographic and health-related factors might be the most powerful model in explaining institutionalization of elderly. With an increasing age, the likelihood of frailty rises and people face more difficulties with all kinds of activities of daily living, their medical condition (especially dementia), and cognitive functioning (see Table 2). The availability of an informal caregiver enables people to continue living in their private households. In this regard, the quality and size of the social network plays a crucial role, as well as having children living nearby or living with a spouse.

---

16For instance, dressing, walking, bathing, eating, or getting in or out of bed.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Studies finding a significant positive impact in multivariate models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Lower: Asakawa et al. 2009; Einiö et al. 2012; Castora-Binkley et al. 2014; Thomeer et al. 2015</td>
</tr>
<tr>
<td>Housing (not home owner)</td>
<td>Einiö et al. 2012; Gaugler et al. 2007; Grundy and Jitlal 2007; Luppa et al. 2010; Martikainen et al. 2009; McCann et al. 2012; Thomeer et al. 2015</td>
</tr>
<tr>
<td>Living alone (without partner or children)</td>
<td>Angelini and Laferrère 2012; Désesquelles and Brouard 2003; Gaugler et al. 2007; Grundy and Jitlal 2007; Kasper et al. 2010; Laferrère et al. 2012; Martikainen et al. 2009; McCann et al. 2012; Rodríguez-Sánchez et al. 2017; Thomeer et al. 2015</td>
</tr>
<tr>
<td>Childlessness (incl. grandchildren)</td>
<td>Bravell et al. 2009; Hays et al. 2003; Kasper et al. 2010; Laferrère et al. 2012; Noël-Miller 2010; Rodríguez-Sánchez et al. 2017</td>
</tr>
<tr>
<td>Small social networks, informal caregivers not available</td>
<td>Hays et al. 2003; Laferrère et al. 2012; Luppa et al. 2010; Maxwell et al. 2013; Noël-Miller 2010; Thomeer et al. 2015</td>
</tr>
</tbody>
</table>

Table 1: List of socio-demographic variables that have a positive statistical impact on the likelihood of living in institutions for the elderly (Schanze and Zins 2019)
In contrast, people who live alone face a greater risk of institutionalization. In addition, socio-economic variables such as education, house ownership, or income also explained institutionalization in a number of studies. With the exception of house ownership, the results for those variables are mixed across studies. Some studies found that a rising level of education and income increase the probability of institutionalization, whereas other studies find the opposite effect. These contradictory results could be partially explained with national context factors.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Studies finding a significant positive impact in multivariate models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive impairment</td>
<td>Agüero-Torres et al. 2001; Castora-Binkley et al. 2014</td>
</tr>
<tr>
<td></td>
<td>Einiö et al. 2012; Gaugler et al. 2007; Hays et al. 2003</td>
</tr>
<tr>
<td></td>
<td>Kasper et al. 2010; Luppa et al. 2010; Maxwell et al. 2013</td>
</tr>
<tr>
<td></td>
<td>Nihtilä et al. 2008; Nöel-Miller 2010; Rodriguez-Sánchez et al. 2017</td>
</tr>
<tr>
<td></td>
<td>Thomeer et al. 2015; Toot et al. 2017</td>
</tr>
<tr>
<td></td>
<td>Luppa et al. 2010; Nihtilä et al. 2008; Riedel-Heller et al. 2000</td>
</tr>
<tr>
<td></td>
<td>Toot et al. 2017</td>
</tr>
<tr>
<td></td>
<td>Agüero-Torres et al. 2001; Angelini and Laferrière 2012</td>
</tr>
<tr>
<td></td>
<td>Einiö et al. 2012; Gaugler et al. 2007; Grundy and Jitlal 2007</td>
</tr>
<tr>
<td></td>
<td>Hays et al. 2003; Luppa et al. 2010; Maxwell et al. 2013</td>
</tr>
<tr>
<td></td>
<td>Nihtilä et al. 2008; Rodriguez-Sánchez et al. 2017; Toot et al. 2017</td>
</tr>
<tr>
<td>Medical conditions</td>
<td>Bravell et al. 2009; Gaugler et al. 2007; Hays et al. 2003</td>
</tr>
<tr>
<td></td>
<td>Kasper et al. 2010; Laferrière et al. 2012; Nöel-Miller 2010</td>
</tr>
<tr>
<td></td>
<td>Rodriguez-Sánchez et al. 2017; Thomeer et al. 2015; Toot et al. 2017</td>
</tr>
<tr>
<td>Limitations in activities of daily living (ADL)</td>
<td>Castora-Binkley et al. 2014; Laferrière et al. 2012; Nöel-Miller 2010</td>
</tr>
<tr>
<td></td>
<td>Thomeer et al. 2015; Toot et al. 2017</td>
</tr>
<tr>
<td>Limitations in instrumental activities of daily living (IADL)</td>
<td>Thomeer et al. 2015; Toot et al. 2017</td>
</tr>
<tr>
<td>Physical dependency and mobility difficulties</td>
<td>Angelini and Laferrière 2012; Désesquelles and Brouard 2003</td>
</tr>
<tr>
<td></td>
<td>Hays et al. 2003; Thomeer et al. 2015; Toot et al. 2017</td>
</tr>
<tr>
<td>Self-rated health</td>
<td>Castora-Binkley et al. 2014; Einiö et al. 2012; Luppa et al. 2010</td>
</tr>
<tr>
<td></td>
<td>McCann et al. 2012; Nöel-Miller 2010</td>
</tr>
</tbody>
</table>

Table 2: List of health-related variables that have a positive statistical impact on the likelihood of living in institutions for the elderly (Schanze and Zins 2019)

In a revision of his model, Andersen added the health care system as explanatory factor to it (1995). The long-term care system and insurance schemes differ across countries and could influence the distribution of socio-economic variables among institutionalized residents.\textsuperscript{17} Laferrière and colleagues found signifi-

\textsuperscript{17}As most of the cited studies on institutionalization are restricted to one country, they cannot analyze macro-level variables in a multilevel approach.
cant negative effects of means-tested payment schemes and comparably large co-payments of the overall costs for care on the likelihood of institutionalization in a cross-national comparison (2012). Moreover, the national culture and predominant family values can also influence institutionalization. Geerts and Bosch aimed to operationalize the long-term care culture of various European countries (2012). However, they could only partially confirm the hypotheses that transition to informal care at home is higher in more familialistic countries whereas transition to formal care is more frequent in countries with needs-based entitlements (ibid.).

To sum it up, all the above-mentioned factors that help to explain institutionalization also determine the statistical distinctiveness of the institutionalized population compared to the population living in private households. As a consequence, those factors might be biased in surveys of private households.

**Recommendations:**

- Bias in survey estimates is caused by two factors: (1) The size of the excluded (or insufficiently covered) parts of the population and (2) its statistical distinctiveness (see Tables 1 and 2). Check carefully whether your target population and the topics your survey covers might suffer from bias due to these two factors.

- Take into account that the bias can affect different groups within your target population to a different extent. In most European countries, a larger share of the elderly population lives in institutions. In some countries also a significant fraction of young people is institutionalized.

- Census data are the first source to estimate bias in some socio-demographic variables. Eurostat offers an easy-to-handle tool to download data on the institutionalized populations (https://ec.europa.eu/CensusHub2) in nearly all European countries. Data of other surveys might help to estimate potential bias in additional variables (check the SERISS Survey Inventory to identify these surveys). You can also run a (semi-)synthetic simulation to quantify the potential bias (e.g., Schanze and Zins 2019).
4 The SERISS Expert Survey

In 2008, researchers working on the European Health Interview Survey (EHIS) decided to set up a Task Force which aimed to investigate whether institutionalized residents could take part in the survey. As part of their work, the Task Force surveyed participating countries to learn more about the sampling frames they used and the experiences they made (Beukenhorst et al. 2011). The SERISS project followed a similar approach. A questionnaire was sent to survey researcher who worked on a recent survey program that included institutionalized residents.

Relevant survey programs were identified with the SERISS Survey Inventory, which will be briefly presented in the following section. In addition, this Chapter describes the contents of the expert questionnaire and depicts the response rate and the sample composition.

4.1 The SERISS Survey Inventory

The SERISS Survey Inventory and its underlying methodology was introduced in Deliverable 2.16 (Schanze 2017). The inventory currently covers nearly 300 surveys conducted in roughly 30 European countries and in Australia, Canada, Israel, the USA, and Russia. Half of the surveys in the inventory (N = 153) covered institutionalized residents. 80 of those surveys are cross-sectional and 73 are panel surveys, 46 of them only covering institutionalized residents and the remaining 107 surveys combining institutions and private households. The overwhelming majority of surveys in institutions was conducted at a national (or regional) level, only 7 cross-national survey programs were found.

Many surveys interviewed residents of retirement and nursing homes (48.4% of all relevant survey programs), fewer surveys covered prisons (13.7%), or refugee accommodations (8.5%). Student dorms, homeless shelters, or military barracks were addressed by less than a handful of surveys each. In addition, 26 survey programs (17%) aimed to cover the entire institutionalized population without a further differentiation of the target population in institutions.

4.2 Data collection

On the basis of the SERISS Survey Inventory, we compiled a list of survey programs that included institutionalized respondents. We decided to cast a wide net and attempted to contact persons for all the surveys that have been conducted after 2005. In total, we contacted 140 survey managers working in 102 different surveys in 30 countries. For the few cross-national survey programs, notably SHARE, ISSP, LFS and EHIS, we contacted the national survey teams separately because we were interested in the country-specific experiences with data collections in institutions. The fieldwork period of our expert survey started in late November 2017 and lasted until end of January 2018. In these two months, two reminders were sent to all the experts who had not reacted to our invitation until these points of time.

---

Please note, this observation does not imply to be representative for the universe of surveys, because the inventory of surveys is not a random sample of surveys or a census of surveys.

As of November 2017

---
In the course of data collection, five survey managers informed us that they were not eligible for the expert survey, since their survey programs did not cover any institutionalized population. In addition, three survey managers told us that their surveys included institutionalized respondents, but they could not identify these cases in the data and during the data collection. As a consequence, they could not provide us with detailed information about this group of respondents. At the end of data collection, 44 out of the 140 contacted survey managers had completed the survey, adding up to a response rate of 32.6% of all the eligible surveys.

4.3 Questionnaire

The questionnaire was designed by the SERISS workpackage team in 2017 and covers several aspects of sampling of the institutionalized population and fieldwork in institutions in survey-based research. The questionnaire is comprised of three parts, dealing with the general survey design, sampling, and fieldwork.

The first part of the questionnaire dealt with general features of the surveys and questions about the institutionalized target population that was covered by the survey program. Respondents were given several groups and could tick multiple of those groups and add other groups in addition. If a respondent selected more than one institutionalized group, s/he was given a follow-up question that asked for the largest group in quantitative terms. We wanted to figure out the relative importance of institutionalized groups and make sure that the following questionnaire was related to a single institutionalized group instead of confounding the answers with multiple institutionalized groups. In addition, the first part of the questionnaire asked about the general survey design and the inclusion rules applied by the survey programs for institutionalized respondents. Taking into account the diversity of survey programs in the Survey Inventory, the questionnaire was intended to apply to a wide range of different survey designs. The first part of the expert questionnaire and its filter questions helped us to achieve this purpose.

The second part of the expert survey covered the sampling design. Respondents answered questions about the main sampling frame they used and whether this sampling frame allowed them to identify institutionalized residents prior to the data collection. We also asked the respondents to evaluate potential undercoverage or overcoverage of institutionalized residents in their main sampling frames and followed up with a question about reasons of clear coverage error. Questions on oversampling and techniques of oversampling of institutionalized residents were only given to survey managers of surveys that covered institutionalized residents and community-dwelling households. The second part ended with an open question on recommendations the respondents might want to give to other researchers who want to sample institutionalized residents in the respective country.

The third part of the expert survey dealt with questions on gatekeepers and fieldwork. We asked about specific gatekeepers that had to be addressed while gaining access to the institutionalized respondents and a normative evaluation of this process. We also asked about proxy interviews and whether there was a systematic protocol to determine if the survey had to be taken by a proxy respondent. Apart from question on various aspects of data collection in institutions, such as translation, adaptations of
questionnaires, informed consent and specific pretest, the third part also examined the general expert assessment with respect to the ability of institutionalized residents to take part in a survey interview. We asked the respondents to estimate the fraction of institutionalized residents that could not be interviewed and wanted to know which mode of data collection they would recommend for this population. Moreover, we asked them to rate whether institutionalized respondents were hard-to-sample, hard-to-identify, hard-to-reach, hard-to-persuade, and hard-to-interview. Like in the second part, the third part also ended with an open question, which invited respondents to provide us with additional insights and expertise. All questions included a "Don’t know" option.

Before starting fieldwork, an informal pretest of the contents and the technical implementation was conducted with survey researchers and experts of GESIS - Leibniz Institute for the Social Science, the Hebrew University of Jerusalem and the workpackage leader at that time, Annette Scherpenzeel.

4.4 Sample composition

The 44 survey programs for which we collected information in the SERISS expert survey were conducted in 21 different countries. As depicted in Figure 7, all major types of survey designs are covered in the survey, including 20 cross-sectional survey programs and 24 panel surveys.

Regarding the types of institutions covered by the survey programs, a significant proportion of experts indicated that their survey program covered more than one institutionalized group (38.6%). One third of those 17 cases covered health-related facilities alongside retirement and nursing homes. The remaining two thirds (N = 11) of those experts selected multiple types of institutions without any specific direction. For instance, one survey program covered boarding schools, health-related facilities, retirement and nursing homes, military barracks, prisons and religious institutions according to the expert. We labeled this and other cases as "Institutions generally included" in Figure 8. Taking into account the heterogeneity of the entire institutionalized population, we asked the experts to indicate the largest group if they chose more than one institutionalized group. We also asked them to refer the following questions in the questionnaire to this type of institution. As it can be expected from Chapter 3, a majority of experts named retirement and nursing homes as the largest group in their sample (76.5%).

Figure 8 confirms the importance of retirement and nursing homes. A majority of surveys covered in the SERISS expert survey primarily focused on this type of institutions. In total, 70.5% of the survey managers covered it, in some cases alongside other institutions. In addition to "catch-all" surveys (N = 11), the expert survey also contains survey programs that focused on other types of institutions, such as refugee accommodations, boarding schools, or health-related institutions.

Even though we managed to get completed interviews from a wide range of different survey designs in different countries and different types of institutions, it is important to mention that the results cannot be generalized and only provide us with qualitative insights. When analyzing the survey results, we will only refer to the sample of completed interviews and take those as examples for potential data collections in institutions.
Figure 7: Types of surveys conducted by the survey managers responding to the SERISS expert survey

Figure 8: Types of institutions covered by the survey managers responding to the SERISS expert survey
5 Sampling the Institutionalized Population

The following chapter names a number of sampling frames and sampling procedures that were used by survey programs to select a sample of institutionalized residents. A few questions should be answered before developing a sampling strategy for a hard-to-sample population (Kalton 2009). In Chapter 3, we already described the quantitative and relative size of the target population. In this chapter, we elaborate on coverage issues that might lead to an inappropriate representation of institutionalized residents and possibilities to identify this target population in various sampling frames. Identifiers can be used to oversample institutionalized residents in order to counterbalance potential undercoverage. We also discuss the availability of other surveys that could be used in a piggy-back approach and briefly suggest weighting as a tool to reduce coverage bias. This chapter relies on the SERISS expert survey, literature and publications of relevant survey programs.

Nearly all of the survey programs included in the SERISS Survey Inventory used probability-based sampling methods. Only one survey was identified as using quota sampling to recruit respondents, namely a small qualitative survey among refugees conducted by the German Federal Office for Migration and Refugees (BAMF-RES). Thus, the following chapter only focuses on probability-based sampling methods.

5.1 Sampling frames

Table 3 lists 28 sampling frames in 18 countries used by the survey programs included in our expert survey. Nearly two thirds of the sampling frames listed in Table 3 were classified as general population registers by the survey experts (see Section 5.1.1). Only a small number of 6 surveys sampled institutionalized residents without any residents living in private households. Those surveys did not use population registers but relied on more specialized sampling frames, namely health-related registers (see Section 5.1.2) or other lists of institutions (see Section 5.1.3). Obviously, the given list cannot be exhaustive. Potentially, there are additional or new sampling frames available in your country, well-suited to draw a probability sample of institutionalized residents. The SERISS Survey Inventory could help to find additional sampling frames in your field of research and country.

5.1.1 Population registers

The majority of survey experts who answered our expert survey classified their sampling frames as general population registers. Whenever those registers are available for survey sampling (see Scherpenzeel et al. 2016), they are usually the first option as sampling frame and frequently used if available. In 6 cases the survey experts could not name the sampling frame used in the survey program (missing value). For instance, 11 of the sampling frames listed in Table 3 were also used by the European Social Survey (ESS).
addition, general population registers offer an easy way to sample institutionalized residents alongside the population living in private households, as long as the register does not suffer from coverage issues (see Table 4).

The cross-national *Survey of Health, Ageing, and Retirement in Europe* (SHARE) aims to cover residents in private households and institutionalized residents living in retirement and nursing homes. Each country used a separate sampling frame and should identify the most suitable sampling frame relying on the "basic principles of probability sampling with minimal coverage errors" (De Luca et al. 2015, 77). Since SHARE covers the population older than 50 years, additional information on the age of residents are also required (Lynn et al. 2013), at least if the countries do not want to screen a sample without age identifier.

In total, 7 SHARE country teams participated in the SERISS expert survey (Estonia, Germany, Israel, Netherlands, Poland, Portugal, and Switzerland). 5 of those countries reported the usage of a general population register to draw their sample, while the Israeli team used a telephone register and the Portuguese team used a health insurance register (Sistema Nacional de Saúde). In the case of Portugal, this can be explained with the absence of a population register in the country (Scherpenzeel et al. 2016).

The Swiss SHARE team used the *Stichprobenrahmen für Personen- und Haushaltsbefragungen* (SRPH, Sampling frame for surveys of individuals and households) as sampling frame for their refreshment sample. The SRPH is based on regional registers and available since 2010 for scientific purposes. In the baseline wave in 2004, SHARE used a telephone register (Klevmarken et al. 2005, 66). According to the sampling design forms, the telephone register also included institutions, which were excluded by the Swiss team at that time because they were not defined as elements from the target population (ibid.). In contrast, as the survey expert pointed out, the SPRH clearly suffers from undercoverage of institutionalized residents, because they are excluded from the frame by definition. The Swiss National Statistical Institute claims that "information on the collective households are not sufficient to be considered in the sampling design" (Bundesamt für Statistik 2018, own translation).

The German population register, (*Einwohnermeldeamtsregister*) is frequently used as sampling frame in social surveys. It was also used by five survey programs included in our expert survey. Those experts named the German register as a sampling frame with slight or even clear undercoverage of the institutionalized populations (see Table 4). In 2013, the Federal Registration Law (*Bundesmeldegesetz*) took effect in Germany. The newly introduced law included a conditional blocking notice for those addresses in the register that were defined as worth of being protected (par. 52). Sensitive addresses included (1) prisons, (2) refugee accommodations, (3) hospitals, nursing homes or further institutions taking care of care-dependent or disabled persons, (4) institutions for people affected by domestic violence, and (5) institutions to treat addicted people.

This law created uncertainty among survey researchers working on surveys that aimed to cover institutionalized residents. The German register is organized in a decentralized way and it is unclear whether local authorities apply the law only to new register entries, whether they were doing own research to find...
out about potentially sensitive addresses and whether they were allowed to ignore the blocking notice when publicly financed studies asked for a random sample of residents within their municipality. As the expert working on the NRW80+ survey pointed out, researchers should also check whether their "institution or service provider are [...] eligible to request data from persons in need of care/in LTC facilities". The IMOA project (Improving Health Monitoring in Old Age) tested a sampling design in six different sampling points by drawing 8,000 addresses from those sampling points. Only in one sampling point the local authorities did not include any addresses blocked by the conditional blocking notice.

In addition, the German SHARE team mentioned another reason for undercoverage in the German population register. According to them, persons are very often still registered in their previous address, especially if a spouse or partner still lives at this address. At least respondents are registered in those cases, even though interviewers need to make an additional effort to follow a selected person if the new address is traceable. The Israeli SHARE survey expert reported a similar issue with the population register in Israel. Many residents do not update their addresses once they moved, and many respondents cannot be tracked even after additional efforts. This is the reason why the survey expert recommends to move away from the population register as sampling frame and use „a combination of the geographical list and a screening process“ instead.

The IAB-BAMF-SOEP Refugee Panel was the baseline wave of a panel survey and aimed to interview refugees who arrived in Germany in a given three year period irrespective of the state of their asylum (Brücker et al. 2016). For this target population the researchers could not rely on the German population register, because it does not contain those refugees who are not yet granted asylum and entitled to stay in Germany. Owing to a collaboration with the Federal Office for Migration and Refugees (BAMF) the researchers could make use of the German Central Register of Foreigners (Ausländerzentralregister, AZR) as sampling frame. The AZR lists all foreigners without a German citizenship, who live in Germany, except foreign armed forces (Babka von Gostomski and Pupeter 2008). The addresses are not included in the register but need to be added by local foreigners’ registration offices once a sample of respondents has been drawn (ibid.).

In the SERISS expert survey the researcher working on the German Refugee Panel pointed out the increased mobility of the population living in refugee accommodations. Refugees usually leave their first accommodation very fast, after six weeks to six months according to the laws (Asylgesetz). Thus, the time passing between the sampling of respondents and the first contact attempt should be as short as possible. Moreover, German authorities struggled to keep track of the arriving refugees and their registration, resulting in undercoverage in the AZR (Kühne et al. 2019). To account for the very mobile population and delays in updates of the AZR, respondents were sampled in four distinct tranches between January and June 2016 (Brücker et al. 2016; Kühne et al. 2019). By this technique the researchers hoped to lower the non-contact rate of respondents (ibid.).

In addition to the Swiss and German registers, Table 3 lists population registers from 10 different countries. When asking the survey experts for concrete reasons for clear undercoverage or overcoverage
<table>
<thead>
<tr>
<th>Name of sampling frame</th>
<th>Type of sampling frame</th>
<th>Country</th>
<th>Abbreviation of surveys</th>
<th>Type of institutions covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account data from 12 health insurance companies</td>
<td>Health insurance register</td>
<td>AT</td>
<td>PaBe</td>
<td>Health-related facilities</td>
</tr>
<tr>
<td>Arnona files - Property tax files</td>
<td>Population or civil register</td>
<td>IL</td>
<td>IL_EIS</td>
<td>Educational institutions</td>
</tr>
<tr>
<td>Ausländerzentralregister (AZR)</td>
<td>Population or civil register</td>
<td>DE</td>
<td>IAB-BAMF-SOEP</td>
<td>Refugee accommodations</td>
</tr>
<tr>
<td>Australian Standard Geographical Classification</td>
<td>Geographical listing/database</td>
<td>AU</td>
<td>HILDA</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>Basic Resident Register</td>
<td>Population or civil register</td>
<td>JP</td>
<td>ISSP</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>Contrôle des habitants</td>
<td>Population or civil register</td>
<td>CH</td>
<td>VLV</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Einwohnermeldeamtregister</td>
<td>Population or civil register</td>
<td>DE</td>
<td>IMOA ReGES</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Fichier national des établissements sanitaires et sociaux (Finess)</td>
<td>List of institutions</td>
<td>FR</td>
<td>ANAIS CARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Gemeentelijke Basis Administratie (GBA)</td>
<td>Population or civil register</td>
<td>NL</td>
<td>SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>General practitioner registration lists</td>
<td>Lists of patients</td>
<td>UK</td>
<td>CC75C</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Icelandic National Register</td>
<td>Population or civil register</td>
<td>IS</td>
<td>ISSP</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>List of police custody suites</td>
<td>List of institutions</td>
<td>UK</td>
<td>AS</td>
<td>Prisons</td>
</tr>
<tr>
<td>Local administrative register of Abbiategrasso (Anagrafe)</td>
<td>Population or civil register</td>
<td>IT</td>
<td>InveCe-Ab</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>National Register</td>
<td>Population or civil register</td>
<td>BE</td>
<td>EGIS</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>Population Registry</td>
<td>Population or civil register</td>
<td>IL</td>
<td>IL_CVS</td>
<td>Educational institutions</td>
</tr>
<tr>
<td>Population and housing census (2011) and the Population Register</td>
<td>Population or civil register</td>
<td>EE</td>
<td>LFS</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>Population census, List of communal establishments</td>
<td>List of institutions</td>
<td>IE</td>
<td>NDS</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Population of the Västermalm parish in Stockholm</td>
<td>Population or civil register</td>
<td>SE</td>
<td>SNAC-K</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Postcode address file (PAF)</td>
<td>Geographical listing/database</td>
<td>UK</td>
<td>LFS UKHLS</td>
<td>Educational institutions</td>
</tr>
<tr>
<td>Rahvastikuregister</td>
<td>Population or civil register</td>
<td>EE</td>
<td>EGIS SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Register der Bundesagentur für Arbeit (BA)</td>
<td>Population or civil register</td>
<td>DE</td>
<td>PASS</td>
<td>Refugee accommodations</td>
</tr>
<tr>
<td>Register of nursing and retirement homes (Zorgkaart Nederland)</td>
<td>List of institutions</td>
<td>NL</td>
<td>OII</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Registret över totalbefolkningen (RTB)</td>
<td>Population or civil register</td>
<td>SE</td>
<td>LFS SWEOLD ULF</td>
<td>Institutions generally included</td>
</tr>
<tr>
<td>Sistema Nacional de Saúde (SNS)</td>
<td>Health insurance register</td>
<td>PT</td>
<td>SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Stichprobenrahmen für Personen- und Haushaltsbefragung (SRPH)</td>
<td>Population or civil register</td>
<td>CH</td>
<td>SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Telephone directory database</td>
<td>Telephone register</td>
<td>IL</td>
<td>SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Universal Electronic System of Population Register Number (PESEL)</td>
<td>Population or civil register</td>
<td>PL</td>
<td>SHARE</td>
<td>Long-term care facilities</td>
</tr>
<tr>
<td>Väestörekisteri</td>
<td>Population or civil register</td>
<td>FI</td>
<td>T2011</td>
<td>Institutions generally included</td>
</tr>
</tbody>
</table>
as indicated in Table 4, we yielded rather mixed results. Probably, this can be explained with the absence of definite knowledge on the coverage of this small part of the population. This is also reflected in some contradicting results for registers across experts. However, the table and related assessments of coverage quality might be a valuable starting point for survey researchers who aim to use a register as sampling frame for institutionalized residents.

5.1.2 Health insurance registers and lists of patients

Health insurance registers or list of patients can be very useful for the purpose of sampling institutionalized respondents with or without additional community-dwelling residents. If those lists can be accessed for sampling purposes, they contain a number of auxiliary variables, such as age, gender and health status, which can be used for stratified sampling approaches. For instance, Fink analyzed routine data by a large private German health insurance to learn more about the transition to long-term care housing after a diagnosis of dementia (2014). Potentially this register could also be used to draw a probability sample of institutionalized (and community-dwelling) residents.

Health-related registers were used as sampling frames in a number of surveys included in the SERISS expert survey. Researchers working on the the Austrian Patientenbefragung (PaBe) aimed to interview patients and used data from 12 insurance companies to identify who belonged to their target population (people staying in in-patient care facilities for at least 2 nights aged 14 and older). The Portuguese SHARE team also labeled their sampling frame (Sistema Nacional de Saude) as health insurance register. According to the information in the SERISS expert survey, the first wave in Portugal in 2011 did not include institutionalized residents. However, according to the Sampling Design Form, the National Health System does include institutionalized residents (Lynn et al. 2013, 111). For 6 percent of the units in the sampling frame addresses are missing, moreover, inaccuracies occur when citizens do not update their addresses once they moved. In the SERISS expert survey, the Portuguese expert recommended to consult with the Instituto da Segurança Social, since they have data about all social and health care institutions.

As another example, the U.S. National Health and Aging Trends Study (NHATS) aimed to survey the population older than 65 years and used the Medicare enrollment database as sampling frame in the baseline wave in 2010 (see Montaquila et al. 2012). In a three stage sampling design researchers selected 95 counties as primary sampling units (PSUs) and 655 ZIP codes as secondary sampling units (SSUs) within the PSUs. Within the SSUs, nearly 15,000 Medicare beneficiaries were sampled. In their report the researchers also provided an estimate of the coverage rate of the sampling frame 2010 compared to the 2010 census. According to Montaquila and colleagues 96% of the 40.3 million people older than 65 years were included in the sampling frame (2012). However, information on whether those 1.5 million people are missing randomly or systematically in the Medicare enrollment database are not available.

The Cardiovascular Health Study (CHS) and the National Long-Term Care Survey (NLTCS) are two other examples for U.S. surveys who already used the Medicare enrollment database in the 1980ies to

---

27 They were only included if a baseline respondent moved to a retirement or nursing home between two waves.

28 For instance, because they want to stay with their previous GP or medical centre.
draw a sample of respondents. In England, the public National Health Services asks so-called Councils with Adult Social Services Responsibilities (CASSRs) to run a postal survey among users of health services, including care homes and community-dwelling care. The NHS provides the councils who take part in the Adult Social Care Survey (ASCS) with an extensive guidance document (NHS Health and Social Care Information Centre 2017). Apparently, the councils hold lists of care home users and are asked to separate users into four different strata, one of them covering long-term residential or nursing care users older than 65 years. The NHS also provides a sample size calculator and suggests drawing a reserve sample to account for a larger number of ineligible residents (e.g., because of cognitive impairments) or a higher nonresponse. Potentially, lists of care home users held by the CASSRs could be also used by third parties for the purpose of drawing a random probability-based sample of institutionalized residents in England.

The Cambridge Project for Later Life (CC75C) used patient lists of local general practitioners (GP) to draw their sample. According to the expert, the compilation of these lists have a very good coverage in most geographical areas comparable to a population register. These lists do not suffer from any coverage issues and allow a direct identification of institutionalized residents. In the reply to the open question, the survey expert reported that the option to sample respondents through GPs was still available. It might be a valuable option in regional studies, but potentially more difficult to implement at a national level.

The SERISS Survey Inventory lists further survey programs that used medical lists or insurance lists as sampling frames. The German Study on Ageing, Cognition and Dementia in Primary Care Patients (AgeCoDe) aimed to examine dementia in a group of elderly people older than 75 years over the course of 6 years (Luck et al. 2010). For this specific target population, the researchers relied on the cooperation and expertise of general practitioners. They contacted a number of GPs in 6 urban centers in Germany and between 19 to 26 GPs participated in each of the centers. From the GPs patient lists, the researchers drew more than 6,600 residents, who did not have dementia during the first wave of data collection, did not live in a nursing home, were not fatally ill or audio-visually impaired. More than 3,000 respondents participated in the baseline wave. Those respondents who moved to a nursing home and developed dementia were also included in the analysis (77 out of 256 cases).

A sampling design on the basis of GP patient lists is very often restricted to certain regions. The Age-CoDe study was conducted in 6 cities, the Newcastle 85+ Study (N85+) drew a cohort of old-aged community-dwelling and institutionalized residents living in Newcastle and North Tyneside from lists of patients of 53 GPs, and the Melton Osteoporotic Fracture Study (MoF) used the register of Melton Mowbray’s single GP as sampling frame in the early 1990ies.

In 2006, the LifeLines Cohort Study also made use of GP lists of patients to draw a baseline sample of respondents and combined this sample with a network sampling approach. In the first step, 333,307 inhabitants between the ages of 25 and 50 years were invited via their general practitioners in three Northern regions in the Netherlands (Scholtens et al. 2015). Nearly 25% of the invited participants completed the baseline questionnaire and were asked whether any of their family members would be also willing to participate. The network sampling approach resulted in 64,500 respondents.

---

29 73% of the 812 contacted GPs agreed to forward the invitation to eligible patients.
30 The group of respondents invited via a family member showed a higher response rate of 70.2%.
Table 4: To what extent does the sampling frame used suffer from coverage issues? (Source: SERISS Expert Survey 2018, stars indicating contradicting categories used for a single register)

<table>
<thead>
<tr>
<th>Coverage Issue</th>
<th>Country</th>
<th>Sampling Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear undercoverage</td>
<td>IL</td>
<td>Arnona files</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Einwohnermeldeamt (2 times) *</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>Postcode adress file (PAF) *</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>Rahvastikuregister</td>
</tr>
<tr>
<td></td>
<td>CH</td>
<td>Stichprobenrahmen für Personen- und Haushaltsbefragung (SRPH)</td>
</tr>
<tr>
<td>Slight undercoverage</td>
<td>DE</td>
<td>Ausländerzentralregister (AZR)</td>
</tr>
<tr>
<td></td>
<td>AU</td>
<td>Australian Standard Geographical Classification</td>
</tr>
<tr>
<td></td>
<td>JP</td>
<td>Basic Resident Register</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Einwohnermeldeamt (3 times) *</td>
</tr>
<tr>
<td>No coverage issues</td>
<td>BE</td>
<td>National Register</td>
</tr>
<tr>
<td></td>
<td>NL</td>
<td>Register of nursing and retirement homes (Zorgkaart Nederland)</td>
</tr>
<tr>
<td></td>
<td>SE</td>
<td>Registret över totalbefolkningen (RTB) *</td>
</tr>
<tr>
<td></td>
<td>IL</td>
<td>Telephone directory database</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>Universal Electronic System of Population Register Number (PESEL)</td>
</tr>
<tr>
<td>Clear overcoverage</td>
<td>IL</td>
<td>Population Registry</td>
</tr>
</tbody>
</table>

5.1.3 Lists of institutions

As described in SERISS Deliverable 2.16, surveys primarily interested in the institutionalized population often prefer using lists of institutions instead of screening a gross sample of the general population for members of their target population (Schanze 2017). This does not imply that surveys of private households and institutions should not use lists of institutions as sampling frame. As the German IMOA project concluded, health surveys with a focus on both housing situations might need to draw institutionalized residents with an additional sampling frame (multiple frame approach) to obtain a sufficiently large sample of institutionalized residents. The Estonian SHARE expert also recommended to use the "Social Insurance Board register" for a survey of institutions in Estonia.

In contrast to the previous sections on population registers or health-related registers, this group of surveys does not work with a list of individuals or households, but needs to start at a higher level, namely at the level of institutions as cluster of individuals. This step adds another stage to the sampling design.
In few cases, survey programs can rely on sampling frames of institutions that are readily available. In other cases, survey researchers or the survey agency need to compile this list on their own using various sources. Both groups need to examine the quality of their sampling frame, just as it is done for population registers (see Maineri et al. 2017). Information about the size of the institutions (e.g., the number of beds) is an important additional information in the sampling frame, required for a proper probability proportional to size (PPS) selection of institutionalized residents.

In the first group, the French ANAIS study aimed to investigate residents of retirement and nursing homes, whereas the French CARE study (*Enquêtes Capacités, Aides et Ressources des seniors*) wanted to combine institutions and private households. CARE used a multiple frame approach and ran different modules in institutions (CARE-Institution) and private households (CARE-Ménage). ANAIS and CARE both used the *Fichier national des établissements sanitaires et sociaux* (FINESS) as sampling frame. FINESS is a governmental register of health-related and social institutions, namely long-term and short-term care institutions for the elderly, institutions for disabled adults and children, general and specialized hospitals, or nursing colleges. Institutions located in one of these fields need to register themselves, which is why the sampling frame does not suffer from coverage issues (see Table 4) and should be very up to date. The register is publicly available without charges. Since the register is also used to manage the institutions (e.g. inspections), it can be expected that institutions de-register if they cease to exist (potential overcoverage).

As a second example, the Swiss *Survey on the Health Status of Elderly Living in Institutions* (ESAI) used the SOMED (*Statistik der sozialmedizinischen Institutionen*) as sampling frame to draw a sample of respondents older than 65 years, who lived more than 30 days in a retirement or nursing home. SOMED belongs to the Swiss system of administrative statistics and is used to collect detailed annual information about retirement and nursing homes, institutions for disabled people, institutions for addicted people, and rehab hospitals (Swiss Federal Statistical Office 2005). Similar to FINESS, SOMED should be very exhaustive and up to date due to its administrative purpose. Also, its mission contains the dissemination of data for the public and science (ibid.), hence, it should be available for drawing a sample of institutions and institutionalized residents for scientific purposes.

The study *Adults with Learning Difficulties in England* (ALDE) used several sources in a multiple frame approach to draw its sample of institutionalized residents: the UK government "produced local databases of supported accommodation providers", the *National Care Standards Commission* (NCSC) compiled a list of registered care homes, and the Department of Health compiled lists of NHS long-stay hospitals (Malam et al. 2014, 356). Selected operators of institutions from those three frames were contacted and the outcome revealed different response rates of operators, reaching from only 35% for NHS hospitals to 68% for social service institutions. Within the cooperating institutions, response rates were more equal for the residents between 70% to 83%. Weighting was used to account for unequal selection probabilities (see Section 5.3).

As an example for a survey program that had to compile their own sampling frame of institutions, the cross-sectional study *Onderzoek naar ouderen in instellingen* (OII) aims to examine the social networks, health status and usage of care services by the elderly population living in retirement and nursing
homes. According to researchers working on the OII survey, the compilation of a sampling frame of institutions was a major obstacle. In the early phase of the study, the survey was conducted by a private market researcher. This company used their own register of care and nursing homes to draw a sample of institutionalized respondents. In 2000, 315 institutions were in the gross sample and 216 institutions agreed to participate in the survey (68.6%). This gross sample resulted in a net sample of 981 institutionalized persons. In the following two waves in 2004 and 2008, the gross sample size of institutions was increased to 364 and 507 institutions respectively. The participation rate of institutions declined to 55% in those two waves. For the 2015 wave, the OII researchers decided to commission the Dutch National Statistical Institute (CBS) with the sampling and fieldwork, because the NSI had a better access to relevant information while compiling the sampling frame of institutions. A CBS team compiled a list of institutions, apparently also relying on a website which lists health care services in all Dutch regions.

The gross sample comprised 548 institutions and CBS achieved a participation rate of institutions of 54.4%. In the net sample, 1616 institutionalized persons were included. CBS also decided to draw a reserve sample of institutions and residents within institutions to account for operators who decided spontaneously to end their cooperation or for respondents who deceased or left the institution before the interview could take place. To assess undercoverage in the sample, CBS compared the achieved sample on two occasions with the Dutch long-term care insurance register (AWBZ).

The MEA Migrant Study investigated Syrian refugees in Bavaria. The researchers drew a number of municipalities proportional to the assumed sizes of the target population. Within those municipalities, smaller communities and cities were selected and the researchers contact regional governments (Regionalregierungen) for a full list of refugee accommodations and the number of residents within those accommodations. According to a researcher working on this study, the selection of small areas in multi-stage sampling design made the enumeration of institutions easier.

The SERISS Survey Inventory includes further examples of surveys that compiled list of institutions to draw a sample of institutionalized residents. The Canadian National Population Health Survey (NPHS) was a panel survey, which covered community-dwelling respondents alongside institutionalized respondents until 2004. It followed those panel members who moved from private households to institutions and also drew a sample of 2,600 institutionalized residents for the very first wave in the mid-1990ies. To compile a sampling frame of care facilities and hospitals, Statistics Canada relied on their own lists and asked provincial ministries of health to verify and update those lists (Statistics Canada 1995). Only relevant types of institutions were kept and the lists were stratified by region, by the type of institution and by the number of beds within the institutions. The sample of 2,600 residents was drawn proportionally to the size of the strata (ibid.). The number of respondents drawn within institutions increased with the size of the institutions.

The German Teilhabe Study aims to investigate the participation of disabled persons in various public spheres (e.g., education, labor market, health care, or justice). In addition to a sample of disabled people living in private households, the survey should also covered institutionalized residents. For this purpose, a scoping study was conducted to assess various possibilities to draw the sample (Schröttle and Hornberg 2003). The survey can be combined with another Dutch survey (OZG) in order to draw conclusions for the entire Dutch population.

31The survey can be combined with another Dutch survey (OZG) in order to draw conclusions for the entire Dutch population.
32https://www.zorgkaartnederland.nl
This study came to the conclusion to abstain from using a publicly available online register of institutions for the following reasons: This register does not allow users to obtain a full list of institutions, but they need to work with separate queries for various area codes. Moreover, the register rather contained retirement and nursing homes and it was unclear whether it suffered from undercoverage with respect to the type of operator (public or private). Instead, the survey agency working on the study followed the recommendation of the scoping study and contacted 23 different mother organizations of operators for lists of institutions in different communities (Schröder et al. 2017). In the scoping study a similar approach was suggested for prisoners, another group that should be covered in the study. A second progress report did not contain any additional information regarding interviews in institutions. The Spanish EDAD survey used a two-stage sampling design to cover collective living quarters along-side the population living in private households (in a separate sample). At the first-stage a list of institutions was compiled by merging several sources. Information about the size of institutions (number of beds) and the ownership (public or private) were crucial information to guide the sampling design as strata (Instituto Nacional de Estadistica 2010). At the second-stage, EDAD researchers asked for a list of all residents within selected institutions and randomly drew approximately 11,000 persons. At least 5 residents were selected per institution.

Further survey practices can be found in the SERISS Survey Inventory and in SERISS Deliverable 2.16 (the German Retirement Home Survey (AHS) and the British Communal Establishments Survey (CES)), which also contains more detailed practices of two surveys of homeless populations that faced additional issues with a highly mobile population without any usual place of residence (fio.PSD and ENFAMS, Schanze 2017).

5.2 Sampling methods

After naming and classifying country-specific sampling frames, the following subsection advances sampling practices that can help to increase the number of institutionalized residents in the samples. This section builds on the previous SERISS Deliverable and extends this Deliverable with additional survey practices and results from the SERISS expert survey.

The institutionalized population is hard-to-sample because it is a rare population (see Section 3). In the remaining part of Chapter 5, we briefly introduce methods of screening and piggy-back sampling, the (rare) practice of oversampling and post-stratification weighting to lower the potential bias caused by undercoverage (and a higher nonresponse).

5.2.1 Screening and piggy-back sampling

The institutionalized population lives in institutions and therefore cannot be classified as hidden or hard-to-identify (Schanze 2017). Nevertheless, a number of surveys included in the SERISS Survey Inventory
used surveys or censuses as tools to screen the population and identify members of the respective target population. In these cases the host surveys or the census are used as sampling frames.

Kalton provided an overview of aspects that need to be considered when using a host-survey to draw a sample of hard-to-survey residents (2014, 408 et seqq.): First of all, the census or host surveys have to contain relevant variables that are required to identify members of the target population. Survey researchers act as gatekeepers and might be reluctant to add questions to their questionnaires. Timing also plays an important role, since a time lag between the collection of data for the host survey or the census and the subsequent fieldwork for the actual survey might lead to significant changes in the target population. This is also true for the institutionalized populations. Refugees or prisoners might have already left the institutions when the interviews should be conducted, residents of retirement or nursing homes might have passed away or suffered from a mental or physical deterioration that prevents further interviews.

In addition, Kalton also raises important questions with respect to the original sampling design, the coverage and the response rates of host surveys. A larger coverage and larger response rates lead to larger net samples and a higher probability to identify a sufficient number of respondents for the follow-up survey. It is important to consider that the response rates for a specific group like institutionalized respondents might be lower than the aggregate response rate. As a consequence, censuses with their full coverage and mandatory response are better suited to reduce the impact of nonresponse bias, even thought it might be more difficult to access the data required to draw a sample. Summing it up, Kalton concludes that screening for hard-to-survey groups in large-scale host surveys "is not as frequent as might be expected" (2014, 409).

This last observation also holds for the institutionalized populations, probably because it is not hard-to-identify as mentioned above. Some surveys screen population registers without direct identifiers prior to fieldwork by using a second register to identify institutions in order to exclude them or approach them in a different way. Other surveys screen a (large) host survey or a census to identify their target population, usually very special groups, such as disabled residents in private households and institutions or residents with cognitive impairments and dementia.

To start with the screening of a population register, the German IMOA project analyzed the decentralized German population register. This register does not contain direct identifiers for institutionalization but contains addresses of individuals. Researchers working on the IMOA project drew a sample of 8,000 addresses from 6 sampling points to assess different ways of identifying institutions. A first indicator was a conditional blocking notice attached to an address (see Section 5.1.1). Five out of six local authorities included blocked addresses in their address delivery, however, only three of them also included relevant information with respect to the blocking notice. Since the blocking notice is not implemented in a standardized way, checking addresses with the internet proved to yield better results, even though it was a time consuming task and sometimes led to mistakes because institutions did not have a proper internet appearance. The survey agency working on the NRW80+ survey also raised limitations because of an inaccurate identification of institutions with their own register of institutions.36

34Possibilities to re-use census data must probably be allowed in the underlying census laws and/or a close collaboration with the National Statistical Institute is required.
35E.g., by sending separate advance letters, preparing for gatekeepers, issuing different questionnaires
36See Schanze 2017 for another example, the Communal Establishment Survey
Moving on to the survey practice of using a host survey to identify a specific target population, the U.S. Aging, Demographics, and Memory Study (ADAMS) was a supplement study to the HRS. It was specifically interested in respondents with dementia and sampled 1,700 HRS respondents older than 70 years. Due to the information available about these respondents ADAMS researchers could stratify their sample by the cognitive functioning of the respondents. The Swiss RESPONS Study (RESidents Perspectives Of Living in Nursing Homes in Switzerland) used a similar design and based its sample on the preceding SHURP study, which evaluated the quality of care in a sample of nursing homes. In 2014, RESPONS selected 51 nursing homes from the SHURP sample, stratified by the linguistic region and the size of the nursing homes. In those nursing homes, approximately 1,000 respondents were selected for survey interviews.

In the case of the English Longitudinal Study of Ageing (ELSA) the piggy-back approach caused problems with respect to institutionalized residents. The Health Survey for England (HSE) was used as sampling frame for the initial sample of the subsequent panel survey. In addition to an increased non-response bias, the Technical Report also named the absence of institutionalized respondents in the HSE sample as disadvantage of this method (Taylor et al. 2007). Obviously, the survey researchers abstained from drawing a supplementary sample in care and nursing homes because ELSA "would not have attempted to follow them" (ibid., 9). As a consequence of the limitation in the sampling design, ELSA is only able to analyze factors leading to a move from a private household into institutions but cannot draw any conclusions for the institutionalized population as such.

Two other surveys included in the SERISS Survey Inventory relied on the census to draw their samples. The Irish National Disability Survey (NDS) aimed to interview disabled residents living in private households and institutions. To identify members of their target population it relied on the 2006 Irish population census. The census enumerated 370,500 disabled Irish citizens, 26,400 (7.1%) of those living in institutions such as nursing homes, hospitals or children’s homes (Central Statistics Office 2008).

In 1999, the French Étude de l’histoire familiale (EHF) interviewed respondents living in student dorms, prisons and workers dormitories alongside the general population. The survey managers distributed their self-completion questionnaires with the 1999 census forms in 2% of the census district. Their effort resulted in a high response rate (86%) and a large sample size (380,000 respondents). Prisoners belonged to the original target population of the census, which is why the EHF researchers decided to cover this part of the institutionalized population as well. They interviewed 1,700 prisoners in 28 prisons in a face-to-face interview mode. In addition, the survey researchers inserted simplified questions of their questionnaire into another survey (HID) to cover retirement and nursing homes.

5.2.2 Oversampling institutionalized residents

To avoid undercoverage of institutionalized residents, a stratification of the population and a sampling design that disproportionately samples residents from certain strata with a higher sampling fraction can be a good way to increase the number of respondents from a potentially undercovered population (Kalton 2009). To allow an effective oversampling, sampling frames need to contain an identifier for institution-
alized individuals or addresses of institutions. Table 5 lists answers given by the experts in the SERISS expert survey whether the sampling frame their survey program used allowed a direct or indirect identification of institutionalized residents.

Table 5: Does the sampling frame allow the identification of institutionalized residents? (Source: SERISS Expert Survey 2018)

<table>
<thead>
<tr>
<th>Country</th>
<th>Sampling Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL</td>
<td>Arnona files</td>
</tr>
<tr>
<td>JP</td>
<td>Basic Resident Register</td>
</tr>
<tr>
<td>CH</td>
<td>Contrôle des habitants</td>
</tr>
<tr>
<td>DE</td>
<td>Einwohnermeldeamtrechtsgesetz</td>
</tr>
<tr>
<td>NL</td>
<td>Gemeentelijke Basis Administratie (GBA)</td>
</tr>
<tr>
<td>IT</td>
<td>Local administrative register of Abbiategrasso (Anagrafe)</td>
</tr>
<tr>
<td>EE</td>
<td>Population and housing census (2011) and the Population Register</td>
</tr>
<tr>
<td>SE</td>
<td>Population of the Västermalm parish in Stockholm</td>
</tr>
<tr>
<td>EE</td>
<td>Rahvastikuregister</td>
</tr>
<tr>
<td>SE</td>
<td>Registret över totalbefolkningen (RTB)</td>
</tr>
<tr>
<td>CH</td>
<td>Stichprobenrahmen für Personen- und Haushaltsbefragung (SRPH)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>Sampling Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>DE</td>
<td>Einwohnermeldeamtrechtsgesetz</td>
</tr>
<tr>
<td>EE</td>
<td>Registret över totalbefolkningen (RTB)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>Sampling Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>Account data from 12 health insurance companies</td>
</tr>
<tr>
<td>AU</td>
<td>Australian Standard Geographical Classification</td>
</tr>
<tr>
<td>DE</td>
<td>Einwohnermeldeamt (3 times)</td>
</tr>
<tr>
<td>FR</td>
<td>Fichier national des établissements sanitaires et sociaux (Finess) (2 times)</td>
</tr>
<tr>
<td>IS</td>
<td>Icelandic National Register</td>
</tr>
<tr>
<td>UK</td>
<td>Population Registry Postcode address file (PAF)</td>
</tr>
<tr>
<td>DE</td>
<td>Register der Bundesagentur für Arbeit (BA)</td>
</tr>
<tr>
<td>PT</td>
<td>Sistema Nacional de Saúde (SNS)</td>
</tr>
<tr>
<td>PL</td>
<td>Universal Electronic System of Population Register Number (PESEL)</td>
</tr>
<tr>
<td>EE</td>
<td>Rahvastikuregister</td>
</tr>
<tr>
<td>FI</td>
<td>Väestörekisteri</td>
</tr>
</tbody>
</table>

Oversampling was rarely used by the survey programs included in the SERISS expert survey. Only two out of 27 potential surveys used oversampling for the institutionalized residents (the German MEA-Migrant and the Swiss VLV). The Swiss Vivre-Leben-Vivere (VLV) survey oversampled the strata of residents older than 90 years and aimed to reach as many respondents as in the youngest strata aged between 65 to 69 years. This approach could be labeled as an indirect oversampling of institutionalized residents, since they are much more prevalent in the oldest age cohorts (see 3). Other surveys identified in the SERISS Survey Inventory followed a similar approach. The U.S. National Health and Aging Trend Study (NHATS) aimed to target Medicare beneficiaries older than 65 years and oversampled persons at older age and Black non-Hispanic persons. The Finish Health 2000 survey (T2000) oversampled people older than 80 years with a double sampling fraction, meaning that these residents had a selection probability twice as high as it would have been the case with a simple random sampling approach.

Some surveys are able to target very specific groups with their sampling design. For instance, the German panel survey on the labor market and social security (PASS) oversampled a part of the population with a high probability of institutionalization. In 2016, PASS conducted more than 800 interviews among refugees, which implies an oversampling of this group of respondents. Another approach consists of an
oversampling at a higher stage in a multi-level sampling design. In those cases surveys do not oversample residents but institutions as such. The American Housing Survey (AHS) compiled a list of assisted living facilities and defined these institutions as an own strata, which was oversampled. The British Resettlement Surveys Reoffending Analysis survey (RSRA) sampled 76 prisons from the total of 136 prisons. To ensure a sufficient participation of the very small group of female prisoners (see Figure 6), this survey oversampled women’s prisons.

5.3 Excursus: Weighting the net sample

Kelfve and colleagues found a significant bias in some items of the 2002 sample of the Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD) when institutionalized respondents were dropped from the sample (2013). This bias still persisted when the sample of respondents living in private households was weighted for age and gender to represent the full sample with all respondents for those two variables. The authors concluded that weighting did not improve the estimation substantially (ibid.). However, this conclusion might be driven by the very high share of institutionalized respondents in the analyzed sample (22.3%) and the complete absence of those respondents as contrasted with samples potentially suffering from undercoverage (see Table 4).

Schanze and Zins conducted a simulation study based on SHARE data to assess the potential improvements survey weights can have in survey samples that suffer from noncoverage or undercoverage of institutionalized respondents (2019). According to their analysis, survey weights for age and gender indeed improved the precision of estimates and lowered the bias in two health-related variables. The weighted samples were even closer to the true value when institutionalized respondents were included insufficiently, especially when the samples were weighted for institutionalization in addition to age and gender. As a lesson learned from this research, survey researchers might be able to decrease the bias in their results by ensuring that institutionalized residents are included in the benchmark statistics used for post-stratification weighting. If information about the prevalence of institutionalized residents in different strata is available (as presented in Figure 3), adding the housing situation as a criterion to the weights might further lower the bias. However, the latter only works if there are at least some institutionalized residents in the data.

A limitation of this approach consists in the quality and the actuality of administrative data available to weight the survey sample. We presented 2011 census data in Chapter 3, data that is more than eight years old. The Estonian LFS expert also mentioned the issue of weighting in his reply to an open answer in our expert survey. According to the expert, "it is not possible to calculate [separate weights for private households and institutions] as we do not have correct and timely (quarterly) numbers of [the] institutionalized population available." If a survey only covers certain types of institutions (e.g., SHARE), the absence of (cross-national) information about the types of institutions are troublesome.

To conclude with a more successful empirical example, the German Refugee Panel used auxiliary information from the sampling frame (AZR), namely the asylum status, country of origin, gender, the date of arrival, and age of selected people, to estimate response propensities and calculate nonresponse weights.
(Kühne et al. 2019). These individual-level variables and some additional regional information were multiplied with the design weights (ibid.).

**Recommendations:**

- The sampling design will differ according to the research interest of the survey program. If the entire population should be covered, population registers are the first option as sampling frames (see Table 3). Health-related registers are a second option if a register is not available or if it suffers from serious undercoverage. A list of institutions might be the best option if the institutionalized residents are at the heart of the research interest and need to be overrepresented in the net sample.

- For some institutionalized populations (esp. for refugees and institutionalized elderly) the time between the sampling and the fieldwork should be as short as possible to account for changes in the population (e.g., due to moves or deceases).

- Some survey agencies maintain lists of institutions, especially if the register they frequently use only contains indirect identifiers of institutions (e.g., addresses). Before starting fieldwork, survey agencies often screen the sample and aim to exclude institutions if they are not part of the target population. On the other hand, those lists could potentially be used to sample (and oversample) institutions.

- A close collaboration with owners of sampling frames is required in some surveys (see refugee studies), but can be also helpful for other institutionalized populations to learn more about potential coverage issues and identifiers that can guide a (stratified or disproportional) sampling design. The SERISS project offers a generic letter to gatekeepers of sampling frames that can be easily adapted (Scherpenzeel 2018a,b).

- Weighting the net sample for the size of the institutionalized population will help to counterbalance undercoverage and nonresponse bias.
6 Gatekeepers

The inclusion of institutionalized populations in research cannot happen without interaction with personnel or staff of the institution, often termed as gatekeepers. In the SERISS expert survey, gatekeepers and the importance to address them properly was one of the most frequent recommendations in an open answer question. According to researchers working on the Dutch OII survey, convincing the operators of institutions to take part in a survey is the most difficult task. Once the operator or a Board decides to take part, the survey can be more easily conducted within the institutions. Interviewers can be introduced to the respondents by the staff of the institutions.

The interaction with gatekeepers usually involves allowing the interviewer to enter the institution and get in touch with the respondents living in the institution. It might also include supplying the researcher with a list of residents to serve as a sampling frame and with information on the number of eligible residents to guide the sampling design. Gatekeepers often serve as mediators between the interviewer and the participants and their cooperation can be extremely beneficial to the research (Fox et al. 2011; Kammerer et al. 2019; Trulson et al. 2004). As the term gatekeepers implies, they have the authority to let in, or to keep out, the interviewers. The gatekeepers can be both physical and organizational, either controlling the physical access to the institution in which the population resides or controlling access to a list (or registry) of the population associated with the institution.

As a side note, it is important to mention that the existence of gatekeepers is not limited to institutions, but sometimes also occurs in private households. A so-called refusal by proxy might be especially likely when interviewers aim to get access to old-aged or cognitively impaired respondents living in private households. Depending on their perception of the capability of elderly persons, relatives of respondents can have a positive or negative impact on the participation in a survey, as a scoping study for the NRW80+ survey pointed out.³⁷ SHARE also tested advance letters for gatekeepers in institutions and in private households (Neuert et al. 2016).

In the following chapter, we will first discuss the interaction with gatekeepers through the issue of the consent required by the researcher of the participants. This discussion will have a theoretic focus and will be based mostly on academic literature reviewed as part of this research. Then, we will examine the dealings with gatekeepers in fieldwork. This section will include a presentation of the results on this topic from our expert survey. It will be followed by a survey of literature discussing more practical issues including guidelines and tips. These two views should be understood as complementary and intertwined.

6.1 The ethical issue of gatekeepers

Gatekeepers are often perceived as a logistic or administrative barrier to be dealt with when attempting to reach the participants, but in essence they also introduce ethical complications (Heath et al. 2007). They are not just "monolithic, neutral and static figures" (Crowhurst and Kennedy-Macfoy 2013) but rather play a role in the power relations surrounding data collection. When they grant the researcher access

³⁷The report of the scoping study is not publicly available.
to the population in their charge, or when they deny it, they are guided by different considerations. Understanding these relationships and considerations can be helpful in guiding the researchers who are requesting access to the population in the institutions.

Although the research’s focus might be the population within the institution, it will very often unwittingly include and expose information about the institution within which the population resides. To some extent, this is true regarding every survey. In the process of a survey or other data collection methods, respondents are asked questions about themselves but also, perhaps inadvertently, are requested to supply information concerning people or entities related to them. While the main respondents have given their consent to participate, the other related parties have not. Homan gives an example of this in his discussion of informed consent in the context of gatekeepers, mainly in children: "pupils are themselves made the unwitting gatekeepers to fields that are spatially removed form that of the interview or questionnaire but are in educational terms approximate to it: they may reveal data relating to their home and family, or to previous classroom experience and teachers, or to relationships within the peer group" (Homan 2001, 330).

Gatekeepers can choose to deny the researcher access to the institution and the population in different ways. This objection could be a straightforward refusal or it can be a more subtle resistance. For example, Wanat describes non-obvious forms of resistance she has met when trying to gain access to teachers or to parents of students through the schools and education systems. The gatekeepers, principals of the schools who controlled access to the participants, used these methods as a way of refusing while appearing to be cooperative: passing on responsibility, controlling communication, constant request for information and forgetting (Wanat 2008). Subtle resistance can be a tactic used in delicate situations in which different levels of the organizational hierarchy might have conflicting interests regarding the research.

A main reason for the institutions choice to deny participation and refuse to participate is unwillingness to face public scrutiny. This has been explicitly mentioned in papers about research in prisons and termed "Fear of the Outsider". In his paper on the gatekeeper’s perspective to outside research, Jeffords explained this as "the belief that academic researchers without real-world experience and insight would view and interpret occurrences on the “inside” as scandals of mismanagement, but what those in the “field” considered common and ordinary" (Jeffords 2007, 89). Trulson and colleagues have been more specific in referring to the population of prisoners: "Dealing with the most violent and least stable individuals in American society is not something the occasional citizen or researcher confronts on a daily basis and the situations they witness may not be understood outside of the context of prison life. These feelings, whether justified or not, contribute to the fear of the researcher" (Trulson et al. 2004, 456). Since the institution determines the setting in which the populations lives, and cannot be ignored when researching this population. However, this does not permit the institutions to refuse outright.

Another reason institutions might refuse to participate is due to the burden and disruption the study causes. Operational demands placed on the institution and its staff as part of the study can be substantial. Also, the study could disrupt the routine of institutional life, which could be not only an inconvenience but also disruptive to the treatment of the institutional population itself. An example of this outlook is given by Jeffords who describes the view of the agency in charge of approving research at a juvenile correction system. He explains that "Any use of state resources (including secondary data requests) and
youth time unlikely to benefit the agency are not considered cost effective but rather disruptive to the agency routine and hence treatment of the youth." (2007, 91). But even the inconvenience and cost of the study to the institution in terms of staff time and space allocation should not be taken lightly. As will be demonstrated below based on the SERISS expert survey, researchers often have many requests, with the institution taking an active role in different parts of the data collection.

An additional concern of institutions might be the fear of legal action, such as lawsuits against the institution. This concern is more a concern of long-term care facilities, when the legal action is taken by a family member who is concerned for his or her relative. After listing the additional effort and responsibilities of the institution’s staff, which serves as a deterrent to participating in a study, Clearly also mentions the legal concern, "There is also [...] even fear of possible litigation as nursing homes are frequently targets of lawsuits by family members." (Cleary 2004, 75). We could not easily find specific cases in which this occurred, but this should belittle the institutions’ concern on this regard. As an additional consequence of the fear of lawsuits, gatekeepers in institutions might want to ask the relatives (as second gatekeepers) for their permission. This will result in a longer time period to get access to the respondent and might also require the interviewers to provide additional information on the contents of the questionnaire. Compared to respondents in private households, this will lead to less interviews with the first contact attempt and might also lead to more non-contacts in institutions because the fieldwork period could be insufficient to clarify whether the guardian agrees with the survey participation.

Gatekeepers working in institutions often hesitate to grant researchers access to respondents because they aim to protect the population entrusted to them when refusing to allow a survey to be conducted among them (Neuert et al. 2016). Sangl and colleagues summarized prior studies that often found the common behavior of nursing home staff "not to recommend a large proportion of nursing home residents for interviews because the staff does not deem them fit or capable for participation; these residents are often classified by staff as 'confused,' 'disoriented,' or as 'having difficulty communicating' and believed to have inadequate cognitive ability" (2007, 66). As an illustration of this motive, a nursing home employee stated in a pretest study conducted for SHARE: "If I know that it’s a person for whom such an interview would be a great burden, then I wouldn’t forbid it, but I would decline it. I would say don’t do it, afterwards the person will be struggling for at least a week and it has an influence on her mental and physical constitution" (Neuert et al. 2016, 21f.). Researchers and interviewers should take the expertise of gatekeepers very serious and potentially adapt their procedures following their opinion (see Chapter 7).

However, sometimes nursing home staff behaves in an overprotective way, and even cognitively impaired residents are able to express their attitudes and opinions if they are able to communicate verbally (Sangl et al. 2007).

Often, the institution’s denial of participation will be the result of a mixture of reasons explained above. The possibility of public criticism, time demands, the disruptions of the routine, an ever-looming concern

38In long-term care facilities the allegations were usually related to the health care level and negligence of the staff. The issue of conducting a survey or research in an institution was not mentioned as a cause for a lawsuit.
39The unpublished NRW80+ scoping study found that 34.4% of interviews in institutions could be achieved with only one contact attempt, which was the case in more than half of all the interviews in private households. Moreover, for 17% out of 306 respondents living in retirement and nursing homes no interview could be conducted because it was unclear whether the guardian would agree, whereas only 0.3% of the sampling units in private household resulted in this outcome. A longer fieldwork period should help to lower this number.
over lawsuits and even an honest concern about the well-being of its residents - all together, they present a deterrent to the institutions to let their residents participate in a survey. However, in continuation to the ethical principal outlined above, the question to be asked is about the agency of the institutionalized population, and whether this protectiveness is a form of paternalism. Why is the institution denying that agency? Apart from domiciliary rights, does it have any rights, a legal authorization? No matter the reason the gatekeepers claims, in essence they are denying the institutionalized population the option to agree to participate on a voluntary basis in the research. Ethically, they are denying the population’s agency to choose to participate.

The other concern is that gatekeepers *assume* consent on behalf of the population in their charge. In these cases, the gatekeeper could influence or pressure the respondents to participate through their authority or based on the trust placed in them. Methodological literature on the subject of gatekeepers demonstrates this when discussing a survey of children contacted through their school, which serves as the gatekeeper (Heath et al. 2007; Homan 2001). In a class setting, children might not understand the difference between a survey questionnaire and an assignment or exam. Even in cases when the survey is understood to be voluntary, the authority of the institution and the cooperation of the institution with the researchers might influence the population’s choice to participate.

As for the institution, it might be interested in the research precisely because the knowledge gained could help the institution itself. Researchers might be able to convince the managers of the institution of the benefits to be gained from the research. As an exceptional case, Trulson and colleagues recount the case of Jackson, who did research in the Texas prison system. Jackson expressed his surprise to be given unhindered access to the prisoners, to which the prison director replied, "How can we find out what we’re doing wrong if we don’t let people like you in to tell us?" (Trulson et al. 2004, 455). While willing cooperation of the institution is not usually the case, it does occur. In these cases, researchers still need to seek voluntary consent and willing participation from the respondents.

### 6.2 The operational aspect of gatekeepers

The SERISS expert survey included multiple-choice and open questions about gatekeepers and gaining access to institutions. The survey had an exploratory focus and its results can illustrate the range of experiences and protocols implemented. We will open this section by recounting some of the results of the multiple-choice questions relating to gatekeepers and the institutions’ personnel. Next, we will present the issues that have come up in the open-ended questions. The survey results will be followed with a review of literature with a more procedural focus, including tips and guidance to getting past gatekeepers.

The responses to the first question on the topic of gatekeepers are presented in Figure 9. As can be seen, in many cases the interviewers contacted more than the respondents themselves. A separate question inquired specifically about an advance letter, and 20 of the 44 experts in our survey said they used an additional advance letter for institutions and/or the family of the institutionalized residents.
Regarding the contact with the institution, the survey also included a question about the process of gaining access or approval, its results presented in Figure 10. Fifteen respondents chose the answer option "Don’t know", the most popular answer option. The high rate of "Don’t Know" responses (a third of the responses) is puzzling and we can only speculate why so many chose it. As for the more informative responses, it is encouraging to see the process more often described as short, whether unstructured or structured. The literature cited chapter also varies on this general description but highlighting the longer and unstructured stories. Meaning, some research reports long and/or disorganized procedures to gaining access (Fox et al. 2011; Trulson et al. 2004) while others focus on the standardization of the procedure (Jeffords 2007).

Our expert survey included two main open questions on sampling and fieldwork, intended to get at information that was not covered in the main sections of multiple-choice answers. Although we did not inquire about gatekeepers specifically, 14 experts referred either to gatekeepers or to the institution and its cooperation in a wider sense. Their answers included two main themes: the level of involvement required of the institution and strategies for getting approval.

First, many of the responses related to the involvement and cooperation of the institutions. Some surveys required very little investment from the institution while others needed more cooperation and assistance. However, a general undertone of the details was the understanding that gatekeepers involved in the research and this involvement should be acknowledged by the survey agency and/or researchers. We will start by giving a few examples of the specific ways that the institutions noticeably assisted in conducting the survey and then give examples of the more general involvement of the institution.

---

40this could be a due to publication bias, also known as "file drawer problem".
In surveys where institutionalized persons are sampled as part of the general population using a general register, there seems to be less effort to involve the institution. Other surveys would request the institution’s assistance in various aspects such as sampling or legal guardianship. A researcher might request the institution’s help of supplying a listing of residents from which to draw a sample. As a researcher working on the Irish *National Disability Survey* in 2006 described it, “we worked very closely with the staff in the homes to get their assistance with a list of residents to select the sample from and with setting-up the interviews”. In some cases respondents may have a legal guardian which has to consent to the interview, especially if they are unfit due to age and frailty or are underage. This could be a burden to the institution who might not be in close contact with the legal guardian (Jeffords 2007). *SHARE-Switzerland* gave the following example: “If participants live in a nursing home, a special letter has to be written to the direction of the nursing home to check who should be informed about the interview (e.g. children, guardian,...) and to check the feasibility of the interview.”. In these actions, the institution can be seen as supporting the survey agency in conducting the survey.

Some of the responses noted that the process of getting the approval of the institution to interviewing one of its residents was imperative. They expressed understanding that the institution can be uncooperative, if it is not a willing participant. This was evident by comments about the relationship with the institution. For example, *Gezondheidsenquête (HIS)* of Belgium wrote succinctly “Cooperation of the director of the institution is necessary.” *SHARE-Israel* mentioned specific actions to keep up good relationships with the institution: “To be very polite towards the respondent, proxy and the institution staff; To be aware of and respect the institution regulations (visiting hours etc.)... If needed, have all papers that the institution requires for allowing to conduct the interview ready when you come”. *SHARE-Germany* also mentioned specific actions: “We developed a targeted information leaflet for nursing home staff... This was on request of the interviewers, who strongly felt this would help them convince the staff to get to the respondent. After fieldwork, the interviewers reported it was indeed helpful.” Neuert and colleagues documented and published the results of tests of this advance letter to gatekeepers (2016).

The second theme of the responses was the best strategy to gain approval, specifically promoting a
top-down or a bottom-up approach. Respondents described and recommended different strategies when attempting to get approval for conducting the survey. In the SERISS expert survey, the minority opinion was to use a bottom-up approach. Only NRW80+ wrote, "use bottom-up approach to contact head of nursing home/LTC instead of top-down approach (via sponsors/Träger)". The other responses on this issue argued for a top-down approach, which was more common. The expert working on the Dutch Ouderen in Instellingen (OII) wrote "In the Netherlands, the institutions must first agree to participate in the study before a random sample of their residents can be drawn. Participation of institutions is thus voluntary. I would recommend getting the support of either a branch organization or (local) government to try and increase the willingness to participate in the survey." Another example is the the Swiss Vivre/Leben/Vivere (VLV) survey that wrote "We benefited from the support of the local (cantonal) authorities in charge of ageing and health, what was a great help with a few institutions directors who were reluctant to receive our interviewers." In addition, the Arrestee Survey in England and Wales suggested to "Develop good relationships with a governing body that allows you to get access to gatekeepers within those institutions." These strategies might be country-specific or specific to the type of institution, depending on the relevant regulation.

An useful reference on this issue is Lindsay’s article on recruiting participants for survey research through institutions (2005). Her research is not about institutionalized population in the sense we are discussing, but rather she is trying to reach a population through the institutions in which they are employed: factories, shops, and other businesses. In each organization, she had to present the research to different audiences, gatekeepers in this context, such as human resource managers, occupational health and safety managers, union officials, supervisors, heads of department, teachers, and the young workers themselves. In the article she describes different strategies to approaching each organization and the number of times this resulted in success, i.e. gaining access. The first strategy she presented was starting from the top and contacting general managers or human resources managers, which led to access of 7/14 of the companies contacted. The second strategy was to begin from the middle by contacting managers or coordinators relevant to her research topic, by which she gained access to 4/10 of the companies. Another strategy she described as "starting from the outside" in which she contacted trade unions outside of the workplace and attempted to recruit them to help promote her research, but this strategy was also met with limited success. These strategies, and some specific cases described in greater detail, are meant to demonstrate the time, effort and emotional energy spent in the process of gaining access. Based on her experience, it seems that non of the strategies she used in her research was particularly better in terms of cooperation and recruitment. Therefore, she does not recommend either a top-down or a bottom-up approach exclusively.

We want to conclude this chapter with two examples of studies that structured their interaction with the institution and its gatekeepers. These are meant to illustrate different issues to keep in mind as well as the level of detail that can be applied to guide interviewers and secure a standardized behavior towards gatekeepers.

Figure 15 (in the Appendix) shows the recruitment protocol which was devised as part of the Newcastle 85+ study. In this research, the institutionalized population was included alongside the population

---

41 see more about the Newcastle 85+ study in (Davies et al. 2010).
living in private households. About 10% of the sample lived in care homes (Davies et al. 2010, 4). For those respondents, a separate recruitment protocol was devised, which clearly refers to the institution gatekeepers. In addition, the inclusion of both institutionalized population and population living in private households was a basis for comparison between them. The researchers confirmed that "our experiences with residents of care homes confirmed previous recommendations that the additional time required to recruit this very important sub group of older people is substantial" (2010, 8).

Another example of the more technical and day-to-day aspects of negotiating access with gatekeepers is a paper by Trulson and colleagues (2004) and the closely related paper by Fox et al. (2011). Trulson and colleagues used records collected in U.S. prisons and kept by them, while Fox et al. interviewed inmates for their study in the USA. They differentiated between gaining access and cooperation, as another demonstration of the point we made above. The researchers in these studies needed much assistance from the institutional personnel, in various stages of the process. They gave advice on making initial contact, creating a good impression, explaining the research in the appropriate terms and more. Figure 16 (in the Appendix) displays part of the tips given by Trulson and colleagues.

This chapter dealt with the importance of gatekeepers through recognition of their different considerations, positions and roles in surveys. Taking into account the multiple positions of the gatekeepers, the researchers and the survey agencies should design a protocol that would deal with all of the possible issues involving the gatekeepers in the project, from the contact strategy through the approval process to the actual interviews. This could help insure the smooth operation of fieldwork and data collection.

**Recommendations:**

- Engage with the institutions respectfully and seek their cooperation. Take care to explain your requests appropriately with non-technical language. Take into account the level of involvement you are demanding from them. This could be help with sampling, staff time, change of routine, etc.

- Consider different strategies to gain access, mainly bottom-up or top-bottom. The best strategy might depend on the institution, country, topic and survey. Examine experiences and strategies made by other studies similar to yours (see SERISS Survey Inventory).

- Try to see the issue from the institutions’ point of view and discuss their concerns. The institutions’ might be concerned with multiple issues which could be dealt with separately - time and effort, public scrutiny and legal concerns.

- Create a protocol to assist the interviewers and structure their interactions with the institutions’ management and staff (securing standardization across interviewers). Make sure they have the proper forms and materials requested by the institution.

---

42Fox et al. advised on additional aspects of survey design.
7 Survey interviews with institutionalized respondents

After a sample has been drawn and access to respondents is secured, the actual survey interview is the next crucial step. The SERISS project aimed to examine the ability to reach institutionalized residents. We decided that reaching respondents also entails considerations on a variety of fieldwork procedures. In the following chapter, we aim to highlight some of the procedures and discuss potential adaptations of standardized procedures to allow surveys to "reach" institutionalized residents. The survey adaptations were found in the relevant literature and technical reports. Moreover, we also included some aspects in our SERISS expert survey.

The challenges posed to survey researchers strongly differ across different types of institutions. In prisons, it might be much more important to account for (safety) requirements demanded by the gatekeepers and the confidentiality of interview contents, whereas interviews with people living in refugee accommodations require the translation of questionnaire and concepts into multiple languages and the employment of culturally sensitive interviewers. In retirement and nursing homes, decisions should be made whether cognitive screening is used, whether questionnaires and questions should be designed in different, easier ways, and whether proxy interviews will be allowed.

On the basis of experiences made in two regional surveys in Berlin43, Kammerer and colleagues developed a model on how to reach so-called hard to reach respondents among the elderly population (2019). The TIBaR model is constituted by four subsequent steps: It starts by building up trust between the study and the gatekeepers and participants. Advance letters, reputable research organizations, positive interactions with gatekeepers who can function as valuable multipliers and an emphasis on data protection and informed consent (see below) are very important in this respect. Monetary and/or non-monetary incentives are the second step of the TIBaR model (see below). As the third step, individual barriers should be identified. In the cases of the two regional surveys among elderly residents mainly consisted of "mobility limitations of various kinds, language difficulties and responsibilities for the care of relatives" (Kammerer et al. 2019). As the final step of the TIBaR model, survey researchers should be responsive in the light of the given barriers, e.g., by allowing the presence of a comfort person during the interview or by scheduling multiple meetings (see below).

Survey researchers usually emphasize the importance of standardized procedures to ensure the comparability of survey results across respondents and interviewers and (sometimes) across countries. Respondents are treated equally with the same methods, also due to financial and logistic reasons (Kühn and Porst 1999). Adaptations or deviations from standardized procedures always face a trade off between uniform standards and methods that account for special requirements of some groups of respondents.

Larger cross-national survey programs usually stick to default standards. This approach is voiced in the open reply of a survey expert working on the European Health Interview Survey (EHIS): "institutionalized people are dealt with in the same way as other people selected for interview. No specific rules are to be applied." Another survey expert working on the ISSP confirmed that their survey program "did not include many of the additional assistance commonly provided for surveys among institutionalized per-

---

43 NEIGHBOURHOOD and OMAHA
44 Trust, Incentives, Barriers, and being responsive (TIBaR)
Institutionalized populations also appear to be hard-to-survey because the interview as such is more difficult. Indeed, 47.5% of the survey experts agreed that they are hard-to-interview (see Figure 14. Nevertheless, a majority of residents in retirement and nursing homes (approximately 60 to 70 percent) are still able to answer questionnaires. This group of respondents is very motivated and shows high response rates around 90 percent (Groom et al. 2009; S. S. D. ONS 2002; Sangl et al. 2007; Wetzels et al. 2008). In our expert survey, one researcher confirmed in his answer to an open question that "refusal rates are particularly low compared to persons in private dwellings". For instance, the German Refugee Panel achieved a response rate of about 71.5% (Kühne et al. 2019), and 88% of the prisoners took part in the Adult Psychiatric Morbidity Survey (APMS) in England and Wales (Singleton et al. 1998). Survey programs that try to alter their standards to get more inclusive can benefit from motivated respondents.

7.1 Cognitive screening

Institutionalized populations in long-term care institutions tend to be older and more affected by cognitive decline, illness or dementia (see Table 2). Therefore, those factors should receive more attention before running a survey in institutions for the elderly (Lang 2014). A data collection in retirement and nursing homes faces a trade-off between a higher participation rate and the validity of the survey responses (Sangl et al. 2007). In the course of deciding whether a respondent is able to participate, survey researchers risk a higher number of falsely blocked respondents when they only rely on the opinions of gatekeepers working in institutions (see Chapter 6). Testing the capacity of potential respondents with a brief cognitive screener is a more objective way to determine whether a respondent is capable to answer survey questions or whether a proxy interview is required.

In the SERISS expert survey, four respondents indicated that they used screening questions. Two respondents noted that they used the Mini-Mental State Examination (MMSE) and one of these also used the Cambridge Cognition Examination (CAMCOG). The two other responses were vague, "it was a small test, incl. 2-3 questions incorporated into the survey" and "I don’t remember the name, but classical questions about time and place". When developing their survey to measure the quality of care, the NCH-AHPS researchers employed eight items "intended to test orientation, recall, and reasoning" (Sangl et al. 2007, 69). Those psychological tasks can also be replaced by vignettes, that require respondents to evaluate and rank brief situations, for instance related to caring and the daily life of residents (ibid.).

When a cognitive screener is used in advance of the actual survey interview, ethical questions need to be taken into account. How should the consent procedure be handled before the interviewer actually assessed whether respondents are capable to answer survey questions? Potentially they need to rely on the assessment of gatekeepers and ask for the consent of guardians or relatives for those respondents who are likely to "fail" in the cognitive tests. In addition, it could put respondents under stress and disturb their well-being if their mental capacity is tested and they feel that they failed.
7.2 Proxy interviews

While interviewing institutionalized residents directly is most desirable, it is not always possible. Reasons for barriers in interviews could be cognitive impairments, language barriers, legal concerns or various other reasons. Those situations might lead to a decision to use a proxy respondent on behalf of the original respondent. Analyses of the survey agency running the NRW80+ survey in Germany found that proxy interviews can help to reduce a higher unit nonresponse in institutions and potential nonresponse bias (results not published, see also Kelfve et al. 2013). The German Ageing Survey (DEAS) decided to allow proxy interviews between two waves, leading to expectations of an increasing number of institutionalized residents as one consequence. Of course, survey programs do not need to go as far as the U.S. National Nursing Home Survey (NNHS) that did only use proxy interviews and did not conduct any direct interviews with institutionalized residents. Nevertheless, offering proxy interviews can be a relevant tool for surveys in institutions.

The SERISS Survey Inventory reveals a frequent practice of proxy interviews across survey programs. Out of 107 surveys, which at least partly included the institutionalized population and published information on their proxy guidelines, 74 surveys are listed that allowed proxy interviews (69.2%). The following section briefly sheds light into underlying motives of surveys to offer the option of proxy interviews. However, it is not always clear when to use a proxy interview and how to determine who should be asked as proxy respondent. The responses in the SERISS expert survey were varied and are shown in Figures 11 and 12, both for all the institutions and specifically for the long-term care institutions.

Regarding the rationale to use proxy respondents, some of the experts included insightful comments on the subject. One comment made by the Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD) was "it is important to allow proxy- or mixed interviews to support the respondent". It seems that at least in this survey, the proxy does neither completely replaces the main respondents nor takes over the interview. Rather, the proxy is there to assist the respondent who might have trouble with some of the questions but perhaps not with all of them. The presence of comfort persons can help respondents to feel at ease, increase the data quality, especially for replies to factual questions. On the other hand, the presence of a third person is likely to have a measurable impact on respondents and their answers to questions about behavior or attitudes. Another strategy consists of mixing proxy interviews and interviews with the original respondent as applied by SHARE. Survey programs can save time by letting informed proxies answer factual questions, e.g., on the housing or the financial situation of respondents, and reduce the burden for respondents.

The survey expert working for the Swiss Vivre-Leben-Vivere (VLV) survey added the following comment in our expert survey: "The proxy procedure was very useful since interviewing only the apt persons is a frequent source of biases in researches on elderly, but there are issues of reliability we anticipated and ethical issues we didn’t completely anticipate." This comment includes multiple points. First, the proxy interviews allow the inclusion of respondents who otherwise would have been systematically excluded, leading to nonresponse bias. The second point is that proxies cannot be treated in the same way as the main respondents themselves, neither in the data collection protocol nor in the data analysis.
How can surveys decide when to abstain from a direct interview and employ a full or partial proxy interview? Figure 11 reveals different decision-makers, namely interviewers, the staff in institutions, relatives or the respondents themselves. 21 out of 31 survey experts stated that their survey allowed decisions to be made by more than one of those parties. Interviewers working for the VLV survey contacted the management of nursing homes in order to get information about the health status of respondents, namely "his or her ability to participate in the survey" (Oris et al. 2016, 38). If the management voiced doubts, a proxy interview was scheduled (ibid.). Asking the institutions’ staff advice or guidance is not a straightforward strategy. As described by Sangl and colleagues, the staff’s opinion on the cognitive abilities of the residents, and specifically the residents’ ability to answer survey questions, are not always correct (2007). Administering cognitive test as described in the previous section could be another alternative, however, potentially a more costly and complicated strategy. It was used only by four survey programs in the SERISS expert survey.

Another important point is that not all proxy respondents are alike, either in terms of accuracy and consistency or in being knowledgeable. Figure 12 presents the answers to the question in the SERISS expert survey about who can serve as proxy. Different proxy relationships can be important in different surveys, depending on the information required. Lee and colleagues found that spouse proxies tend to be more consistent over re-interviewing than other proxy respondents, and also more consistent than the target respondent, when inquiring on levels of disability (2004). In wave 7 of SHARE, proxy respondents were involved in every third interview in institutions, compared to less than 5% in private households. Nearly 90% of the proxy respondents were close relatives of respondents, either spouses or partners (57.7%) or children or children-in-law (29.4%). In addition, other relatives (6.5%), nursing home staff

---

45Classification of institutionalized respondents based on Schanze and Zins 2019.
(1.0%), home helper (2.0%) or friends (1.5%) served as proxy respondents.

As important limitation of proxy interviews it should be mentioned that they are not reliable sources for subjective information, namely for respondent’s attitudes and behavior (Isengard 2002; Sangl et al. 2007). Since proxy interviews cannot replace regular interviews with the respondent, they should only be used for appropriate topics. This might be a reason for many social surveys to rule out proxy interviews, as the added value is rather limited. However, in our expert survey the NRW80+ survey expert argued in favor of not restricting proxy interviews on the hard facts from beginning on. Many proxy informants are willing and able to provide researchers with additional valuable information. However, interviewers should try hard to get in touch with the respondents themselves.

In every case, proxy respondents can be a good (additional) source for factual data, either as a partial response to a survey or in creating a census. For example, the French *Enquête auprès des établissements d’hébergement pour personnes âgées* (Study of Residents in Sheltered Accommodation for Elderly People) and the *Enquête auprès des établissements et services pour enfants et adultes handicapés* (Study on Institutions and Services for disabled children and adults), both run by the French Ministry of Health, used only proxies. The surveys are carried out every 4 year and collect basic information on the characteristics of the institutions, the basic profile of the staff and information on the residents. This information includes the age, sex, date or year of entry and level of disability and dependence.46

---

46See each survey for the complete list of survey-specific information collected: Res-
7.3 Potential adaptations of the questionnaire and survey procedures

Survey researchers need to think about ways to facilitate the response task to account for special requirements of institutionalized respondents and enable their participation. In the following sections, we suggest a few potential adaptations that have come up in the SERISS expert survey, either in the questions or in the open answers of the respondent. This is certainly not an exhaustive list, but can serve as starting point for researchers interested in including the institutionalized population in their survey. As a positive side effect, some of the suggested adaptations might also help to cover hard-to-survey groups living in private households.

7.3.1 Survey mode

The survey mode exerts a strong influence on the contents and the characteristics of a survey. For instance, specific sensitive topics cannot be addressed in a face-to-face survey, or the length of a questionnaire should be shorter in interviews conducted via the phone. Additional mode effects are expected when older respondents are interviewed (see Kühn and Porst 1999). The SERISS expert survey included a question about the survey mode recommended for surveying the institutionalized population. Its results are presented in Figure 13. The most recommended method is the face-to-face interview mode with interviewer (either CAPI of PAPI), selected by 29 survey experts. The second most-recommended option was a mixed mode, where seven experts specified different survey modes or combinations of modes. Other options were less recommended, with two experts opting for phone interviews (CATI), and one expert recommending a self-administered interview via PAPI by mail. No expert recommended Self-administered web survey and five respondents did not recommend a survey mode.

A face-to-face mode or a mixed mode with a face-to-face component was recommended by all the experts working on surveys in health-related institutions and nursing and retirement homes. Obviously, self-administered modes (especially when online) are less suited for an older population, especially in the light of potential visual impairments. For older respondents, but also in other populations (e.g., refugees, prisoners) experienced interviewers are crucial to gain trust by respondents. Interviewers can guide respondents and motivate them throughout the interviewer, they can also provide assistance to respondents for more difficult questions (Lang 2014). Moreover, as Kühne and colleagues pointed out, observations made by interviewers during their visits in refugee accommodations can be extremely valuable when analyzing unit nonresponse of such an understudied part of the population (2019). In the light of those tasks, the survey expert working on the Dutch Pilot 2T survey recommended to "provide an additional training to the interviewers". In this specific survey interviewers were also in charge of drawing a sample of...
residents within the institutions and it was important to reassure interviewers before the fieldwork started.

In contrast, Fox and colleagues reconsidered the personal verbal interview when researching the inmate population. Since the topic of crime and victimization might be sensitive, they decided not to conduct personal interviews where the respondents would be required to verbally respond, therefore being influenced by social desirability concerns. Instead, they used a pencil-and-paper survey which was written for a basic level of comprehension. The survey was also read aloud to further assist respondents who might be functionally illiterate. This survey instrument, along with other measures, were used to ensure the inmates felt safe responding truthfully (Fox et al. 2011).

Large survey programs like the International Social Survey Programme (ISSP) or the Labor Force Survey (LFS) allow countries to cover the institutionalized population, but they do not cover specific institutionalized groups. It is interesting to see that survey experts working on those survey programs recommended other survey modes like CATI or self-administered modes. As mentioned earlier, those programs appear to not adapt their survey modes to a rather small and specific group like the institutionalized population. The experts who recommended mixed modes suggested a variety of options and combinations, ranging from the most resource-demanding to the cheapest and less-guided. Two recommended a face-to-face interview wherever possible but to allow a PAPI interview with a proxy respondent when it was not possible. One expert suggested a face-to-face interview with an online part, while another expert specifically suggested to use the cheaper modes, but to use to more resource-demanding methods, such as face-to-face interviews when necessary. Two other responses allowed for either a self-administered web survey or a self-administered PAPI option.
7.3.2 Informed consent

The respondent’s formal consent, hand-signed or recorded electronically, has become the obligatory standard in survey research. The ethical principal to keep in mind is to respect the personal agency of the institutionalized population and the caution to be used when attempting to deny that agency. Questions to keep in mind in this discussion are, is the participant allowed to give his or her consent and to withdraw it? If gatekeepers and operators of institutions are involved, what are their assumptions regarding the residents’ ability and right to consent? Moreover, institutionalized residents are issued to the rules and authority of the operator of institutions and might feel a greater pressure to take part in a survey. Survey managers have the ethical obligation to ensure that respondents understand their rights, the voluntary nature of social research, and the aims and contents of the survey.

The scoping study for the NRW80+ survey found a significantly lower success rate of first contact attempts in institutions compared to private households. This can be explained with the necessity to get the consent of relatives before starting the survey interview. This is confirmed by the German SHARE team in our expert survey. The gatekeepers "can often formally not decide [to] let the respondent participate if he/she […] is under guardianship. This must be done by family." The NRW80+ expert also raised doubts whether the low number of proxy interviews due to "heavy burden […] for obtaining informed consent" are sufficient to alleviate the problem of underrepresentation of institutionalized residents.

Malam and colleagues briefly described how they gained consent of respondents with learning difficulties as part of the ALDE survey (2014, 360): Interviewers should take their time to explain the aims of the survey and the survey process to respondents, also by making use of showcards. Following this explanation, respondents needed to answer a number of simple questions to check if they understood the nature of the survey. If they understood the contents, they were asked to provide their consent. If the condition was not fulfilled, interviewers moved on by asking a close relative or independent support person for their consent.

7.3.3 Wording and simpler language

Questionnaire adaptation can happen in different degrees and for different needs. One simple way could be to adapt the readability of the questionnaire or assisting devices such as showcards by using larger print. However, making a greater effort to adapt the questionnaire will facilitate the inclusion of the institutionalized populations.

A more substantial adaptation could be altering question wording and question types. Respondents need to process the question content, retrieve relevant information and translate those information to a given answer scale - a cognitively demanding task (Kühn and Porst 1999; Schwarz 2007). The development of the Nursing Home Consumer Assessment of Health-care Providers and Systems (NHCAHPS) demonstrates the use of cognitive interviewing to help researchers learn how respondents understand and answer questions. The researchers tested different types and formats of questions. They found that rating ("How much do you like...”) worked better than reports ("How often do you like...”) and that a non-specific present time period provided the most reliable responses compared to a single day or a reference period
of multiple days (Sangl et al. 2007, 70f.). They also found the 0 to 10 response scale was easier to handle than a scale with labels (worded response categories) while a scale with "Yes/Sometimes/No" was a good alternative in the field tests. Methodologists should consider different forms, types and orders of questions and answers to find the best questionnaire for their population of interest and study goals. The French Étude de l’histoire familiale (EHF) fielded a number of simplified questions as part of the HID survey\(^\text{47}\) in retirement homes. And the Irish National Disability Survey (NDS) used additional two questionnaires for the adult and infant population living in institutions with fewer items (see Central Statistics Office 2010).

An alternative approach could be to use a standard questionnaire but to allow the interviewers of the institutionalized population leeway in rewording or explaining the questions when interviewing them. Conversational interview approaches In one of the answers to the open questions in the SERISS expert survey, Ouderen in Instellingen (OII; Netherlands) wrote the following: "Interviewer instruction is critical. Interviewers must have a good understanding of what information each question is meant to obtain. The institutionalized population can have difficulty understanding certain questions and need help from the interviewers. Interviewers must be given room to formulate questions slightly differently."

Different considerations could lead to different adaptation strategies. In the case of a new survey or instrument development, researchers might be more inclined to change the wording of the questions. In surveys including both the privately-residing and the institutionalized population, or in ongoing surveys interested in maintaining comparability, the second approach might be better. Testing existing or new questionnaires with pretests and focus groups can help to identify misunderstandings or redundant parts in the questionnaire.

Regarding redundancy, not all questions apply to the specific living circumstances of institutionalized residents. Survey researchers might abstain from asking all questions to institutionalized respondents and need to program respective filter questions. The Swedish SWEOLD survey followed this approach according to a researcher. For instance, institutionalized residents did not receive questions regarding informal care. SHARE uses a system of different respondents, some of them answering question on the financial and housing situation on behalf of their households. Residents of nursing homes usually do not answer those modules on finances and housing.

### 7.3.4 Timing and multiple meetings

As another important factor raised by several experts in the SERISS expert survey, the timing of fieldwork should be planned more generous if surveys aim to include institutionalized residents. One expert pointed out that more time might be needed to involve the legal representatives or guardians of institutionalized residents. In another case in the Netherlands, a decision by residents committees were required and those committees did not meet very frequently.

While there are advantages to conducting the interview in one meeting, when interviewing the institutionalized population this might change. In an open question in the SERISS expert questionnaire, two respondents pointed out the difficulty of interviewing elderly respondents in long-term care facilities due

\(^{47}\text{Handicaps-incapacités-dépendance}\)
to deteriorated health. As explained by SHARE-Israel: "Be aware that respondents residing in institutions might be more ill and have less cognitive skills than the rest of the sample population and prepare for it, even prepare to conduct the interview in more than one appointment". In the words of the Ouderen in Instellingen (OII; Netherlands) respondent: "due to the fragile health of the residents they tire easily. Interviews should therefore be spread over two or three days if needed."

We already mentioned the approach to lower the time passing by between a sample has been drawn and the fieldwork starts. As another aspect of timing, panel surveys need to think about the timing of their waves. Panel mortality is expected to be higher among institutionalized respondents, and the survey expert working on the German Ageing Study (DEAS) recommended to keep the time between two waves as short as possible if the number of institutionalized respondents should be larger.

7.3.5 Translation

Translating the survey questionnaire is a widespread practice in all surveys - in multilingual populations, cross-national surveys or when implementing a standardized survey instrument in a new language. When considering the institutionalized population, the translation of questionnaires is a very important topic specifically for surveys among refugees and prisoners. In the SERISS expert survey, we found that surveys including refugee accommodations were the most likely to translate the questionnaire into a large number of additional languages. All 5 surveys in our expert survey which included specifically refugee accommodations responded that they translated the questionnaire. Some translated the questionnaire to Arabic, others to more languages and as many as eight languages. Eight other surveys reported that they translated the questionnaire into additional languages such as Spanish and Russian.

The German Refugee Panel translated its German source questionnaire into English and five additional languages using two translators for a better quality of translations (Kühne et al. 2019). During the interview, the CAPI questionnaire showed the German version and the respondent’s language on the screen. Additionally, audio files were prepared for illiterate respondents. The entire set-up aimed to reduce language barriers and increase the quality of data (ibid.).

Other examples of translations can be found in the SERISS Survey Inventory. A British survey research with 400 refugees from 5 countries was carried out for Department for Work and Pensions slightly before 2004. The countries of origin were Sri Lanka (Tamil refugees), Iraq (Arabic speaking and Iraqi Kurdish), Kosovo, Turkey (Turkish and Kurdish) and the Somali regions. The questionnaire was translated and checked by multiple translators and bilinguals. The British Survey of New Refugees (SNR), with the data collection taking place between 2005 to 2009 translated its English source questionnaire into the ten languages recorded as most commonly spoken by refugees according to the UK Border Agency Caseworker Information Database (Cebulla et al. 2010). This highlights one aspect of the translations: The refugee population which is under study is constantly changing and translations need to be continuously updated to include new languages.

Even surveys in retirement homes and health-related institutions might need to translate their questionnaires depending on the characteristics of their target population. The English Adult Social Care Survey

---

48 Arabic, French, Shona, Somali, Amharic, Farsi, Swahili, Tigrinya, Turkish and Urdu.
(ASCS) provides questionnaires in 17 different languages.\textsuperscript{49} The translations are based on the basic questionnaire for community-dwelling residents. Linking back to the section on questionnaire adaptations, the amount of translation work easily increases if various versions of the same questionnaire needs to be translated.

### 7.4 Incentives

In the SERISS expert survey, we asked about the policies the survey programs implemented with respect to incentives. Following the results, more than half of all survey programs did not offer any incentives. 11 programs issued conditional incentives (28.2\%) and another 6 programs used unconditional incentives (15.4\%).\textsuperscript{50} A large majority of those survey programs with incentives used monetary or quasi-monetary incentives. We also asked for the exact amount of money paid to respondents. After converting national currencies to euro, the amounts can be compared with caution, especially in the light of different national price levels. However, the 11 respondents indicating the exact amount provided numbers between 2.40 (SHARE Poland) to 43.30 (SHARE Switzerland). In average, institutionalized respondents received approximately 14 as incentives.\textsuperscript{51} Refugees are a very deprived group and should profit more from monetary incentives than most other parts of the population. However, the German Refugee Panel did not use monetary incentives, because it turned out during the pretest that the incentives had to be deducted from respondents’ social benefits. Instead, interviewers handed over small presents in advance to the interview “in order to avoid the impression that the gift is payment or even bribery” (Kühne et al. 2019).

In addition to monetary and non-monetary incentives, altruistic motives might be highly relevant to institutionalized residents. As the expert working on the Swiss VLV survey pointed out in our expert survey, “the interviews were difficult because of the health status of the residents, but when it could happen it was in most cases a nice moment for the interviewed, breaking life monotony and solitude, speaking with someone giving her/him attention”. The opportunity to reveal the personal opinion and to be an active member of the society are motivating factors (Kammerer et al. 2019), even more when the survey addresses topics that are of relevance for institutionalized residents, such as age, health, housing situation.\textsuperscript{52}

### 7.5 Ethical considerations of surveying the institutionalized population

While dealing with vulnerable groups residing in institutions, researchers should also take into account ethical aspects of their survey. For instance, refugees might have suffered from war, injuries, a dangerous flight, and traumas in their native countries. They left their homeland, probably against their

\textsuperscript{50}5 survey managers did not know
\textsuperscript{51}We abstain from additional analyses given the low number of valid answers provided.
\textsuperscript{52}On the basis of multiple studies, Kühn and Porst concluded that the motivation of older respondents is strongly influenced by their personal involvement (1999).
will, might have lost or left behind close relatives and have to get used to a new country and difficult living circumstances in a communal accommodation. Respondents might answer questions about their own background with an increased social desirability, if governmental organizations are project partners and respondents believe that the survey was tied to the asylum process. Residents living in prisons are sometimes asked to report about illegal behavior before or during their imprisonment. Such a self-accusation is very sensitive and prisoners might be afraid of legal consequences when they do not trust the anonymization. Residents of long-term care homes might feel their impairments even stronger while answering the survey questionnaire and hope for concrete improvements of their personal situation when answering questions on the quality of life or care, even when the survey just aims to record the status quo as it is.

Designing a survey with methods that enable the target population to participate can be achieved by involving the target population during the survey cycle. As a good example of such a practice, we would like call attention to the Adults with Learning Difficulties in England study that we think demonstrates how combining these adaptations could lead to including the institutionalized population in a profound way.

The study involved people with learning difficulties in all stages of the research from design to reporting, as can be seen in the excerpt in Figure 17 (in the Appendix). People with learning difficulties were also included as part of the research team. While the study did not target specifically institutionalized population, 31% of the sample of adults with learning difficulties live were living in some form of supported accommodation and including them was important (Emerson et al. 2005, 27). The survey included a combination of the adaptations suggest above, and others. As can be seen in Figure 17, The report as well as the survey used simple language that is easy to understand as well as images to get the message across. The German Teilhabe study targeted a similar group of citizens and also published a report of results in easy language.

As part of the ethical considerations, we would also like to flag the topics of data protection and anonymity of respondents. Survey researchers need to decide whether an identifier in the data will allow users to differentiate between institutions and private households. This can be critical, because institutionalized residents belong to a small group with a prevalence between 1% to 3% in the total population. Regarding their characteristics, they are special in many ways and can be more easily identified as outliers as a consequence. In addition, operators of institutions know their clients very well and could be motivated to learn more about the opinions voiced by their residents in the publicly available data set. Without going too much into detail, common methods of data anonymization can be very relevant for data on institutionalized respondents, including the suppression of certain values, an aggregation of smaller categories or even changes of specific values.

As our final argument in this section, ethical considerations might also lead survey researchers and National Statistical Institutes to the decision to extend their coverage from private households to institutions.

---

53 All the information presented about the study can be found at: http://webarchive.nationalarchives.gov.uk/201205032221457/http://www.dh.gov.uk/en/Publicationsandstatistics/PublishedSurvey/ListOfSurveySince1990/Generalsurveys/DH_4081207

From a survey research perspective, citizens somehow lose their right to express their opinions and participate as active members in their society once they moved to an institution.

**Recommendations:**

- Since the institutionalized population is very diverse, different adaptations are required depending on the specific institutionalized group.

- A face-to-face mode is likely to be the best survey mode for most institutionalized populations. Experienced and sensitive interviewers are crucial to build up trust between them and gatekeepers and respondents. They can make decisions whether a proxy is required, assist their respondents and collect additional information about the environment and the respondent’s behavior during the interview.

- The trade off between standardization and adaptation to specific needs is the main question that needs to be answered. However, adapting survey procedures does not imply to abandon standards but rather to implement standards that are more reactive and inclusive.

- If institutionalized residents struggle to cope with standards imposed by a survey program, a certain group of the population living in private households will face the same troubles. Hence, surveys of the community-dwelling population can benefit by some of the strategies suggested above to improve their coverage of hard-to-interview respondents living in private households and decrease a likely bias caused by unit nonresponse.
8 Discussion

On the basis of the results presented in this deliverable as well as in SERISS Deliverable 2.16 (see Schanze 2017), the following chapter aims to approach replies to the questions whether it is necessary and feasible to cover the institutionalized populations in general social surveys.

8.1 Is it necessary to include institutionalized populations?

A general answer to this question would be most probably misleading, since it really depends on the research interest of a given study and the definition of its target population. Chapter 3 with its quantitative description of the highly diverse institutionalized population in Europe provides the relevant background while answering this question.

On the basis of work done in SERISS Workpackage 2 since 2016, we think that it is not necessary to include the entire institutional population in general social surveys in Europe. The institutional population is very diverse and heterogeneous, leading to a wide range of possible survey adaptations required to cover it adequately. Most groups, like prisoners, refugees, students, or monks, are relatively small in nearly all countries and should not change any survey estimates in general social surveys. Following a cost-benefit approach, the costs to cover the entire population would probably exceed the benefits.

However, we definitely recommend to consider an inclusion of institutions for the elderly population in general social surveys to avoid bias of survey estimates. In Europe, residents of retirement and nursing homes are the largest group within the institutionalized population, and with a continuing demographic change in most European societies this group is expected to get larger and more relevant (Rodrigues et al. 2012). Moreover, this part of the institutionalized population is unequally distributed across age groups, with a larger cluster of institutionalized residents within the oldest age cohorts. Especially those analyses that aim to draw conclusions for the elderly population are more likely to produce biased results if they only rely on a sample of private households. Tables 1 and 2 provide an indication of topics that might be biased when institutionalized residents are excluded.

To add the international perspective, cross-national survey programs need to pay special attention to the institutionalized populations. In a number of cross-national survey programs, such as the ISSP, LFS, and SHARE, a number of countries cover institutionalized residents, whereas other country teams exclude them from their target population. This could be either caused by a deliberate decision or by methodological issues, for instance because the sampling frame suffers from noncoverage or undercoverage. Even a target population of “private households” will differ because national laws and cultures lead to a diverging usage of institutions. To give an example, Chapter 3 showed rather similar patterns of institutionalized residents within specific age cohorts across countries. However, in France, 16% of the women older than 80 years are institutionalized, compared to less than 5% among women in the neighboring country Italy (see Figure 3). Analyses of female healthy aging in Europe might come to the conclusion

---

55In the case of SHARE, this observation holds true for the baseline wave.
that French women are healthier than Italian women, simply because a higher share of less mobile and less healthy women live outside the target population in institutions in France. Since cross-national survey programs want to publish comparable data, we recommend a discussion of the definition of the target population and related methodological issues to avoid distortions in analyses.

In addition to general social surveys, specific research questions might require an inclusion of institutionalized residents. A number of survey programs will continue to cover institutionalized residents. With the increasing number of refugees arriving in Germany from 2015 onward, some survey programs started to cover this group in order to learn more about their backgrounds and aspirations. The institutionalized population can be highly relevant and interesting if questions on trust in justice (prisoners), migration and integration (refugees), education (student dorms and boarding schools), or religion (monasteries) are asked.

8.2 Is it feasible to include institutionalized populations?

If survey researchers come to the conclusion that an inclusion of institutionalized residents is desirable or even necessary, questions about the feasibility will follow. The Health Survey for England (HSE) included institutionalized residents once in the 2000 wave. According to a researcher, it was originally intended to repeat this exercise in a 5 year cycle. In fact this approach was not followed, the HSE still covers only private households. The researcher named the higher costs of interviews in retirement and nursing homes as one reason, often caused by the necessity to use a proxy interview instead of interviewing the respondent directly. Moreover, the absence of a high quality and up to date sampling frame of institutions and relevant auxiliary information like the size of the institutions led to difficulties in drawing a random sample. Getting back to the previous section, the researcher also raised the question as to what extent policy makers were really interested in getting good data and information on transitions into care homes and the health status of the institutionalized population.

In the SERISS expert survey, we asked the survey experts to share their general experiences on surveying the institutionalized population. Figure 14 summarizes the results as to which extent four dimensions of the hard-to-survey classification also apply to institutionalized respondents. Apparently, they are not so hard-to-sample and hard-to-persuade to take part in survey interviews. However, most probably due to gatekeepers working in institutions, a majority of survey experts thinks that institutionalized residents are indeed hard-to-reach. Also, the interview itself might be more difficult than in many private households. Nearly 60% of the survey experts attached the hard-to-interview label to institutionalized residents. Notwithstanding the large fraction of survey researchers emphasizing certain difficulties, there is still a significant number of researchers who reject the given labels for institutionalized residents.

On a more abstract level, Willis questioned the nomenclature of "hard-to-survey" populations and argued "for a subtle but important shift in investigator viewpoint" (2014, 175). Indeed, some groups pose stronger challenges for survey research, however, sometimes the label hard-to-survey "mainly reflect[s] the separation between researcher and respondent" that should be bridged with adequate survey methods (ibid.). For instance, Kühne and colleagues came to the conclusion that refugees are indeed hard-to, but not impossible to sample (2019). The framing of the institutionalized population as a "hard-to-survey"
population and the adherence to “traditional” definitions of the target population might prevent survey researchers to extend the coverage of survey programs. To give an example, survey researchers could be afraid that interviews in prisons might be too dangerous for interviewers. In fact, an ONS survey in England and Wales did not conduct survey interviews with 15 prisoners out of a sample of 3,500 prisoners “because [those 0.42%] were too dangerous or were too disturbed at that time” (Singleton et al. 1998, 6). And another expert working on a German survey among refugees (PASS) was positively surprised that “access to refugees was much easier than expected”.

The SERISS Survey Inventory lists more than 150 survey programs that included institutionalized residents in different parts of the world. More than 100 of those surveys covered institutionalized residents alongside the community-dwelling population. This substantial list of survey programs gives a first impression that it is definitely not impossible to survey residents living in institutions. In line with an expert opinion voiced by a survey researcher working on the UKHLS survey, researchers of the German IMOA project came to the result that an adequate inclusion of elderly institutionalized residents is only possible with a significant additional effort. Even this effort might not be sufficient to achieve data which would allow separate analyses for this group. The usual sample sizes with 2,000 or fewer respondents is sim-

---

56This image of prisons might be also strongly influences by its portrayal in TV, literature, and films.

57For the matter of completeness, the survey expert working on the English Arrestee Survey (AS) raised the problem of “extremely angry respondents, who had a known history of violence, or were under the influence of drugs or alcohol and incapable of giving an interview.” This could probably be explained with the fact that respondents were arrested very recently. The expert also added that respondents “were normally very keen because it got them out of their cell and was something to do.” The AS achieved a response rate of about 24%.
ply too small for the rare group of institutionalized residents. A better representation can be achieved by drawing separate samples for community-dwelling residents and institutionalized residents, as it is done in the French CARE survey or the Dutch OII survey. The UKHLS expert suggested to include institutionalized residents either in surveys with large samples or in longitudinal surveys that can accumulate institutionalized residents over time. However, even if separate analyses are not possible with usual sample sizes, the IMOA researchers added that excluding institutionalized residents from health surveys or surveys with health as one of the topics is not justifiable.

9 Conclusion

This deliverable dealt with the institutionalized populations, a part of European societies which is often excluded from social surveys and even from many administrative data collections. The lack of data results in a knowledge gap, and the available information about this group is scattered across various sources (see SERISS Survey Inventory). Although the institutionalized population is often labeled as hard-to-survey, the large number of survey programs covered by the SERISS Survey Inventory proves that they are not impossible-to-survey. We analyzed the practices used by those survey programs and their experiences, especially by running an expert survey among researchers of those survey programs. The deliverable started by providing a top-down definition of institutions and the institutionalized population. Figure 1 might help researchers to develop practical and concise guidelines for their interviewers to identify institutions and either include or exclude them from the survey sample. We emphasized the necessity to communicate the definition of the target population to data users as clearly as possible in order to avoid biased conclusions.

Mainly relying on the European census, we described the institutionalized population in Europe. The census data were collected in 2011 and cover only very few hard-facts, notably not including the type of institution. In most European countries, the institutionalized population is a rare group among younger and middle-aged citizens, but reaches a substantial share around 10% within the older age cohorts older than 70 years. Those residents live in retirement and nursing homes. In a number of countries, a significant share of young residents is also institutionalized, mainly living in educational institutions, or military barracks. Since bias is determined by the relative size of the excluded population and its statistical distinctiveness, we also described factors that explained institutionalization of old-aged residents in many studies (see Tables 1 and 2). Researchers working with those variables and related variables face a peril to obtain biased results if institutionalized residents are excluded, especially when the research question focuses on older age cohorts.

Chapter 5 describes different types of sampling frames used by various survey programs and provides a non-exhaustive overview of sampling frames in different countries. Those sampling frames can be population registers, health-related registers held by insurance companies or general practitioners, or other lists of institutions. The FINESS and SOMED registers of institutions are two examples for appropriate sampling frames in France and the Switzerland, however, usually survey researchers need to compile

---

58 At least it was not included in the harmonized data at the Eurostat level.
their own lists of institutions. Combining such a list of institutions with a population register can be one way to oversample institutions and allow users to analyze this part of the population with a sufficiently large number of cases.

Apparently the restricted access to the institutionalized population is one of the biggest issues for survey researchers (see Figure 14). We described underlying reasons for gatekeepers to prevent survey interviews within their institution. Those motives can be manifold and reach from fear of legal actions or the necessity to keep certain routines, to an overprotection of their vulnerable clients and the necessity to get in touch with legal guardians and/or relatives first. On the other hand, gatekeepers can be extremely helpful in securing cooperation by respondents and informing interviewers about potential barriers (e.g., language barriers, cognitive impairments). Either a top-down approach or a bottom-up approach might lead to success while convincing gatekeepers to collaborate with the survey program. In every case, survey researchers should prepare their interviewers to communicate with gatekeepers in a non-technical language and equip them with separate advance letters.

Finally, we briefly introduced potential adaptations of fieldwork procedures. The adaptations need to follow the challenges posed by specific groups of the heterogeneous institutionalized population. A face-to-face interview mode should be most appropriate to cover refugees, residents of retirement and nursing homes, or patients in health-related institutions. Experienced, sensitive, and well-prepared interviewers are crucial while gaining access to institutions and conducting the survey interview within institutions. Offering proxy interviews might be necessary to allow interviewers to collect information about a larger number of respondents, at least some hard-facts that can be used to analyze the potential impact of item nonresponse on the bias in results. Moreover, we also touched additional aspects, namely informed consent, the questionnaire contents and question formats, the timing of interviews, translation, incentives and ethical considerations.

Summing it up, survey researchers need to estimate the expected financial, logistic and methodological costs of covering institutionalized residents. They need to trade off potential adaptations of sampling and fieldwork methods caused by an extension of the coverage against the necessary standardization and comparability of survey results across respondents, across countries, and over time. This deliverable aims to support researchers in the process of assessing the necessity to include institutionalized residents. We recommend to social surveys to consider an extension of the target population to include at least those residents living in retirement and nursing homes. Including other types of institutions without further adaptations will probably make a definition of the target population easier ("entire general population"), even though it would not be necessary with respect to bias. If an inclusion of institutionalized residents is deemed necessary, we hope that the contents presented in this deliverable also support researchers while making the inclusion of institutionalized residents feasible. As a positive side effect, it will also help survey programs to cover certain respondents living in private households, thereby starting to close the gap between researchers and so-called "hard-to-survey" respondents.

Disclaimer: The views and opinions expressed in this article are those of the authors and do not necessarily reflect the position of GESIS – Leibniz Institute for the Social Sciences, the Hebrew University of Jerusalem or any other organizations or researchers associated with the SERISS project.
References


Fink, Anne (2014). “Dementia and long term care - an analysis based on German health insurance data”.


Goffman, Erving (1957). “Characteristics of Total Institutions”.


ONS, Social Survey Division (2002). “Technical report A pilot survey of people living in communal establishments Key points”.


Scherpenzeel, Annette (2018a). *Background to the letter to request access for sampling for European survey infrastructures*. Tech. rep. Deliverable 2.3a of the SERISS project funded under the European Union’s Horizon 2020 research and innovation programme GA. URL: www.seriss.eu/resources/deliverables.


RECRUITMENT PROTOCOL
CARE HOME; NURSING / RESIDENTIAL

Note: First contact letters to care home managers/consultees etc should be posted out as priority within your mail list, as recruitment from this setting may require more time than direct contact with participants.

Figure 15: Recruitment protocol for people residing in care homes as part of the Newcastle 85+ study.
Appendix A:

Tips and Suggestions for Breaking In To Prison Organizations:
A Summary

Getting Permission and Initial Access: How to Deal with Gatekeepers
I. Tip One: Get a Contact
A. Get an established contact.
   1. Pick them wisely.
B. Endorsement letters are a double-edged sword.
C. Formalize an official meeting.
D. Keep in mind the agency you are asking to do research in.
II. Tip Two: Establish Yourself and Your Research
A. The initial meeting.
   1. Act professional
      a. Be on time or early.
      b. Be formal.
      c. Do not use slang.
      d. Actively listen.
   2. Dress for the occasion
      a. Dress to your audience.
      b. No ear, tongue, lip, or eyebrow rings.
   3. Establish the importance of the research
      a. How does the study benefit the agency?
      b. You are invading their turf.
         a. You provide an independent and neutral view.
         b. Incorporate their research needs.

Maintaining Access: Developing and Cultivating Relationships
III. Tip Three: Little Things Count
A. Formal activities
   1. Be on time.
   2. Show up regularly and at the same time.
   3. Show up when it is convenient for them, not you.
   4. Ask for help during “off” times.
   5. Stay out of the way.
   6. Keep regular contact with both line staff and supervisors.
   7. Follow your timeline.
      a. Draft “research progress” documents monthly
B. Informal activities
   1. Invite gatekeepers to lunch.
   2. If you are given advice, follow it for the sake of following it.
   3. Participate in other informal activities, especially if asked.

Agency Exit
IV. Tip Four: Make Sense of Agency Data by Keeping in Contact
A. Obtain a detailed understanding of agency data.
   1. Keep in contact to facilitate interpretation.
   2. Arrange bi-monthly meetings, go to lunch, send an email, or place a call.
V. Tip Five: Deliver Competent Readable Reports on Time
A. Get the report done on time.
   B. Include an executive summary.
   C. Avoid “but and if” statements.

Figure 16: Part of the summary the tips and advice by Trulson and Marquart (2004).
The Survey

In 2001 the Government said (in a ‘White Paper’ called Valuing People) that it would carry out a national survey of people with learning disabilities in England. It wanted to find out three things:

- What is life like for people with learning difficulties?
- What support do they get?
- What do they want and need?

It wanted to do this in a way that involved people with Learning disabilities at all stages of the research. To do this the Department of Health made sure that people with learning disabilities:

- were part of the committee that decided who would be asked to do the survey
- were included as full members of the research team and the advisory group
- could participate in the survey by answering as many questions as possible
- were involved in writing the report of the survey
- got to see the results of the survey. (This report is one of the ways we are doing this.)

Figure 17: An excerpt out of the full report of Survey of adults with learning difficulties in England 2003/4