



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

Deliverable Number: 2.3a

Deliverable Title: Background to the letter to request access for sampling for European survey infrastructures.

Work Package: 2. Representing the population

Deliverable type: Report

Dissemination status: Public

Submitted by: SHARE ERIC

Author: Annette Scherpenzeel (MEA)

Acknowledgements:

We sincerely thank Prof. Jaak Billiet, University of Leuven, for the original idea of using the Belgian ESS case as basis for this deliverable, and for making available his archive of letters and documents. We also thank Ruud Luijkx and Marjan van Lier (Tilburg University) who carried out the documentation and coding of the Belgian archive. Sarah Butt, Jan-Lucas Schanze, Tom Emery, Johanna Bristle and Michael Bergmann (WP2 members of SERISS) all contributed with many helpful comments.

Date Submitted: August 2018

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.





www.seriss.eu  @SERISS_EU

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 policymaking 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: Scherpenzeel (2018). *Background to the letter to request access for sampling for European survey infrastructures*. Deliverable 2.3a 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

Background to the letter to request access for sampling for European survey infrastructures

Content

- Summary 4
- 1. Introduction 4
- 2. Case study: How researchers in Belgium re-gained access to the population register for ESS sampling 5
 - 2.1 What happened in Belgium 5
 - 2.2 Data and Method 5
 - 2.3 Results 6
 - 2.3 Conclusion 8
- 3. Explanation of the arguments that were incorporated in the D2.3 letter and memo 9
- 4. Next steps: How the template letter and memo can be used 10
- References 12

Summary

SERISS Deliverable 2.3 is an attempt to improve the quality of sampling for European survey infrastructures by obtaining access to population registers in all countries. The core part of the deliverable consists of a letter from the survey infrastructures to relevant register authorities, making the case for survey access. The letter is a template, which, on the one hand, can be used by national teams and national funders, and, on the other hand, can be sent jointly by all directors of the four cross-European surveys cooperating in SERISS to exert more pressure. This report describes how the letter was developed, and how arguments were found to which register authorities are responsive. Furthermore, the report explains how the letter can be used in practice.

1. Introduction

The overall purpose of Task 2.1 in SERISS is to improve the quality and efficiency of survey sampling practices by understanding more about how sampling is conducted by the four major face-to-face surveys taking part in SERISS and taking steps to exploit potential synergies for sample frame generation/access. The specific goal of D2.3 is to use SERISS as a means of overcoming potential barriers to accessing population registers held by statistical offices (often the gold standard frame for survey sampling) by establishing closer links between the survey infrastructures and the statistical offices. The initial scoping work reported as part of SERISS Deliverable D2.1 (Scherpenzeel et al, 2016), has already allowed us to determine a) which countries have population registers that are suitable for survey sampling b) in which case access to these registers is a problem. The next step in furthering collaboration on sampling between the major surveys taking part in SERISS and the National Statistical Offices (NSIS) or other population register owners is therefore to take pro-active steps to negotiate access to these registers. Discussions among the SERISS Expert Network (involving the surveys and representatives from NSIS) concluded that this was best achieved in the first instance by sending of a joint letter or memo from the survey infrastructures to relevant register authorities making the case for survey access with arguments to which such authorities are sensitive and which are relevant to the objections/hesitations of the register authorities.

This conclusion was further supported by the report of an authority in the field of sampling and survey research, Prof. Jaak Billiet of the University of Leuven, Belgium. Prof. Billiet, together with other researchers, was successful in achieving direct register access for drawing a sample in Belgium after this had previously been denied in 2005 by the register authority on advice of the Belgium privacy commission. The archived letters that were exchanged between the researchers and the commission have been analysed in the framework of SERISS WP2, to identify which arguments were most relevant to the commission and led to the reversal of the decision. The selected arguments were, in the next step, used to construct the letter and memo that constitute the core of D2.3 and that can be sent to the register authorities in the countries where access to population registers for sample drawing is currently not possible.

This report gives some background information to the letter and memo. The aim of the report is:

- Describe the Belgium situation as a case study and present the results of the analysis of the Belgium documents provided by Prof. Jaak Billiet
- Give an overview of arguments to which the register authorities seemed responsive
- Document how the successful arguments were incorporated in the D2.3 letter and memo
- Describe the next steps to take and how the letter and memo can be used

2. Case study: How researchers in Belgium re-gained access to the population register for ESS sampling

2.1 What happened in Belgium

Access to the national population register which had been obtained for the sampling for the European Social Survey (ESS) rounds 1 (2002/03) and 2 (2004/5), as well as for other studies in Belgium was changed in 2005, not allowing direct sampling anymore. The alternative method proposed by the Belgium privacy commission, advising the Ministry of the Flemish Community, was that the population register (Rijksregister) itself (one of its services) would send an information letter to the sampled persons, and provide contact information to the researchers only for those persons who explicitly consented to participate in the survey. One of the founders of the ESS and an authority in the field of sampling and survey research was successful in achieving direct register access again by writing a series of letters explaining the scientific relevance of proper sampling and how this can only be done by using the population register for person samples.

2.2 Data and Method

The letters and documents used to improve the register access for ESS sampling in Belgium have generously been made available by Prof. Billiet to the SERISS network and have been studied within the framework of WP2. Researchers and students from the Tilburg University, associated with the European Values Study (EVS), analysed and coded all letters from both sides (Prof. Billiet and the Belgium privacy commission / register authorities) using the following coding scheme¹:

¹ The coding scheme was developed by the author of this report based on a first reading of the documents.

Table 1. Coding scheme for the arguments in the Belgium case study

Type/Content of argument	Code
General scientific value	1
General societal value	2
Response rates and response bias	3
Comparability across countries, time and other studies	4
Sampling error, sample design effect, weighting	5
Other methodological argument	6
Individual data protection, anonymity, privacy	7
Rights of citizens/respondents, their freedom of choice and information	8
Other Ethics argument	9
Transfer or processing of data by third parties (juridical)	10
Other juridical or administrative argument, referring to laws, regulations, statutes	11
Costs and efficiency	12
Other	13

Furthermore, all (Flemish) letters and arguments were summarized and documented in English in an Excel file, distinguishing between arguments used by the researchers and by the commission. The file also lists which argument was used in response to which argument from the opposite side. The Excel file can be obtained on request.

2.3 Results

In total, 5 letters were written by Prof. Billiet and other researchers to the commission (all before the decision was reversed) and 7 letters by the privacy commission, the Ministry of the Flemish community and its study department, and/or the Rijksregister to the researchers (2 letters before the decision was reversed, and 5 after that to work out the agreements for access for sample drawing). Figure 1 and table 2 show what the most used arguments were from the two sides, using the codes in Table 1 and the Excel documentation file².

² Only the letters written before the decision was reversed were used for table 2 and figure 1.

Figure 1. Word cloud graphs representing the type of arguments used by the Belgium privacy commission and by the researchers. Graph is based on the data presented in table 2, the titles indicating the argument types were shortened in the word cloud graphs.

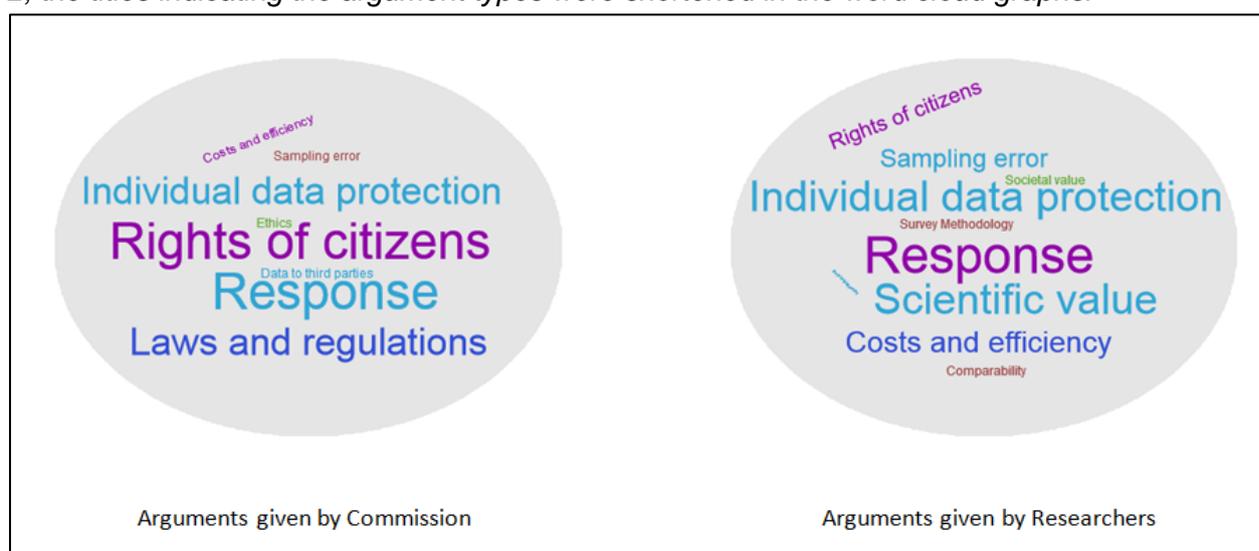


Table 2. Frequency of using different argument types by the Belgium privacy commission and the researchers, in letters preceding the reversal of the decision about access to the Rijksregister for sample drawing. Letters sent between December 2005 and June 2006³.

Code	Argument type	Number of times argument type was used by commission	Number of times argument type was used by researchers
1	General scientific value	0	10
2	General societal value	0	3
3	Response rates and response bias	4	12
4	Comparability across countries, time and other studies	0	3
5	Sampling error, sample design effect, weighting	1	6
6	Other methodological argument	0	3
7	Individual data protection, anonymity, privacy	3	10
8	Rights of citizens/respondents, their freedom of choice and information	4	5
9	Other Ethics argument	1	0
10	Transfer or processing of data by third parties (juridical)	1	1
11	Other juridical or administrative argument, referring to laws, regulations, statutes	3	0
12	Costs and efficiency	1	7
13	Other	2	9
Total number of arguments coded		20	69

³ The date of the letters was not always clearly indicated, and sometimes had to be roughly estimated or inferred from the references to other letters.

The size of the words in the two word clouds in Figure 1 indicate the relative frequency in which this argument type was used, either by the commission (left word cloud) or by the researchers (right word cloud). Both the commission and the researchers frequently use arguments that refer to response rates and response bias, but to interpret this similarity one needs the content of the letters: The codes do not suffice. The letters show, in this case, that the researchers strongly emphasized that low response rates and large response bias would be the result of the new policy to restrict access to the register and argued how this would affect the scientific quality and value of the study. Response rates and response bias were not a major argument used by the commission spontaneously: the commission only referred to these arguments in reply to the letters of the researchers. Both parties wrote about individual data protection relatively frequently, but further arguments emphasized by the commission on the one side and the researchers on the other side were remarkably different. The most important arguments repeatedly used and most strongly emphasized in the commission's letters were the rights of citizens and their freedom of choice and information. The importance of this argument was not immediately recognized by the researchers who did not use it nor refer to it in their first three letters. However, the last two letters of the researchers, and especially the letter immediately preceding the reversal of the decision by the commission, *did* mention the rights of citizens and their freedom of choice, so in the end phase they did also take up the core value of the commission. Nevertheless, the arguments which the researchers used mostly were arguments of scientific value and methodological requirements, such as response correction, sampling error, representativeness and comparability. The practical arguments given by both parties in addition to arguments based on fundamental principles also differed: whereas the researchers frequently mentioned the higher costs and inefficiency of other methods of sample drawing, the commission was mainly referring to laws and regulations that would prohibit the use of the register by the researchers. Finally, the last line of table 2 shows that the researchers used a much larger number of arguments (69) than the commission (20) across all letters. This difference reflects the fact that the researchers wrote more letters before the reversal of the decision (5) than the commission (2), probably so because they were actively fighting the decision and the commission was only responding.

2.3 Conclusion

In conclusion, the Belgium privacy commission on the one side and the Belgian researchers on the other side seemed to weigh the different aspects of the use of the person register for sample drawing very differently. Both sides kept repeating their own arguments, not taking up the central, important values of the other party well. They only seemed to find common grounds in the importance of individual data protection. However, in the later stage of the letter exchange, the researchers did make a small movement towards taking up the core value of the commission. We do not know if this had any effect on the reversal of the decision of the commission or whether the decision was reversed on other grounds. Nevertheless, we might learn from this to put ourselves in the shoes of the authorities who we want to convince and to try to think from that perspective. That could mean putting somewhat less emphasis on scientific value and methodological arguments and more on individual data protection and citizen's right in the design of the letter and memo that is to be sent to unwilling gatekeepers of registers.

It is clear that the results of this case study apply to one specific country only, but the general impression from the network events and from contact with country teams is that the

reasons given by other register owners for denying access to registers for samples often also relate to data protection, citizens' right to privacy and freedom of choice.

3. Explanation of the arguments that were incorporated in the D2.3 letter and memo

The arguments which the Belgium authorities seemed to value most were individual data protection, citizens' rights and freedom of choice, and legal aspects (laws and regulations). The researchers, however, considered the scientific and societal value of a high quality survey very important and argued with methodological requirements. In the memo which is aimed to be sent to register authorities and advising bodies, we have made a combination of these major arguments from both sides. The memo puts slightly more emphasis on the commission's main arguments but tries to link the researcher values to the direct benefit of citizens as well. The complete memo, together with the template letter, is included in annex 1. We shortly summarize each argument given in the memo here, to explain why it was chosen and written in this way:

Memo argument 1: Overview of access to registers in European countries for scientific samples

The first section of the memo summarizes our findings of SERISS D2.1 on sampling frames used across Europe (Scherpenzeel et al, 2016), and emphasizes that the majority of the person population registers in European countries do provide access to person data for the sample drawing in our surveys. This argument is given to show that this access has not been contradictory to the earlier European data protection regulations nor to many national data protection laws. In addition, it is also meant as a social validation argument, stating that "others do it too". Finally, the section explains why this situation which exists in so many countries is beneficial for science, policy makers and citizens and why extending it to all countries would increase these benefits.

Memo argument 2: The use of person register information for scientific sample drawing complies with the European General Data Protection Regulation (EU-GDPR)

The second argument is completely focused on individual data protection, laws, and regulations. This is likely to be the main argument of interest for privacy commissions and register authorities, if they think in similar ways as in the Belgium case. It especially focuses on the EU-GDPR which now applies in all EU Member States. This new regulation has made many national institutes and data protection officers uncertain, therefore this memo argument tries to re-assure that providing personal data from registers for scientific purposes is allowed under specified conditions. This section, in combination with section 3, also re-assures that these conditions are met and that data protection in our surveys has always been strictly regulated.

Memo argument 3: The citizen's right to privacy protection and freedom of choice is treasured

Citizens' rights were strongly emphasized in the letters of the Belgium privacy commission: Representatives of the commission felt the responsibility to protect the rights of citizens. This

can be a very strong argument for ministries, ethics committees and data protection officers who might be advising a register owner in general. Therefore, we have put this argument strongly forward in the memo. Citizens' rights and freedom in fact constitute an ideal link to mention the ethics and values of the academic world, as was also done in the last letter of the Belgium researchers before the reversal of the decision. Thus, this section emphasizes that citizens' rights are a key value of universities and scientific partners who are responsible for these studies. The success of surveys depends on the trust people have in scientific studies and researchers who respect the citizens' rights. Moreover, the letter states that the freedom of choice also means giving citizens the opportunity to - on a voluntary basis - take part in the survey and voice their opinions.

Memo argument 4: Citizens benefit from reliable scientific research based on representative samples from a population register

The benefit argument brings together the central value of the Belgium privacy commission (citizen rights) with that of the Belgium researchers (reliable research demands representative person samples). It explains what the methodological advantages are of a probability sample from a person register, but since methodological arguments as such are unlikely to impress register authorities, these advantages are directly linked to the need for reliable statistics which serve society and policy, and thus to the benefit of citizens.

4. Next steps: How the template letter and memo can be used

The memo and letter developed as a result of D2.3 (and made available as D2.3b) are intended to serve as a template for researchers working on high-quality national or cross-national social surveys and seeking access to national population registers for sampling purposes. The current template is for a memo sent jointly by the directors of the four major cross-national surveys that are part of SERISS and making the joint case for sample access for these surveys (scenario 3 as described below). However, the letter and memo can be used in different ways as described below. The text and signatories of the letter can be adapted accordingly to specific national contexts and data protection cultures (including for use in non-EU countries where GDPR is not the primary legislation)⁴. The letter and memo can be used in the following ways:

1. Country teams of each survey can individually use the template letter as an example for the formulation of their own specific letter with appropriate adaptations, sign it themselves or with the director of their survey, and send it to the register they want to use, including the general memo. In addition, a country team could also ask national funders, e.g. ministries, to sign and/or send the (adapted) letter.
2. Country teams across (some of) the four surveys can jointly send the letter, with specific adaptations where needed, to the register in their country signed by the directors of the (up to four) surveys and including the general memo. National funders of any of the participating surveys could also be asked to (co-)sign the letter or send it themselves.
3. The central management teams of the four surveys can together select a number of countries in which they want to make a joint, central effort to improve the sampling

⁴ Terminology should also be adapted to specific national contexts/languages where necessary, for example the name „GDPR“.

situation and send the letter and memo signed by all four directors (possibly signed in addition by the country leaders).

4. The SERISS sampling expert network can extend the letter to include more surveys and stakeholders, and send the letter and memo to Eurostat and/or try to get it published as an open letter on the Eurostat website. The aim of this would be to give a sign of European-level support for the request to give access to population registers to draw samples for high quality scientific studies. This effort could possibly be accompanied by a “white list” of such studies and would be more effective if done in combination with any of the foregoing possibilities.
5. The letter and memo will be published on the SERISS website, providing the opportunity to other survey researchers to consult it and use it as basis for their own efforts to convince register authorities, in Europe and elsewhere, to allow sample drawing from the population registers.

The sampling expert network will discuss which of the above steps are considered most feasible and effective for specific survey and/or country contexts, and establish a time line to jointly carry them out.

We kindly request any team, organisation or researcher to inform us if and how they used the letter. In this way we can keep track of its usefulness and effect. Please email to: seriss@city.ac.uk

References

Scherpenzeel, A., Maineri, A. M., Bristle, J., Pflüger, S.-M., Mindarova, I., Butt, S., . . . Luijkx, R. (2017). Report on the use of sampling frames in European studies. Deliverable 2.1 of the SERISS project funded under the *European Union's Horizon 2020 research and innovation programme*(GA No: 654221).

Scherpenzeel (2018). *Letter to request access for sampling for European survey infrastructures*. Deliverable 2.3b 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