Sample Design Summary: ESS Round 10

Country:	Slovakia (SK)
NC:	Denisa Fedakova (dfedak@saske.sk), Michal Kentos (kentos@saske.sk)
Other Experts:	<name> (<email address>)
Survey Institute:	<institute name>
Sampling Expert:	Olena Kaminska (olena@essex.ac.uk)
Country Contact:	Achim Koch (Achim.Koch@gesis.org)
Reference Survey:	ESS 9
Date:	19/10/2020

Status:

- [ ] Pre sign-off
- [ ] Signed off
- [x] Post sign-off amendment
- [ ] Final (post-field work)

1.1 Target Population

Number of residents aged 15 or older in the country:	4594153
Source and reference date:	Statistical Office of the Slovak Republic, December 31th 2019

1.2 Population Coverage

All persons aged 15 and over who are residents of Slovakia. Some population groups will be excluded from the target population: people living in localities with less than 10 dwellings, homeless people, people who are residents of Slovakia but working most of time outside the country. There are 1320 dwellings in localities with under 10 dwellings, which constitutes 0.062% of all dwellings in the country (2,143,425).

2. Summary of the Sample Design

This is a two-domain design with 4 stage design used in larger municipalities and 3 stage design used in smaller municipalities.

Sampling Domain 1: Larger Municipalities (usually over 1000 dwellings)

The sampling design for this domain has a 4-stage design.

At the first stage municipalities are selected by stratified sample. The stratification is based on geographical region (8) and municipality (15+ population habitat) size classes (5). Within the strata the municipalities are selected by a systematic sample with probability proportional to their 15+ population.

At the second stage, sampling points (streets), are selected from each sampled municipality. One sampling point is selected from each municipality for each time the municipality has been selected by the systematic sampling algorithm at the first sampling stage. The sampling points are selected with probability proportional to the number of their dwellings (mail boxes).

For each selected sampling point a complete listing is made of all the dwellings in the sampling point (each dwelling in a building will be listed separately). At the third sampling stage 7 dwellings are selected from each selected sampling point by a simple random sample.

At the fourth sampling stage one 15+ person will be selected within a dwelling using a Kish Grid (FMS).

Sampling Domain 2: Smaller Municipalities (usually under 1000 dwellings)

In smaller municipalities that usually have under 1000 dwellings a three-stage sample design will be implemented.

At the first stage municipalities are selected by stratified sample. The stratification is based on geographical regions (8). Within the strata the municipalities are selected by a systematic sample with probability proportional to their 15+ population.

At the second stage in each sampled municipality 7 dwellings are selected through SRS.

At the third stage one person is selected from each sampled dwelling using a Kish Grid (FMS).

3. Sample Design Details

3.1 Sampling Domain 1: Larger Municipalities

Around 81.5% of the population lives in larger municipalities

First Sampling Stage

unit: Large municipalities

frame: A list of all large municipalities in Slovakia

size: 266

strata: 8 regions × 5 municipality (habitat) size classes. There are 31 non-empty strata because in every region at least 1 one municipality size classes is empty. Within each stratum, the number of localities (municipalities) is known.

allocation: The allocation (see Appendix Table 2) is proportional to the size of population 15+ in the strata (see Appendix Table 3 for the population sizes of the strata and Table 4 for the allocation of the sample size).

algorithm: A systematic sample with probability proportional to the 15+ population size of municipalities. The algorithm can (and will in some strata) select a municipality more than one time. Number of distinct municipalities selected in each stratum is shown in Table 5. In total, 446 selections are made, corresponding to 266 distinct municipalities. The municipalities were ordered according to statistical codes of municipalities (regions – districts – municipalities). Municipalities within districts are mostly ordered alphabetically.

Second Sampling Stage

unit: Sampling points. Sampling points are streets or combination of geographically adjacent streets if a street has fewer than 7 dwellings.

frame: All sampling points in the selected municipalities

size: 446

strata: Where multiple sampling points are selected – implicit stratification by the dwelling size of a sampling point

allocation: Not applicable

algorithm: Where multiple sampling points need to be selected in one municipality systematic sampling with probability proportional to the number of dwellings in the sampling points will be used. In municipalities where one sampling point needs to be selected a random sample with probability proportional to the number of dwellings will be used.

Third Sampling Stage

unit: Dwellings

frame: All dwellings within selected sampling points are listed by NC staff in the field (sampling and fieldwork will be separated). This list is then used as sampling frame for dwellings.

size: 3122, 7 dwellings are sampled from each selected sampling point

strata: No stratification

allocation: Not applicable

algorithm: Simple Random Sample

Fourth Sampling Stage

unit: Persons

frame: All eligible persons within selected dwellings

size: 3122, one person is selected from each sampled dwelling

strata: No stratification

allocation: Not applicable

algorithm: Kish-Grid (FMS)

3.2 Sampling Domain 2: Small Municipalities

Around 18.5% of the population lives in small municipalities (with under 1000 dwellings in each). This corresponds to the first column in Tables 2-5 in Appendix.

First Sampling Stage

unit: Small municipalities

frame: A list of all small municipalities in Slovakia

size: 97

strata: 8 regions. Within each stratum, the number of localities (municipalities) is known.

allocation: The allocation (see Appendix Table 2) is proportional to the size of population 15+ in the strata (see Appendix Table 3 for the population sizes of the strata and Table 4 for the allocation of the sample size).

algorithm: A systematic sample with probability proportional to the 15+ population size of municipalities. Number of distinct municipalities selected in each stratum is shown in Table 5. In total, 97 selections are made, corresponding to 97 distinct municipalities. The municipalities were ordered according to statistical codes of municipalities (regions – districts – municipalities). Municipalities within districts are mostly ordered alphabetically.

Second Sampling Stage

unit: Dwellings

frame: Smaller municipalities will be treated as one sampling point each – so all dwellings in the municipality will be listed and used as a sampling frame.

size: 679, 7 dwellings are sampled from each selected municipality

strata: No stratification

allocation: Not applicable

algorithm: Simple Random Sample

Third Sampling Stage

unit: Persons

frame: All eligible persons within selected dwellings

size: 679, one person is selected from each sampled dwelling

strata: No stratification

allocation: Not applicable

algorithm: Kish-Grid (FMS)

4. Planning the Sample Size

History of Planned and Realised Values

ESS	$p.\overline{b}$	$\overline{b}$	$p.\rho$	ρ	$p.\text{Deff}$	Deff	$p.\text{Deff}_c$	Deff_c	$p.\text{Deff}_p$	Deff_p
5	7.00	7.42	0.02	0.158	1.462	2.504	1.12	2.014	1.305	1.243
6	7.00	7.39	0.05	0.171	1.722	4.020	1.32	2.093	1.305	1.921
9	5.18	2.71	0.1	0.16	2.127	1.481	1.418	1.171	1.5	1.265
9, domain 1 (large)		2.67		0.16						
9, domain 2 (small)		2.9		0.16						

ESS	$p.\text{rr}$	rr	$p.\text{ri}$	ri	$p.\text{n}_{gross}$	n_{gross}	$p.\text{n}_{net}$	n_{net}	$p.\text{n}_{eff}$	n_{eff}
5	0.70	0.747	0.02	0.007	2500	2500	1750	1856	1200	741
6	0.70	0.739	0.00	0.002	2506	2506	1754	1847	1019	459
9	0.74	0.395	0	0.02	2800	2800	2072	1083	974	732
9, domain 1 (large)		0.399		0.046						
9, domain 2 (small)		0.437		0.051						

Parameters of the Planned Gross Sample Size

Domain	Achieved interviews per cluster ($\bar{b}$)	Intraclass Correlation Coefficient (ρ)	Design Effect due to Selection Probabilities (Deff_p)	Response Rate (rr)	Ineligible Rate (ri)	Effective Sample Size (n_{eff})
1 Large	3.61	0.16	1.265	0.54	0.046	880
2 Small	3.92	0.16	1.265	0.59	0.051	200
Total	3.66	0.16	1.265	0.55	0.047	1080

Design Effect

Domain 1

$$\begin{aligned} \text{Deff}_c &= 1 + (\bar{b} - 1) \times \rho \\ &= 1 + (3.61 - 1) \times 0.16 \\ &= 1.417 \end{aligned}$$

$$\text{Deff}_p = 1.265$$

$$\begin{aligned} \text{Deff} &= \text{Deff}_p \times \text{Deff}_c \\ &= 1.792 \end{aligned}$$

**results have been rounded to 3 d.p.s*

Domain 2

$$\begin{aligned} \text{Deff}_c &= 1 + (\bar{b} - 1) \times \rho \\ &= 1 + (3.92 - 1) \times 0.16 \\ &= 1.467^* \end{aligned}$$

$$\text{Deff}_p = 1.265^*$$

$$\text{Deff} = \text{Deff}_p \times \text{Deff}_c = 1.856^{*}$$

** results have been rounded to 3 d.p.s*

Gross Sample Size

Domain 1

$$\text{Min. } n_{net} = \text{Deff} \cdot n_{eff} = 1.792 \times 1223 = 2191^{**}$$

$$\text{Min. } n_{gross} = \frac{n_{net}}{\text{rr} \times (1 - \text{ri})} = \frac{2191}{0.54 \times (1 - 0.046)} = 4253^{**}$$

$$\text{Planned } n_{gross} = 3122$$

*** results have been rounded to 0 d.p.s.*

Domain 2

$$\text{Min. } n_{net} = \text{Deff} \cdot n_{eff} = 1.856 \times 278 = 516^{**}$$

$$\text{Min. } n_{gross} = \frac{n_{net}}{\text{rr} \times (1 - \text{ri})} = \frac{516}{0.59 \times (1 - 0.051)}$$

	=	922^{**}
---	---	---
Planned n_{gross}	=	679
---	---	---

** *results have been rounded to 0 d.p.s.*

Total Planned n_{gross} = $3122 + 679 = 3801$

Remarks

To obtain the required 1,500 effective sample size a 5175 gross sample size would be needed. The budget restrictions allow for a maximum of 3,801 gross sample size. This leads to the expected 1,080 total effective sample size. The domain assignment takes into account different values of RR between domains (0.54 in domain 1 and 0.59 in domain 2), ineligibility rate (0.046 and 0.051) and design effects (1.792 and 1.856). These were estimated using data from the round 9. The RR was scaled up to an average expected RR of 0.55 given the change in data collection organization. A special permission was obtained for lower effective sample size, neff, of 1,080.

5. Sampling Design Data File (SDDF)

Variables to be included in the SDDF

Variable	Description/Value
IDNO	Respondent identification number
CNTRY	SK
PROB1	Expected selection frequency of municipalities (i.e. the PSUs)
PROB2	Conditional inclusion probability of sampling points within municipalities
PROB3	Conditional inclusion probability of dwellings within sampling points
PROB4	Conditional inclusion probability of persons within dwellings
PSU	Municipality identification number
SSU	Sampling point identification number
STRATEX1	An indicator of 39 strata
DOMAIN	An indicator of whether municipality is large (over 1000 dwellings) or small (under 1000 dwellings)
STRATIM1	Serial number of the order of municipalities for systematic sampling at the first sampling stage
POPSIZE	Population 15+ size of municipalities
STRTVAL1	Number of dwellings per street for larger municipalities / per smaller municipality
OUTCOME	Summary field outcome. Value = 1 for respondent (data in questionnaire data file), 2 for eligible non-respondent (refusal, non-contact, unable to be interviewed due to language, health, etc), 3 for ineligible (outside of survey population: died, moved abroad, vacant address, etc)
REGION	Indicator of the 8 regions used for stratification
MUNSIZE	Municipality size (number of dwellings) – 6 categories
POPDEN	population density of municipality
AGE1	proportion younger than 25 years within municipality
AGE2	proportion 25 – 64 years within municipality
AGE3	proportion 65 years or older within municipality
NATION	proportion Slovakian nationality within municipality

Probabilities of Selection

Domain 1: Large Municipalities

For large municipalities:

1. $\text{PROB1}_{i|h} = \frac{N_{ih}}{N_h} * m_h$
2. $\text{PROB2}_{j|ih} = \frac{b_{jih}}{B_{jih}}$
3. $\text{PROB3}_{l|jih} = \frac{7}{B_{jih}}$
4. $\text{PROB4}_{k|ljih} = \frac{1}{N_{ljih}}$

Domain 2: Small Municipalities

For small municipalities:

1. $\text{PROB1}_{i|h} = \frac{N_{ih}}{N_h} * m_h$
2. $\text{PROB2}_{j|ih} = 1$
3. $\text{PROB3}_{l|jih} = \frac{7}{B_{jih}}$
4. $\text{PROB4}_{k|ljih} = \frac{1}{N_{ljih}}$

Statistic	Description
N_h	Population 15+ size of h-th stratum
N_{ih}	Population 15+ size of the i-th municipality of the h-th stratum
m_h	Number of selected sampling points within the h-th stratum
B_{jih}	Number of all dwellings in the j-th sampling point of the i-th municipality of the h-th stratum
b_{jih}	Number of selected dwellings in the j-th sampling point of the i-th municipality of the h-th stratum
N_{ljih}	Number of eligible persons of the l-th dwelling of the j-th sampling point of the i-th municipality of the h-th stratum

Appendix

Table 2: Distribution of population aged 15+ (in 1000) by regions and municipality size categories

Region \ Municipality Size (dwellings)	Small -999	Large -4999	-18399	-49999	-99999	100000+	Total
Bratislava region	13,1	90,1	64,9	18,7	0,0	365,8	552,6
Trnava region	75,4	199,9	108,1	43,8	55,4	0,0	482,6
Trenčín region	84,3	157,1	100,2	161,7	0,0	0,0	503,3
Nitra region	109,6	212,8	62,5	131,2	65,5	0,0	581,5
Žilina region	85,6	222,2	89,4	116,8	68,3	0,0	582,2
Banská Bystrica region	147,7	132,5	121,9	81,1	67,3	0,0	550,5
Prešov region	185,6	184,3	113,0	119,0	75,8	0,0	677,6
Košice region	149,3	155,4	70,3	84,5	0,0	204,3	663,7
Total	850,5	1354,3	730,3	756,8	332,2	570,0	4594,2

Source: Statistical Office of the Slovak Republic, December 31th 2019

Note: The numbers in this table represent 15+ population size. The municipality size categories (e.g. -999) represent number of dwellings in a municipality.

Table 3: Number of municipalities by regions and size categories

Region \ Municipality Size (dwellings)	Small -999	Large - 4999	- 18399	- 49999	- 99999	100000+	Total
Bratislava region	22	42	7	1	0	1	73
Trnava region	134	105	9	2	1	0	251
Trenčín region	177	85	9	5	0	0	276
Nitra region	219	121	8	5	1	0	354
Žilina region	185	114	11	4	1	0	315
Banská Bystrica region	422	77	13	3	1	0	516
Prešov region	542	105	12	4	1	0	664
Košice region	338	89	9	3	0	1	440
Total	2039	738	78	27	5	2	2889

Source: Statistical Office of the Slovak Republic, December 31th 2019

Table 4: Number of selected sampling Points in regions and municipality size categories

Region \ Municipality Size (dwellings)	Small	Large				
	-999	-4999	-18399	-49999	-99999	100000+	Total
Bratislava region	1	11	8	2	0	44	66
Trnava region	9	24	13	5	7	0	58
Trencín region	10	19	12	19	0	0	60
Nitra region	12	25	7	16	8	0	68
Žilina region	10	26	11	14	8	0	69
Banská Bystrica region	17	16	15	10	8	0	66
Prešov region	21	22	13	14	9	0	79
Košice region	17	18	8	10	0	24	77
Total	97	161	87	90	40	68	543

Source: Statistical Office of the Slovak Republic, December 31th 2019

Table 5: Number of selected municipalities in regions and municipality size categories

Region \ Municipality Size (dwellings)	Small	Large				
	-999	-4999	-18399	-49999	-99999	100000+	Total
Bratislava region	1	11	6	1	0	1	20
Trnava region	9	24	8	2	1	0	44
Trencín region	10	19	8	5	0	0	42
Nitra region	12	25	7	5	1	0	50
Žilina region	10	26	10	4	1	0	51
Banská Bystrica region	17	16	12	3	1	0	49
Prešov region	21	22	12	4	1	0	60
Košice region	17	18	8	3	0	1	47
Total	97	161	71	27	5	2	363

Source: Statistical Office of the Slovak Republic, December 31th 2019