Assessment of socio-demographic sample composition in ESS Round 8 and 9

Achim Koch

Roberto Briceno-Rosas

GESIS Leibniz Institute for the Social Sciences, Mannheim

achim.koch@gesis.org

roberto.briceno-rosas@gesis.org

May 2021

Assessment of socio-demographic sample composition

in ESS Round 8 and 9

– Executive Summary –

Approach:

- Comparison of six demographic variable distributions (gender, age, marital status, work status, nationality, and household size) from ESS 8 and 9 with external benchmark data from the EU LFS.
- Comparison of results with those of analogous assessments for ESS 5, 6, and 7.
- Comparison of sample composition without and with applying ESS post-stratification weights.

Results:

- Using LFS data as an external benchmark, we find indications of misrepresentation of demographic groups in the samples of the ESS. The level of misrepresentation varies between countries and variables.
- The basic patterns of misrepresentation in ESS 8 and 9 are rather similar to the patterns observed for previous rounds of ESS (e.g. underrepresentation of younger age groups and of non-nationals; overrepresentation of females and of married persons).
- Using the ESS post-stratification weight normally decreases the level of misrepresentation. The size of the reduction differs between countries and variables. In a few cases, however, applying the post-stratification weight increases the dissimilarity between ESS and LSF.

Interpretation and recommendation:

- It seems natural that differential response propensities of demographic subgroups are the main factor behind the patterns of misrepresentation observed. A specific subgroup will be underrepresented in a country if that group is particularly difficult to contact or less willing to consent with a survey request.
- In addition, we cannot preclude that the behavior of interviewers also affects sample composition. If interviewers, for instance, do not always adhere to the survey standards set in the ESS (e.g. by not following the ESS contact schedule or by substituting reluctant target persons by persons more willing to participate), this might also contribute to demographic misrepresentation.
- Although applying post-stratification weights is a cost-efficient approach to correct for demographic misrepresentation, it is not a perfect remedy for dealing with the patterns of misrepresentation observed in the ESS.
- Aiming for balanced response rates during fieldwork therefore continues to be an important goal. ESS National Coordinators should be aware of the specific patterns of misrepresentation in their country. They should discuss potential reasons of mis-representation and consider measures to improve in the upcoming round. To that end, the present results were fed back to the countries of ESS round 10.

Assessment of socio-demographic sample composition in ESS Round 8 and 9[1]

Achim Koch & Roberto Briceno-Rosas

GESIS – Leibniz Institute for the Social Sciences, Mannheim/Germany
May 2021

Contents

[1] The CST of the ESS requests that the following citation for this document should be used: Koch, A. & Briceno-Rosas, R. (2021). Assessment of socio-demographic sample composition in ESS Round 8 and 9. Mannheim: European Social Survey, GESIS.

1. Introduction

The European Social Survey (ESS) is an academically driven cross-national survey that has been conducted every two years across Europe since 2002. The ESS aims to produce high-quality data on social structure, attitudes, values, and behaviour patterns in Europe. Much emphasis is placed on the standardisation of survey methods and procedures across countries and over time. Each country implementing the ESS has to follow detailed requirements that are laid down in the ESS Survey Specifications. These standards cover the whole survey life cycle. They refer to sampling, questionnaire translation, data collection, and data preparation and delivery. As regards sampling, for instance, the ESS requires that only strict probability samples should be used; quota sampling and substitution are not allowed. Each country is required to achieve an effective sample size of 1,500 completed interviews, taking into account potential design effects due to the clustering of the sample and/or the variation in inclusion probabilities. Regarding data collection, the ESS specifies – among other things – that face-to-face interviewing is the only mode allowed. Targets are set for the response rate (70%)[2] and the noncontact rate (3% maximum). The fieldwork period is specified (September until December of the survey year), the personal briefing of interviewers is required, and a detailed call schedule for the interviewers is laid down.

The purpose of setting these standards is to achieve accurate and comparable survey data. An important aspect of survey quality refers to the quality of the realised samples in terms of representation of the target population. The sample in each ESS country should reflect the target population of the ESS adequately, which means that sampling, coverage, and nonresponse errors should be minimised. Quality control activities in the ESS are mainly directed at compliance with the prescribed data collection procedures. In each survey round, for instance, it is checked whether a country achieved the target response rate, whether the interviewers were adequately briefed, whether the call schedule was adhered to, etc. The (implicit) assumption is that a country that follows the ESS survey procedures and achieves a high response rate will also achieve a sample of good quality.

In the present paper we assess empirically to what extent ESS samples represent the ESS target population. We analyse the socio-demographic sample composition in ESS countries by comparing ESS variable distributions with suitable external benchmark data, for which we choose the European Union Labour Force Survey (LFS). The analyses refer to ESS 8 and ESS 9, which were fielded in the years 2016/17 and 2018/19, respectively. Our analyses provide an indication of the degree of over-/underrepresentation of certain demographic subgroups in ESS samples. In the past, a similar analysis has been conducted for ESS 5, ESS 6 and ESS 7 (Koch et al. 2014; Koch 2016; Koch 2018). The present analysis carries on this exercise.

[2] The ESS 9 Specifications for participating countries state: "In addition, a minimum target response … has been set at 70%. Ideally, all countries should aim for this 70%. Acknowledging – based on previous experiences in the ESS – that reaching this 70% target response rate is very challenging in many countries, all countries are expected to plan and budget fieldwork in order to reach a response rate higher than in the previous round." (European Social Survey 2018, p. 31)

2. Assessment with external benchmark data

The comparison of survey results with independent and more accurate information about the population parameters is a well-known method to analyse sample quality and the degree of nonresponse bias (Groves 2006). For this approach, no information at the individual level is required. There needs to be another survey or administrative record system containing estimates of variables similar to those being produced from the survey. Then, the survey estimates can be benchmarked with information from the other data source, the so-called gold standard. The difference between estimates from the survey and the other data source can be used as an indicator of bias.

The advantage of this method is that it is in theory relatively simple to implement. Usually, the method is not too expensive since it does not require collecting additional data. The drawback is that normally only a limited set of variables can be compared. In order to draw valid conclusions about nonresponse bias, the benchmark data have to be quite accurate, i.e. they should not be severely affected by, for instance, measurement or nonresponse errors. In addition, the measurements of the relevant variables should match closely between the two data sources (equivalent measurements). Both data sources have to refer to the same target population, and also the reference period should be as close as possible. If these conditions hold, differences between the survey data and the benchmark data might arise from three sources of error: sampling error, coverage error, and nonresponse error.

It goes without saying that no benchmark information is available for the ESS key survey variables – this is the reason why the ESS exists! Comparisons have to be restricted to several socio-demographic variables. The results, however, are important beyond these variables. Socio-demographic characteristics are intrinsically important since they are – potentially – related to many attitudes and behaviours. For this reason, some of these variables are used to construct post-stratification weights. Since 2014, post-stratification weights are also provided for the ESS (European Social Survey 2014; Lynn & Anghelescu 2018).

For a cross-national survey like the ESS, the most promising candidate to act as a valid standard for such a comparison is the European Union Labour Force Survey (LFS). Most of the countries that participate in the ESS also conduct the yearly Labour Force Survey for Eurostat.

3. The European Union Labour Force Survey

The European Union Labour Force Survey (LFS) is a large sample survey among residents in private households in Europe.[3] It is an important source for European statistics about the situation and trends in the EU labour market. The LFS is currently fielded in 35 European countries. These include all Member States of the European Union, the United Kingdom, three EFTA countries (Iceland, Norway, and Switzerland), and four EU candidate countries (Montenegro, North Macedonia, Serbia, and Turkey). The sampling units are dwellings, households or individuals depending on the country-specific sampling frames. Each quarter, more than 1.7 million interviews are conducted throughout the participating countries to obtain statistical information for some 100 variables. The quarterly sampling rates in the countries vary between 0.2% and 1.7%.

[3] http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey

The EU LFS is conducted by the National Statistical Institutes across Europe and is centrally processed by Eurostat (for details of national implementation see Eurostat 2018a, 2018b, 2019a, 2020). The National Statistical Institutes of the Member States are responsible for designing national questionnaires, drawing the sample, conducting interviews and forwarding results to the Commission (Eurostat) in accordance with a common coding scheme. As a rule, the data are collected by interviewing the sampled individuals directly, but proxy interviews (through a responsible person in the household) are also possible. Moreover, part of the data can also be supplied by equivalent information from alternative sources, such as e.g. administrative registers (mainly social insurance records and population registers).

Table 1: Timing of fieldwork in ESS 8 and ESS 9

Country	ESS 8: % of interviews completed in year		ESS 9: % of interviews completed in year	
	2016	2017	2018	2019/2020
AT	100.0		93.0	7.0
BE	92.1	7.9	86.9	13.1
BG			100.0	
CH	94.7	5.3	95.8	4.2
CY			70.9	29.1
CZ	100.0		48.2	51.8
DE	88.4	11.6	77.9	22.1
EE	89.5	10.5	89.9	10.1
ES		100.0		100.0
FI	80.5	19.5	85.6	14.4
FR	70.9	29.1	56.7	43.3
HR				100.0
HU		100.0		100.0
IE	6.5	93.5	18.1	81.9
IS	27.4	72.6		
IT		100.0	3.9	96.1
LT		100.0		100.0
LV				100.0
NL	87.1	12.9	93.5	6.5
NO	99.9	0.1	44.0	56.0
PL	81.7	18.3	65.5	34.5
PT	19.1	80.9	10.3	89.7
SE	91.6	8.4	67.1	32.9
SI	99.5	0.5	96.5	3.5
SK				100.0
UK	91.1	8.9	87.3	12.7

Source: ESS 8, ed. 2.1, variables ‘inwyys’ and ‘inwyye’ (start/end of interview, year)
ESS 9, ed. 02, variable ‘inwyys’ (start of interview, year)
Highlighted: Countries with all interviews or the majority of interviews completed in the following year(s)

The present comparison with the LFS is conducted for the eighth and ninth survey round of ESS. 23 countries in total participated in ESS 8. In ESS 9, 27 countries are included in the

second data edition which was used for the present analysis.[4] Among these countries, 21 (ESS 8) and 25 (ESS 9) countries also participated in the LFS. Only Israel and Russia (ESS 8) and Montenegro and Serbia (ESS 9) were not part of the LFS and had to be excluded from our analyses. As a rule, fieldwork in each ESS country should take place between September and December of the survey year (i.e. in 2016 for ESS 8, and 2018 for ESS 9). Unfortunately, not all countries managed to adhere to this schedule. In seven countries in ESS 8, and 11 countries in ESS 9, most or all interviews were completed only in 2017 or 2019/2020, respectively (see Table 1).

Table 2: Basic characteristics of LFS 2016, LFS 2018, ESS 8 and ESS 9*

Country	LFS 2016			LFS 2018			ESS 8	ESS 9
	Participation compulsory	Response rate (%)	Proxy rate among 15-74 years old respondents (%)	Participation compulsory	Response rate (%)	Proxy rate among 15-74 years old respondents (%)	Response rate (%)	Response rate (%)
AT	yes	94.7	24.3	yes	93.0	27.9	52.2	50.8
BE	yes	71.6	17.8	yes	81.0	27.0	56.8	57.6
BG				no	80.1	30.0		69.4
CH	no	80.8	3.0	no	79.7	2.7	52.2	51.8
CY				yes	95.7	38.1		53.4
CZ	no	79.8	42.3	no	77.6	43.0	68.4	67.4
DE	yes	97.4	23.9	yes	97.3	23.2	30.6	27.6
EE	no	69.8	36.8	no	71.9	16.3	64.5	62.7
ES	yes	87.4	51.8	yes	84.4	51.0	67.7	53.8
FI	no	69.6	4.1	no	65.6	3.9	56.8	51.8
FR	yes	80.7	26.6	yes	79.7	27.1	50.8	48.1
HR				no	57.6	51.5		43.2
HU	no	80.7	42.6	no	75.5	41.6	42.4	40.7
IE	no	72.7	50.2	no	63.4	46.8	64.5	62.0
IS	no	73.2	0.8				44.0	
IT	yes	86.7	20.3	yes	85.7	33.1	49.5	51.9
LT	no	78.7	35.8	no	78.4	36.1	64.0	59.2
LV				no	65.2	38.1		38.9
NL	no	53.0	45.5	no	50.6	45.0	53.0	49.6
NO	yes	81.6	17.2	yes	84.3	17.3	52.9	43.3
PL	no	62.4	37.2	no	57.8	36.5	69.4	60.4
PT	yes	84.3	48.3	yes	83.6	48.2	45.1	34.9
SE	no	57.0	3.1	no	52.8	2.8	42.1	39.0
SI	no	78.6	49.7	no	78.7	54.2	55.8	64.1
SK				yes	82.4	51.1		39.6
UK	no	55.4	34.9	no	48.9	35.2	43.0	41.0
Mean		76.0	29.3		74.8	33.1	53.6	50.5

* 26 countries which took part in ESS 8 or in ESS 9, and in LFS
Source LFS: Eurostat 2018a, 2019a;
Source ESS 8: Wuyts & Loosveldt 2019;
Source ESS 9: ESS website: Notes on data and fieldwork (09.09.2020)

[4] In the 3rd data edition of ESS 9, another two countries were added (Denmark and Iceland). Due to time restrictions, these countries could not be included in the present analysis.

Table 2 documents a few basic parameters (participation compulsory, response rate[5] and rate of proxy interviews) for LFS 2016 and 2018. For the sake of comparison, the response rates of ESS 8 and ESS 9 are also included.

Among the 26 countries, participation in the LFS was mandatory in 10 countries (see Table 2). The LFS response rates vary between 48.9% (United Kingdom in 2018) and 97.4% (Germany in 2016). Accordingly, the LFS, too, has a severe nonresponse problem in some countries. The consequences for the nonresponse error of the LFS cannot be assessed here. However, two points can be made in favour of still using LFS as a benchmark for the ESS. First, in each country except the Netherlands and Poland, the LFS response rate is in both survey years (often considerably) higher than the ESS response rate. The average LFS response rate among the countries analysed is 76.0% in 2016, and 74.8% in 2018. The respective rates in ESS are 53.6% and 50.5%. Second, it has to be taken into account that the LFS data itself are weighted to adhere to the population distribution. (Nearly) all countries used population information on gender, age and region in their weighting procedure (Eurostat 2018a, 2019a). Several LFS countries included additional variables (like employment status or nationality). Thus, at least the distributions of these variables should validly reflect the countries' population.

Apart from the question of nonresponse error, the measurement error properties of the LFS data might also be questioned. In some LFS countries, a large number of proxy interviews were conducted. The percentage of proxy interviews varies between 0.8% (Iceland in 2016) and 54.2% (Slovenia in 2018). On average across all countries, around 30% of all interviews in LFS 2016 and LFS 2018 were proxy interviews. We cannot empirically assess what this means for the quality of the LFS data. However, it seems justifiable to assume that the basic demographic information which we use for our analyses will not noticeably be impaired by this problem (Köhne-Finster & Lingnau 2009; Thomsen & Villund 2011; Zühlke 2008).

4. Data and variables

For our analyses we used ESS round 8 (edition 2.1) and round 9 (edition 2.0) data[6], and anonymised EU LFS 2016 (edition 2018) and 2018 (edition 2019) data[7]. We used the yearly datasets of LFS, including the so-called 'structural' variables (Eurostat 2019b). Comparisons between ESS and LFS were possible for variables which were either measured in an identical way or, if this was not the case, where the measurements could be recoded to a common standard. This was true for six variables: gender, age, marital status, work status, nationality, and household size. We deliberately did not include a variable like education, which is difficult to measure in a comparable way in a cross-national context (Ortmanns & Schneider 2016). Table 3 shows the variables and the respective categories which we distinguished, plus their source variables in ESS and LFS.

[5] In the LFS most countries calculate response rates on the household level, only in a minority of countries response rates are calculated on the person level (which is the standard in ESS).

[6] European Social Survey Round 8 Data (2016). Data file edition 2.1. NSD – Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 9 Data (2018). Data file edition 2.0. NSD – Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC.
The ESS ERIC, Core Scientific Team (CST) and the producers bear no responsibility for the uses of the ESS data, or for interpretations or inferences based on these uses.

[7] All results and conclusions are those of the authors and not those of Eurostat, the European Commission or any of the national authorities whose data have been used.

Table 3: Variables of the ESS – LFS comparison

Variable	**Categories**	**ESS source variable**	**LFS source variable**
Gender	• Male • Female	gndr	sex
Age	• 15-24 years • 25-34 years • 35-44 years • 45-54 years • 55-64 years • 65-74 years • 75 years and older	agea (recoded)	age (recoded)
Marital status	• Not married • Married (incl. registered partnership)	maritalb (3-6 = 0) (1-2 = 1)	marstat (0-1 = 0) (2 = 1)
Work status	• Not in paid work in the last 7 days • In paid work (for at least one hour) in the last 7 days	pdwrk + crpdwk	wstator (3-5 = 0) (1-2 = 1)
Nationality	• National of country • No national of country	ctzcntr (1 = 0) (2 = 1)	national (non-nationals recoded in one category) (0 = 0) (1-21 = 1)
Household size	Respondent lives in household comprising • 1 person • 2 persons • 3 persons • 4 persons • 5 or more persons	hhmmb (recoded)	hhnbpers (recoded)

The ESS interviews persons aged 15 years and over resident within private households, regardless of their nationality, citizenship or language. In order to achieve comparable target populations, we excluded persons under 15 years in the LFS. In addition, persons living in an institutional household (which were surveyed in a few LFS countries) were excluded. In Iceland, Norway, and Sweden, LFS data are only available for persons aged 74 years or younger. The LFS sample in Estonia does not include persons 75 years and older living alone in a household. For these four countries, we restricted the ESS (and LFS) analyses to persons aged 74 years or younger.

ESS data were weighted with the design weight (DWEIGHT). This weight corrects for differences in selection probabilities between sampling units in a country. The design weights are computed as normed inverse of the inclusion probabilities. LFS data were weighted with the standard weight variable COEFF, as recommended by Eurostat. COEFF corrects for differences in selection probabilities. In addition, it includes a post-stratification adjustment to adapt the LFS data to known population characteristics. In (nearly) all LFS countries, data on gender, age, and region were used for the adjustment. A number of countries included additional data in weighting, like information on unemployment or nationality (see Eurostat 2018a, 2019a). Using weighted data for the LFS thus should reduce both sampling errors and errors due to nonresponse or noncoverage – at least for the variables included in the weighting procedure.

When determining the categorisation of the variables, we tried to make sure that the proportions of persons in the different categories were of a reasonable size in all countries. Apart from one variable (nationality), this could be achieved. Table 4 shows the minimum and maximum values among the 21 countries in the LFS 2016 and the 25 countries in the LFS 2018 for the variables and categories included in the analysis. It is noteworthy that the proportion of non-nationals is very low in some countries. In five out of the 21 countries in 2016, and in seven out of the 25 countries in LFS 2018, the percentage of non-nationals is less than 2.0%. These are Bulgaria, Croatia, Czech Republic, Hungary, Lithuania, Poland, Portugal (only in 2016), and Slovakia.

In addition, it should be noted that in the standard LFS data files no information on household size has been made available for five countries (Finland, Iceland, Norway, Sweden, and Switzerland).[8] Thus, the analyses with the variable household size had to be restricted to 16 countries in ESS 8, and 21 countries in ESS 9. Due to a routing error, no information on marital status is available for Portugal in ESS 8. The same applies to Latvia in ESS 9. As regards the variable work status, all persons 75 years and older in Finland, Hungary and Latvia were classified as 'not in work' in the LFS. We did the same for the respective data in ESS 8 and ESS 9.

Table 4: Minimum and maximum values of the analysed variables
(21 countries in LFS 2016; 25 countries in LFS 2018)

Variable / category	**LFS 2016**			**LFS 2018**		
	# of countries	**Minimum (%)**	**Maximum (%)**	**# of countries**	**Minimum (%)**	**Maximum (%)**
Female	21	49.0	54.9	25	49.0	55.2
15-24 y.	21	11.3	16.8	25	10.5	16.4
25-34 y.	21	13.0	19.7	25	12.7	20.2
35-44 y.	21	14.2	20.0	25	14.2	19.9
45-54 y.	21	14.5	19.0	25	14.6	18.7
55-64 y.	21	13.5	17.6	25	13.7	17.8
65-74 y.	21	9.9	15.1	25	10.3	14.9
75+ y.	17	7.0	12.8	22	7.1	13.1
Married	21	40.5	60.0	25	40.3	60.3
In paid work	21	43.7	81.2	25	44.6	68.5
Non-national	21	0.2	23.8	25	0.2	24.1
1p-hh	16	10.0	24.6	21	9.4	26.3
2p-hh	16	26.4	38.3	21	22.0	38.4
3p-hh	16	14.9	29.0	21	14.8	28.4
4p-hh	16	14.8	24.5	21	14.4	24.9
5+p-hh	16	6.3	19.6	21	6.3	21.7

[8] For Iceland, Norway, and Switzerland, no household information at all is available in the LFS. For Finland and Sweden, data is available only in separate country-specific files for a special household subsample. See Eurostat 2019b, p. 41.

5. Patterns of misrepresentation

Which socio-demographic groups are over- or underrepresented in the ESS samples? Table 5 and Table 6 display the direction and size of differences between ESS and LFS estimates for the six variables included in our analyses. For dichotomous variables (gender, marital status, work status, nationality), the differences for only one category are shown. For age and household size, differences for all categories are provided. Green cells indicate an overrepresentation of the respective category in a country in the ESS, while red cells indicate an underrepresentation. Thus, it can easily be checked whether the patterns of misrepresentation are similar across countries.

To provide an indication of whether the observed differences between ESS and LFS are within the limits of sampling error, we calculated 95% confidence intervals for the ESS estimates. We incorporated sample design indicators (PSU, STRATUM, and DWEIGHT) into the analyses in order to obtain design-unbiased estimates of standard errors (Kaminska 2020; Lynn 2019). The confidence intervals were estimated using the complex sample procedure of SPSS. When the confidence intervals do not overlap with the percentage from the LFS, we interpret this as an indication of a significant over- or underrepresentation with respect to that specific estimate.[9]

For each of the six variables, significant differences between ESS and LFS estimates show up in at least around half of the countries in both rounds. The only exception is the variable work status, where in ESS 9 only 9 out of 25 countries exhibit a significant difference. The patterns of the differences are rather similar in ESS 8 and ESS 9. Broadly speaking, we can state that – in case significant differences occur – the following patterns prevail:

- Females tend to be overrepresented.

- Younger age groups (in particular, persons who are between 15 and 34 years old) tend to be underrepresented.
  Conversely, persons who are between 55 and 74 years old tend to be overrepresented.

- Married persons (including persons living in a registered partnership) tend to be over-represented.

- Persons in paid work tend to be overrepresented.

- Non-nationals tend to be underrepresented in nearly all countries.

- When it comes to household size, the most notable pattern is that persons living in two-person households tend to be overrepresented.
  In return, persons living either alone or in large (5 or more person) households tend to be underrepresented.

[9] We could not estimate the sampling errors of the LFS estimates. Due to the rather large sample size, they tend to be small (see the examples in Eurostat 2020, p. 14ff). In addition, the post-stratification weighting applied in the LFS should eliminate sampling error, at least for the characteristics used as control (see section 4 above).

Table 5: Differences between ESS 8 and LFS 2016 estimates (in percentage points) *

					Age					In paid	Non-			HH-size		
	Female	15-24 y.	25-34 y.	35-44 y.	45-54 y.	55-64 y.	65-74 y.	75+ y.	Married	work	national	1p-hh	2p-hh	3p-hh	4p-hh	5+p-hh
AT	3.3	-2.2	-2.5	1.1	1.9	3.4	1.5	-3.3	9.5	5.1	-7.1	-2.8	5.8	-0.3	-0.8	-2.0
BE	-1.5	1.8	-1.9	0.8	0.9	-0.1	0.9	-2.5	2.2	3.5	-2.4	0.3	-1.5	-1.8	1.5	1.3
CH	-2.5	1.2	-1.3	-1.7	-1.0	1.6	2.6	-1.5	5.1	1.2	-5.1					
CZ	0.0	5.8	1.7	-0.6	0.8	1.2	-2.6	-6.2	-2.7	6.6	-1.3	-3.3	1.5	4.3	-0.5	-2.0
DE	-3.9	2.2	-1.7	0.2	0.2	1.9	0.3	-3.2	2.3	3.1	-4.5	-5.9	-0.5	1.8	2.5	2.1
EE	0.8	-1.8	-1.6	-1.2	0.7	1.8	2.2		0.8	0.9	-3.6	-0.9	4.9	-1.0	-1.4	-1.6
ES	-1.0	0.0	-3.3	-1.9	2.0	2.5	2.0	-1.3	2.5	4.5	-3.7	-3.0	1.1	-0.1	2.2	-0.2
FI	-1.2	-2.5	-0.8	-0.7	0.9	1.4	1.4	0.3	1.2	0.0	-1.2					
FR	-0.9	-2.2	-3.2	-0.9	1.5	2.9	1.7	0.2	7.9	-0.3	-2.6	-1.1	3.7	-0.9	1.4	-2.9
HU	4.7	-2.7	-2.0	-2.3	-1.1	2.5	3.0	2.7	2.9	4.0	-0.4	7.6	7.8	-2.4	-6.2	-6.8
IE	1.2	-3.3	-3.9	-2.0	4.0	2.8	2.6	-0.2	7.5	-2.6	-4.4	0.7	4.3	-0.9	-3.0	-1.0
IS	0.8	-4.8	-1.3	-2.0	2.0	2.6	3.5		7.7	-3.0	-2.7					
IT	-0.8	1.8	0.1	-1.2	0.8	0.5	0.3	-2.3	-0.9	6.2	-1.8	1.5	3.4	-1.0	-3.2	-0.7
LT	1.9	0.4	-3.6	2.4	0.8	1.2	3.1	-4.5	6.1	2.1	-0.4	-13.2	6.4	7.1	2.8	-3.0
NL	4.3	0.1	-2.8	-0.1	-0.6	1.9	2.4	-1.0	4.7	-1.8	-2.1	-6.3	-0.3	1.7	3.8	1.1
NO	-3.2	-0.8	-1.4	-1.0	1.5	0.4	1.4		2.6	4.1	-1.8					
PL	0.1	0.1	-0.6	-1.5	0.4	1.3	1.8	-1.4	-4.1	0.0	-0.2	1.6	-0.4	1.5	-0.7	-2.0
PT	4.1	-2.0	0.1	-3.7	1.0	3.3	3.3	-2.0		1.2	-0.1	-0.2	4.0	-5.2	1.8	-0.3
SE	1.3	-4.7	-3.9	0.4	-0.8	3.7	5.3		7.3	1.0	-3.8					
SI	3.5	0.8	-1.2	-2.8	0.5	2.3	0.3	0.1	2.3	-1.1	-1.2	-7.5	1.0	-1.5	1.6	6.2
UK	3.4	-2.9	-1.7	0.1	0.6	1.5	3.0	-0.6	4.9	-0.9	-2.4	3.2	0.3	-1.3	-0.6	-1.7
# sign. diff.	7+ / 3-	4+ / 9-	1+ / 10-	1+ / 5-	2+ / 0-	9+ / 0-	14+ / 1-	1+ / 6-	13+ / 2-	8+ / 2-	0+ / 19-	3+ / 7-	8+ / 0-	3+ / 2-	4+ / 3-	2+ / 6-

* green = overrepresentation, red = underrepresentation, dark green / dark red = LFS estimate outside 95% confidence interval of ESS estimate

Table 6: Differences between ESS 9 and LFS 2018 estimates (in percentage points) *

					Age					In paid	Non-			HH-size		
	Female	15-24 y.	25-34 y.	35-44 y.	45-54 y.	55-64 y.	65-74 y.	75+ y.	Married	work	national	1p-hh	2p-hh	3p-hh	4p-hh	5+p-hh
AT	1.5	-3.7	-3.2	1.1	3.2	1.2	1.6	-0.3	6.9	1.3	-7.3	-1.4	7.1	-1.4	-0.8	-3.6
BE	-0.1	0.2	-0.6	-1.0	0.6	0.6	0.8	-0.5	0.2	4.1	-2.0	-1.5	1.0	-0.7	1.3	-0.1
BG	3.0	-1.9	-7.1	-2.8	-0.2	1.8	6.6	3.6	3.1	-7.3	0.1	3.0	4.2	-2.7	-3.2	-1.3
CH	-1.1	1.0	-0.9	-0.1	0.2	0.4	1.0	-1.6	1.7	1.6	-4.4					
CY	-1.3	-5.2	-8.3	-0.3	2.0	5.6	2.2	3.9	10.5	-0.5	-11.2	-2.5	3.1	-1.5	-0.5	1.4
CZ	4.5	3.8	-2.9	-3.2	6.3	0.5	-0.8	-3.7	0.4	3.6	-1.3	-6.8	-0.5	5.3	2.1	-0.1
DE	-2.2	1.1	-1.3	-1.5	-0.9	2.3	2.2	-2.0	2.7	1.5	-5.1	-5.3	1.3	1.0	1.4	1.6
EE	2.6	-1.3	-3.2	-0.7	0.6	1.2	3.2		-0.2	-0.4	-3.8	-6.0	3.3	-0.9	2.0	1.6
ES	-2.0	1.2	-1.0	-1.3	0.5	1.2	0.5	-1.1	-2.1	4.6	-2.2	-2.2	1.3	1.6	0.9	-1.6
FI	0.5	-2.2	-2.8	-0.1	0.4	1.9	2.2	0.4	1.3	-0.4	-1.0					
FR	0.8	-2.4	-2.5	-0.3	1.5	1.5	2.2	0.0	9.5	2.4	-1.5	0.0	0.7	-1.3	1.4	-0.6
HR	6.8	-0.9	-2.5	-0.8	2.9	3.0	2.5	-4.1	3.1	0.6	0.0	2.3	7.2	0.9	-5.8	-4.5
HU	4.8	-3.3	-1.7	-2.2	1.1	0.6	2.8	2.7	1.8	3.3	-0.5	8.4	5.7	-0.4	-6.6	-7.1
IE	1.4	-4.9	-4.8	-1.4	2.0	3.3	4.8	1.1	7.1	-5.7	-4.0	-1.1	1.3	-0.3	-1.1	1.2
IT	0.9	0.3	-1.2	-1.7	-1.2	0.7	2.7	0.5	-0.5	2.7	-1.9	1.8	4.3	-1.1	-3.4	-1.5
LT	13.2	-4.9	-4.5	-2.5	0.8	6.3	4.5	0.3	4.7	-1.1	-0.8	0.0	10.0	-1.5	-6.2	-2.3
LV	10.5	-4.5	-5.8	-3.7	4.6	2.5	4.6	2.3		-2.1	-5.3	1.1	7.1	-0.9	-4.5	-2.8
NL	-0.4	-0.9	-2.2	-0.1	1.1	1.8	1.2	-1.1	2.9	1.6	-1.3	-2.8	0.9	-0.6	1.4	1.2
NO	-4.4	-1.2	-3.0	-0.5	1.9	2.0	0.9		2.3	7.4	-5.8					
PL	0.4	1.6	-1.1	-0.8	-1.3	0.3	1.4	-0.1	-3.0	-0.7	-0.4	2.3	1.9	-1.6	0.2	-2.7
PT	3.8	-1.3	-0.8	-0.8	3.3	1.5	0.5	-2.4	2.5	0.0	2.9	-1.1	0.5	-4.5	1.1	3.9
SE	0.2	-5.0	-3.3	-1.2	-1.0	3.4	7.3		6.5	2.6	-4.2					
SI	3.1	1.4	-1.6	-0.5	-0.7	-0.1	1.4	0.0	-0.1	-1.3	-0.8	-3.6	3.9	-2.6	-0.9	3.1
SK	2.0	-4.8	-6.5	-3.9	0.3	5.2	9.0	0.7	7.3	0.4	-0.1	6.8	12.4	-4.9	-7.3	-6.9
UK	1.9	-5.0	-2.7	0.4	1.3	3.8	2.5	-0.3	6.9	1.0	-2.8	0.6	-0.7	-0.7	1.5	-0.6
# sign. diff.	10+ / 2-	1+ / 13-	0+ / 16-	0+ / 7-	7+ / 0-	12+ / 0-	17+ / 0-	3+ / 5-	12+ / 1-	7+ / 2-	1+ / 20-	6+ / 8-	10+ / 0-	1+ / 4-	2+ / 7-	4+ / 8-

* green = overrepresentation, red = underrepresentation, dark green / dark red = LFS estimate outside 95% confidence interval of ESS estimate

In 2020, each of the countries participating in ESS 10 received individualised feedback on the results of the present assessment of sample composition. The feedback included some suggestions on how countries might improve sample composition in the upcoming round by implementing a targeted survey design. Administrating targeted survey procedures to population subgroups can help to achieve response rates which are better balanced. The feedback document is displayed in the Appendix.

6. A summary measure of ESS-LFS differences

In order to arrive at a summary measure for the consistency of ESS and LFS variable distributions, we calculate the index of dissimilarity (Duncan & Duncan 1955) for each socio-demographic variable of our analysis:

$D = ½ \sum_i^n |ESSi - LFSi|$
with n = number of categories,
ESSi = percentage in category i of ESS,
LFSi = percentage in category i of LFS.

Table 7: Index of dissimilarity between ESS 8 and LFS 2016 variable distributions

Country	Gender	Age	Marital status	Work status	Nation-ality	House-hold size	mean
AT	3.3	8.0	9.5	5.1	7.1	5.9	6.5
BE	1.5	4.5	2.2	3.5	2.4	3.2	2.9
CH	2.5	5.5	5.1	1.2	5.1		3.9
CZ	0.0	9.5	2.7	6.6	1.3	5.8	4.3
DE	3.9	4.9	2.3	3.1	4.5	6.4	4.2
EE	0.8	4.7	0.8	0.9	3.6	4.9	2.6
ES	1.0	6.5	2.5	4.5	3.7	3.3	3.6
FI	1.2	4.0	1.2	0.0	1.2		1.5
FR	0.9	6.3	7.9	0.3	2.6	5.0	3.8
HU	4.7	8.2	2.9	4.0	0.4	15.4	5.9
IE	1.2	9.4	7.5	2.6	4.4	5.0	5.0
IS	0.8	8.1	7.7	3.0	2.7		4.5
IT	0.8	3.5	0.9	6.2	1.8	4.9	3.0
LT	1.9	8.0	6.1	2.1	0.4	16.3	5.8
NL	4.3	4.5	4.7	1.8	2.1	6.6	4.0
NO	3.2	3.3	2.6	4.1	1.8		3.0
PL	0.1	3.6	4.1	0.0	0.2	3.1	1.8
PT	4.1	7.7		1.2	0.1	5.8	3.8
SE	1.3	9.4	7.3	1.0	3.8		4.6
SI	3.5	4.0	2.3	1.1	1.2	8.9	3.5
UK	3.4	5.2	4.9	0.9	2.4	3.6	3.4
mean	2.1	6.1	4.3	2.5	2.5	6.5	3.9
min	0.0	3.3	0.8	0.0	0.1	3.1	1.5
max	4.7	9.5	9.5	6.6	7.1	16.3	6.5

The index of dissimilarity (D) is a measure widely used in research on segregation. The range of the index is between 0 and 100. In the present context, a value of 0 indicates that there is no

dissimilarity between the LFS and the ESS in the relative shares of respondents across the categories of a variable. A value of 100 indicates that the two distributions are completely dissimilar (consider, e.g., a dichotomous variable, where the first category comprises 100% in LFS and 0% in ESS, and the second category comprises 0% in LFS and 100% in ESS). The index of dissimilarity measures the percentage of respondents that would need to move between the categories of a variable to produce the same distribution for the two surveys. In contrast to the percentage point differences reported in the previous section, the index of dissimilarity is a non-directional measure. It does not provide an indication of which demographic subgroups are *over*- or *under*represented.

Table 8: Index of dissimilarity between ESS 9 and LFS 2018 variable distributions

Country	**Gender**	**Age**	**Marital status**	**Work status**	**Nation-ality**	**House-hold size**	**mean**
AT	1.5	7.2	6.9	1.3	7.3	7.2	5.2
BE	0.1	2.2	0.2	4.1	2.0	2.3	1.8
BG	3.0	12.0	3.1	7.3	0.1	7.2	5.5
CH	1.1	2.6	1.7	1.6	4.4		2.3
CY	1.3	13.8	10.5	0.5	11.2	4.5	7.0
CZ	4.5	10.6	0.4	3.6	1.3	7.4	4.6
DE	2.2	5.7	2.7	1.5	5.1	5.3	3.7
EE	2.6	5.1	0.2	0.4	3.8	6.9	3.2
ES	2.0	3.4	2.1	4.6	2.2	3.8	3.0
FI	0.5	5.0	1.3	0.4	1.0		1.6
FR	0.8	5.2	9.5	2.4	1.5	2.0	3.6
HR	6.8	8.4	3.1	0.6	0.0	10.4	4.9
HU	4.8	7.2	1.8	3.3	0.5	14.1	5.3
IE	1.4	11.2	7.1	5.7	4.0	2.5	5.3
IT	0.9	4.2	0.5	2.7	1.9	6.1	2.7
LT	13.2	11.9	4.7	1.1	0.8	10.0	7.0
LV	10.5	14.0		2.1	5.3	8.2	8.0
NL	0.4	4.2	2.9	1.6	1.3	3.5	2.3
NO	4.4	4.8	2.3	7.4	5.8		4.9
PL	0.4	3.3	3.0	0.7	0.4	4.4	2.0
PT	3.8	5.3	2.5	0.0	2.9	5.6	3.3
SE	0.2	10.6	6.5	2.6	4.2		4.8
SI	3.1	2.9	0.1	1.3	0.8	7.1	2.5
SK	2.0	15.2	7.3	0.4	0.1	19.2	7.4
UK	1.9	8.0	6.9	1.0	2.8	2.1	3.8
mean	2.9	7.4	3.6	2.3	2.8	6.7	4.2
min	0.1	2.2	0.1	0.0	0.0	2.0	1.6
max	13.2	15.2	10.5	7.4	11.2	19.2	8.0

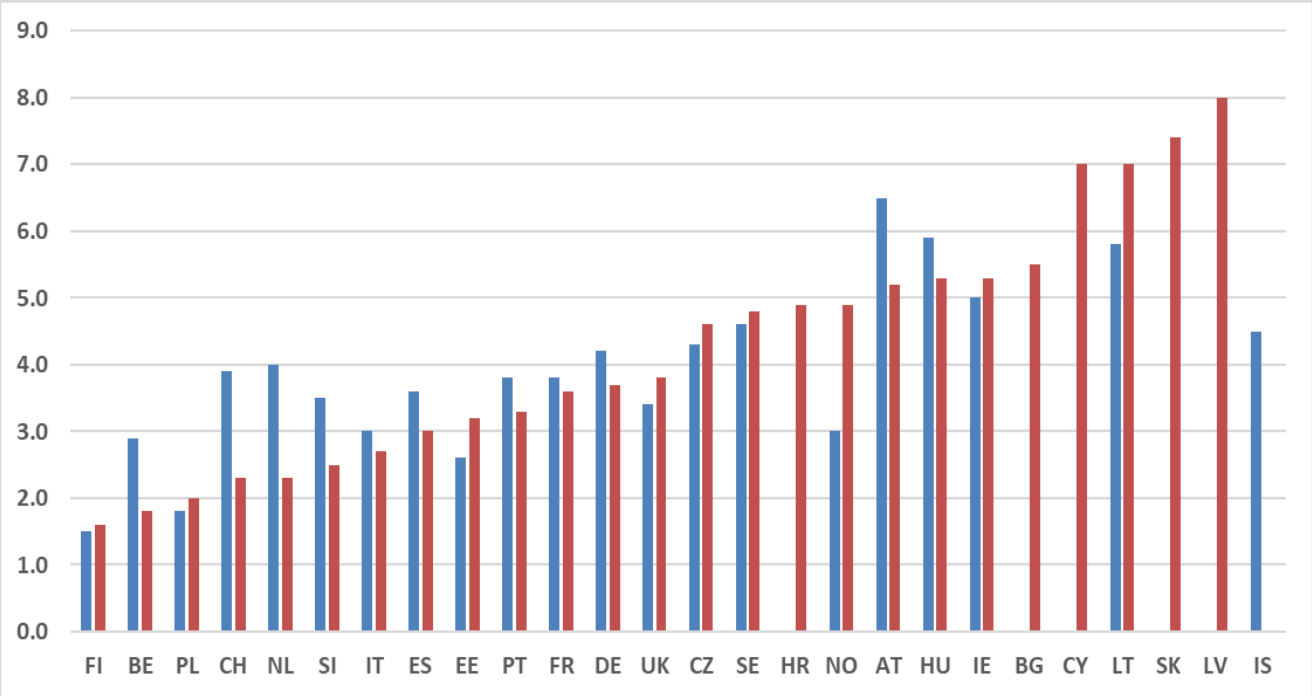
The size of D varies both between countries and between variables (see Table 7 and Table 8). Both in ESS 8 and in ESS 9, the largest dissimilarity pertains to the variable household size (D of 16.3 in Lithuania in round 8, and D of 19.2 in Slovakia in round 9). The mean value of D across all variables and countries is 3.9 in ESS 8 and 4.2 in ESS 9.[10] This means that – on

[10] The index of dissimilarity for household size is not available in five countries (Finland, Iceland, Norway, Sweden, and Switzerland). For these countries, the average value of D is based on the remaining five variables.

average – around 4% of respondents in ESS would have to change categories in order to achieve the same distribution as in the LFS. Accordingly, the average level of misrepresentation in the ESS does not seem to be very high. D is highest for the variables age (mean 6.1 and 7.4, respectively) and household size (mean 6.5 and 6.7, respectively). To some extent, this is the consequence of these two variables having a larger number of categories than the remaining variables.

In ESS 8, the mean value of D across the six variables varies between a low of 1.5 in Finland and a high of 6.5 in Austria (see Figure 1). In ESS 9, the range is between 1.6 (Finland) and 8.0 (Latvia). Countries with a rather high average D typically show values well above average in several variables (see Table 7 and 8).

Figure 1: Average level of dissimilarity (mean D across six variables);
ESS 8 (blue bars) and ESS 9 (red bars)



6.1. Average level of dissimilarity across ESS rounds

To put the results for ESS 8 and 9 into context, we can compare them with the respective information for ESS 5 thru 7. Table 9 (second row) shows that the average level of dissimilarity (average D across all countries) is highest in ESS 9. Whereas in ESS 9 the mean D is 4.2, in all previous rounds this number is slightly lower (3.8 and 3.9, respectively).[11]

When we look at the average level of dissimilarity variable by variable (see Table 9), we find that sample composition in ESS 9 is not uniformly worse than in previous rounds. As regards the variable work status, for instance, the average level of deviation decreases from round 5 to

[11] This does not mean, that individual countries have become worse over time, as the participating countries differ between rounds (see Table 10, below).

9, with round 9 showing the lowest average discrepancy. For the variables age and nationality, however, the largest average discrepancy can be observed in ESS 9.

Table 9: Average level of dissimilarity (mean D across countries), ESS 5 thru 9

	ESS 5	ESS 6	ESS 7	ESS 8	ESS 9
Average D (6 variables)	3.8	3.8	3.9	3.9	4.2
Gender	2.9	2.2	2.2	2.1	2.9
Age	6.2	5.9	6.3	6.1	7.4
Marital status	2.8	3.6	3.8	4.3	3.6
Work status	3.7	3.7	3.1	2.5	2.3
Nationality	2.3	2.5	2.2	2.5	2.8
Household size	5.6	5.3	7.1	6.5	6.7
# of countries	23	24	20	21	25

A glance at individual countries which participated in ESS rounds 5 thru 9 reveals different levels and trends in dissimilarity (Table 10). Several factors might have contributed to these results. Dedicated efforts to improve sample composition by implementing targeted fieldwork efforts might be one reason for the observed differences. In other instances, changes in the sample design or in the survey organization appointed might have come along – sometimes unintentionally – with differences in sample composition. We can neither describe nor explain these differences here in detail. Finding out about the concrete reasons will require a detailed look at the degree and patterns of misrepresentation at the level of individual variables. In addition, more country specific knowledge (about the sampling design, the interviewers deployed, the use of response enhancing measures like incentives, the number and timing of call attempts, etc.) will be required, and often further analyses, e.g. of the ESS contact forms data, will be advisable. Therefore, we will just pick up a few examples, and speculate in general terms about potential reasons for the results in question.

Belgium, Finland, and Poland, for instance, show little variation in D between rounds at a low level of discrepancy (Table 10). Each of these countries used a sample of individuals in every round. A sample of individuals usually has a positive effect on sample quality, compared to samples of households or addresses (Eckman & Koch, 2019). In addition, it seems notable that these countries showed a high consistency in the survey organisation appointed for fieldwork.

Ireland is a country which also displays little variation in the average degree of dissimilarity across rounds, however at a higher level. This higher level might relate to the fact that Ireland used a sample of addresses in every round.

In France, the deviations are considerably higher in ESS 7 than in the other rounds. A closer inspection shows that the underrepresentation of persons living either alone or in a two-person household is particularly high in round 7. Problems with the implementation of the sampling design (inaccessible addresses in large cities), which had been raised for round 7, might have played a role for this result.

Table 10: Average D (across six variables) for countries in ESS 5 thru 9*

Country	ESS 5	ESS 6	ESS 7	ESS 8	ESS 9	mean per country	range
PL	2.0	1.2	1.8	1.8	2.0	1.8	0.8
FI	1.8	1.7	2.9	1.5	1.6	1.9	1.4
BE	2.2	2.0	2.5	2.9	1.8	2.3	1.1
NO	1.3	3.3	2.4	3.0	4.9	3.0	3.6
ES	3.0	2.3	3.1	3.6	3.0	3.0	1.3
DK	3.0	3.5	2.9			3.1	0.6
CH	3.6	3.4	2.7	3.9	2.3	3.2	1.6
SE	2.5	2.3	2.1	4.6	4.8	3.3	2.7
IT		4.3		3.0	2.7	3.3	1.6
EE	3.9	3.1	4.6	2.6	3.2	3.5	2.0
SI	3.1	3.8	4.5	3.5	2.5	3.5	2.0
DE	3.4	3.7	4.0	4.2	3.7	3.8	0.8
NL	4.7	3.8	4.8	4.0	2.3	3.9	2.5
IS		3.4		4.5		4.0	1.1
GR	4.0					4.0	-
UK	3.1	5.4	4.4	3.4	3.8	4.0	2.3
CZ	3.1	5.2	4.9	4.3	4.6	4.4	2.1
FR	3.1	4.4	7.2	3.8	3.6	4.4	4.1
PT	7.6	4.4	3.9	3.8	3.3	4.6	4.3
HU	2.6	4.4	4.9	5.9	5.3	4.6	3.3
IE	4.3	4.1	4.9	5.0	5.3	4.7	1.2
BG	4.6	4.3			5.5	4.8	1.2
HR					4.9	4.9	-
LT	6.4	3.9	4.5	5.8	7.0	5.5	3.1
AT			5.2	6.5	5.2	5.6	1.3
SK	5.5	7.0			7.4	6.6	1.9
CY	8.7	5.3			7.0	7.0	3.4
LV					8.0	8.0	-
mean per round	3.8	3.8	3.9	3.9	4.2		

* Countries sorted in ascending order by mean D across ESS 5 thru 9 (seventh column)

In the Netherlands, the average level of dissimilarity in round 9 is lower than in the previous rounds. This might be related to switching from an address sample to a sample of individuals in ESS 9 (for the potential effects of different sampling designs in the Netherlands, see Kölln et al. 2019).

In Norway, the average dissimilarity in round 9 is higher than in previous rounds. This difference comes along with a change in the survey organisation fielding the ESS (moving from a state-run agency to a commercial agency).

Finally, as a last example, consider Portugal whose deviations were higher in ESS 5 than in subsequent rounds, with the smallest average difference observed in ESS 9. Again, one can speculate to what extent the various changes Portugal made with respect to the survey agency and the sampling design, have contributed to this development. In any case, demographic sample quality has improved over time in Portugal.

6.2. High levels of variable-specific dissimilarity across ESS rounds

Large values of D deserve special attention. If a country exhibits large differences for one or more variables in one round, or large differences across several rounds of ESS, a closer inspection of underlying patterns and potential causes seems indicated.

Table 12 on the next page exhibits a synopsis of results for ESS 5 thru ESS 9 on the level of individual variables. Results with a $D \geq 10.0$ are highlighted in red. Results with $5.0 \leq D \leq 9.9$ are highlighted in yellow.

Table 12 includes 650 data points in total, i.e. 650 measurements of D. Following Biemer et al's. (2018) suggestion, we may classify a $D \geq 10.0$ as a relevant deviation. Table 11 below summarises the results with respect to the occurrence of such large differences. Among the 650 measurements, 39 indices (= 6%) are 10.0 or larger.

Table 11: Number of indices of dissimilarity ≥ 10.0, by variable and ESS round

	ESS 5	ESS 6	ESS 7	ESS 8	ESS 9	Sum	Total # of D	% D ≥ 10.0
Gender	1	-	-	-	2	3	113	3%
Age	3	3	-	-	8	14	113	12%
Marital	-	1	1	-	1	3	110	3%
status								
Work	3	-	-	-	-	3	113	3%
status								
Nationality	1	1	-	-	1	3	113	3%
HH-size	1	1	5	2	4	13	88	15%
Sum	9	6	6	2	16	39	650	6%
Total # of D	132	138	115	120	145	650		
% D ≥ 10.0	7%	4%	5%	2%	11%	6%		

Whereas for gender, marital status, work status, and nationality the share of $D \geq 10.0$ is around 3% each, the rate of critical values is higher for the variable age (12%) and household size (15%). Table 11 and 12 also reveal differences between ESS rounds 5 thru 9. The highest share of critical values of D refers to ESS 9: 11% of all measurements show a dissimilarity index of 10.0 or larger. The respective share is lower for ESS 5 (7%), ESS 6 (4%), ESS 5 (5%), and ESS 8 (2%). Thus, the basic pattern with respect to the incidence of large discrepancies is similar to the one when we look at average values by variable and round.

As we have seen, large discrepancies are the exception. Notwithstanding this, the question to what extent post-stratification adjustments might help to mitigate any misrepresentation is especially relevant in such a case. This will be dealt with in the next section.

Table 12: Variable-specific indices of dissimilarity, ESS 5 thru ESS 9*

cntry	gndr_5	gndr_6	gndr_7	gndr_8	gndr_9	age_5	age_6	age_7	age_8	age_9	marital_5	marital_6	marital_7	marital_8	marital_9
AT			0.9	3.3	1.5			6.6	8.0	7.2			4.1	9.5	6.9
BE	0.5	0.1	2.0	1.5	0.1	3.7	2.6	4.8	4.5	2.2	0.3	0.3	0.5	2.2	0.2
BG	3.0	5.1			3.0	11.4	12.5			12.0	2.6	3.3			3.1
CH	2.5	1.1	0.9	2.5	1.1	4.8	4.1	4.4	5.5	2.6	3.6	3.5	3.8	5.1	1.7
CY	4.1	3.8			1.3	8.9	6.3			13.8	0.9	1.9			10.5
CZ	2.3	2.3	0.9	0.0	4.5	6.1	10.2	9.7	9.5	10.6	1.6	4.9	5.0	2.7	0.4
DE	2.8	1.4	1.9	3.9	2.2	5.8	5.6	6.0	4.9	5.7	1.3	1.8	3.8	2.3	2.7
DK	2.1	1.2	2.6			6.7	5.8	3.3			3.3	4.8	3.2		
EE	4.7	3.6	4.8	0.8	2.6	6.6	5.7	5.8	4.7	5.1	0.5	2.0	3.8	0.8	0.2
ES	0.5	0.5	2.5	1.0	2.0	3.4	3.3	3.6	6.5	3.4	1.9	2.1	1.6	2.5	2.1
FI	0.0	0.4	0.6	1.2	0.5	3.9	4.5	7.8	4.0	5.0		2.4	2.8	1.2	1.3
FR	0.1	1.9	0.2	0.9	0.8	5.6	7.4	9.2	6.3	5.2	6.7	9.3	13.0	7.9	9.5
GR	4.4					5.0					1.5				
HR					6.8					8.4					3.1
HU	0.9	1.9	4.5	4.7	4.8	3.7	4.2	6.8	8.2	7.2	0.2	3.3	4.1	2.9	1.8
IE	2.8	0.9	3.1	1.2	1.4	6.5	5.9	8.1	9.4	11.2	2.8	3.5	4.7	7.5	7.1
IS		0.9		0.8			4.6		8.1			4.3		7.7	
IT		0.3		0.8	0.9		5.6		3.5	4.2		1.3		0.9	0.5
LT	11.6	1.6	4.3	1.9	13.2	7.8	7.1	8.6	8.0	11.9	6.6	2.1	1.7	6.1	4.7
LV					10.5					14.0					
NL	2.3	2.4	4.2	4.3	0.4	7.1	6.5	5.5	4.5	4.2	9.3	7.8	7.4	4.7	2.9
NO	0.5	2.5	3.0	3.2	4.4	2.4	4.3	4.7	3.3	4.8	0.6	3.8	1.5	2.6	2.3
PL	0.6	0.0	2.3	0.1	0.4	3.6	2.0	3.0	3.6	3.3	3.0	0.4	2.6	4.1	3.0
PT	7.4	7.0	0.0	4.1	3.8	12.6	5.4	7.7	7.7	5.3	0.3	0.9	3.3		2.5
SE	2.8	0.3	0.1	1.3	0.2	5.6	4.0	4.3	9.4	10.6	0.8	3.8	1.8	7.3	6.5
SI	2.7	3.2	3.2	3.5	3.1	3.9	4.6	7.1	4.0	2.9	0.9	0.7	0.4	2.3	0.1
SK	5.6	5.0			2.0	13.1	12.1			15.2	9.2	10.1			7.3
UK	3.4	5.5	2.3	3.4	1.9	5.2	7.2	8.4	5.2	8.0	4.1	6.9	6.3	4.9	6.9

cntry	work_5	work_6	work_7	work_8	work_9	nation_5	nation_6	nation_7	nation_8	nation_9	hhsize_5	hhsize_6	hhsize_7	hhsize_8	hhsize_9
AT			0.6	5.1	1.3			5.0	7.1	7.3			14.2	5.9	7.2
BE	3.0	3.4	4.2	3.5	4.1	2.8	2.1	1.5	2.4	2.0	2.7	3.6	2.0	3.2	2.3
BG	5.9	1.5			7.3	0.1	0.1			0.1	4.5	3.1			7.2
CH	0.6	2.3	0.6	1.2	1.6	6.4	6.1	3.9	5.1	4.4					
CY	10.0	6.1			0.5	14.3	11.9			11.2	14.1	1.8			4.5
CZ	1.8	3.0	2.7	6.6	3.6	0.4	0.7	0.8	1.3	1.3	6.2	9.9	10.3	5.8	7.4
DE	1.1	2.8	2.7	3.1	1.5	3.2	4.5	3.7	4.5	5.1	6.2	6.1	5.8	6.4	5.3
DK	1.1	2.3	2.2			1.9	3.3	3.4							
EE	0.5	0.9	3.5	0.9	0.4	5.5	0.7	4.1	3.6	3.8	5.7	5.9	5.4	4.9	6.9
ES	2.7	1.8	2.2	4.5	4.6	5.5	2.9	4.1	3.7	2.2	4.2	3.1	4.4	3.3	3.8
FI	3.1	1.1	3.0	0.0	0.4	0.1	0.1	0.3	1.2	1.0					
FR	1.9	1.6	4.8	0.3	2.4	2.3	1.9	1.3	2.6	1.5	2.3	4.3	14.9	5.0	2.0
GR	7.2					0.2					6.0				
HR					0.6					0.0					10.4
HU	3.7	2.5	2.5	4.0	3.3	0.0	0.6	0.5	0.4	0.5	7.2	14.0	11.2	15.4	14.1
IE	11.1	7.0	4.8	2.6	5.7	0.4	5.0	5.3	4.4	4.0	2.3	2.3	3.4	5.0	2.5
IS		3.4		3.0			3.8		2.7						
IT		9.2		6.2	2.7		4.3		1.8	1.9		4.9		4.9	6.1
LT	2.4	2.4	0.9	2.1	1.1	0.4	0.2	0.3	0.4	0.8	9.6	9.8	11.1	16.3	10.0
LV					2.1					5.3					8.2
NL	0.3	0.6	3.7	1.8	1.6	1.6	1.9	1.1	2.1	1.3	7.5	3.7	7.0	6.6	3.5
NO	1.5	4.2	2.5	4.1	7.4	1.7	1.8	0.5	1.8	5.8					
PL	1.2	1.9	1.1	0.0	0.7	0.2	0.1	0.2	0.2	0.4	3.6	2.8	1.7	3.1	4.4
PT	15.2	7.7	6.6	1.2	0.0	0.5	0.3	0.6	0.1	2.9	9.8	5.4	5.1	5.8	5.6
SE	1.1	0.6	1.8	1.0	2.6	2.1	3.0	2.5	3.8	4.2					
SI	6.8	8.6	8.1	1.1	1.3	0.3	0.2	1.5	1.2	0.8	4.0	5.8	6.8	8.9	7.1
SK	0.6	7.8			0.4	0.2	0.4			0.1	4.1	6.6			19.2
UK	2.7	6.8	3.8	0.9	1.0	2.4	3.7	2.5	2.4	2.8	1.0	2.3	3.4	3.6	2.1

* Yellow: $5.0 \leq D \leq 9.9$; red: $D \geq 10.0$

7. Effect of post-stratification weights

Using post-stratification adjustments is, in principle, a cost-efficient approach to improve survey representativeness. Well-designed post-stratification weights can correct for sampling, coverage, and nonresponse errors. Applying post-stratification weights, however, will not reduce any bias that arises within weighting classes. Weighting is therefore unlikely to compensate completely for survey misrepresentation. In addition, if misrepresentation is large, some weights will also be large. In this case, the use of post-stratification weights will increase the variance of estimates and lead to a loss in precision.

The ESS has been providing post-stratification weights for its users for some time. These weights have been constructed using information on gender, age group, education, and region (for ESS 8, see Lynn & Anghelescu 2018). The post-stratification weights (pspwght) are obtained by adjusting the ESS design weights (dweight) in such a way that they will replicate the distribution of the cross-classification of gender, age group, and education in the population, and the marginal distribution for region in the population.[12] In most countries, the population distributions for the adjusting variables were obtained from the European Union Labour Force Survey.[13] For gender, a simple dichotomy (male vs. female) has been used. Age has been grouped into three categories (15–34 years, 35–54 years, and 55 years or older). Both ESS and LFS use the ISCED classification for measuring education. For weighting, the education measure has been recoded into a three-level variable. The recoding of the variable region generally follows the standard NUTS division of countries. Since regions are country-specific, they require separate specification of recoding procedures for each country.

Table 13 shows the average indices of dissimilarity across countries for the six variables, both without and with applying the ESS post-stratification weights. The level of reduction in dissimilarity by using post-stratification weights varies between variables. The largest relative reduction pertains to the variable gender (around 90%), followed by the variables marital status and age (between 38% and 50%). The smallest reduction pertains to the variables work status, nationality, and household size (between 6% and 13%). That the level of reduction is highest for the variable gender does not come by surprise. Gender is among the control variables included in the post-stratification weight. One usually would expect that the variables included as control in the post-stratification weight will show a more or less perfect fit with the benchmark data. The variable age also has been used as a control for the post-stratification weight. Here, however, the reduction in dissimilarity is much smaller than the one regarding the variable gender. Different categorisations of the age variable may have contributed to this result. For the calculation of the post-stratification weight, only three different age groups have been distinguished. In the present comparison, however, we use a more detailed categorisation with seven age groups.

On average across all countries and variables, the level of dissimilarity between ESS and LFS is reduced by about one-third when the post-stratification weights are used (see last row of Table 13). This applies both to round 8 (mean D of 3.9 vs. mean D of 2.7) and round 9 (mean D of 4.2 vs. mean D of 2.7).

[12] Accordingly, the ESS post-stratification weights are post-stratified design weights. For the sake of brevity, we use the term post-stratification weight in the present paper.

[13] When LFS data was incomplete or absent, these estimates have been taken from other sources: either data provided by the ESS National Coordinators or data obtained from the Office for National Statistics of that country. When data has been taken from LFS, annual estimates have been derived from the respective quarterly data sets. For ESS 8, information from LFS 2016 has been used. In some countries, information on education or region was not included in the weighting procedure (see Lynn & Anghelescu 2018).

Table 13: Effect of using post-stratification weights: average level of dissimilarity (mean D across all countries*); ESS 8 and ESS 9

Variable	**ESS 8 (dweight)**	**ESS 8 (pspwght)**	**Rel. red.****	**ESS 9 (dweight)**	**ESS 9 (pspwght)**	**Rel. red.****
Gender	2.1	0.2	90%	2.9	0.2	93%
Age	6.1	3.8	38%	7.4	4.2	43%
Marital status	4.3	2.5	42%	3.6	1.8	50%
Work status	2.5	2.2	12%	2.3	2.0	13%
Nationality	2.5	2.3	8%	2.8	2.5	11%
Household size	6.5	6.1	6%	6.7	5.8	13%
Mean D across 6 variables	3.9	2.7	31%	4.2	2.7	36%

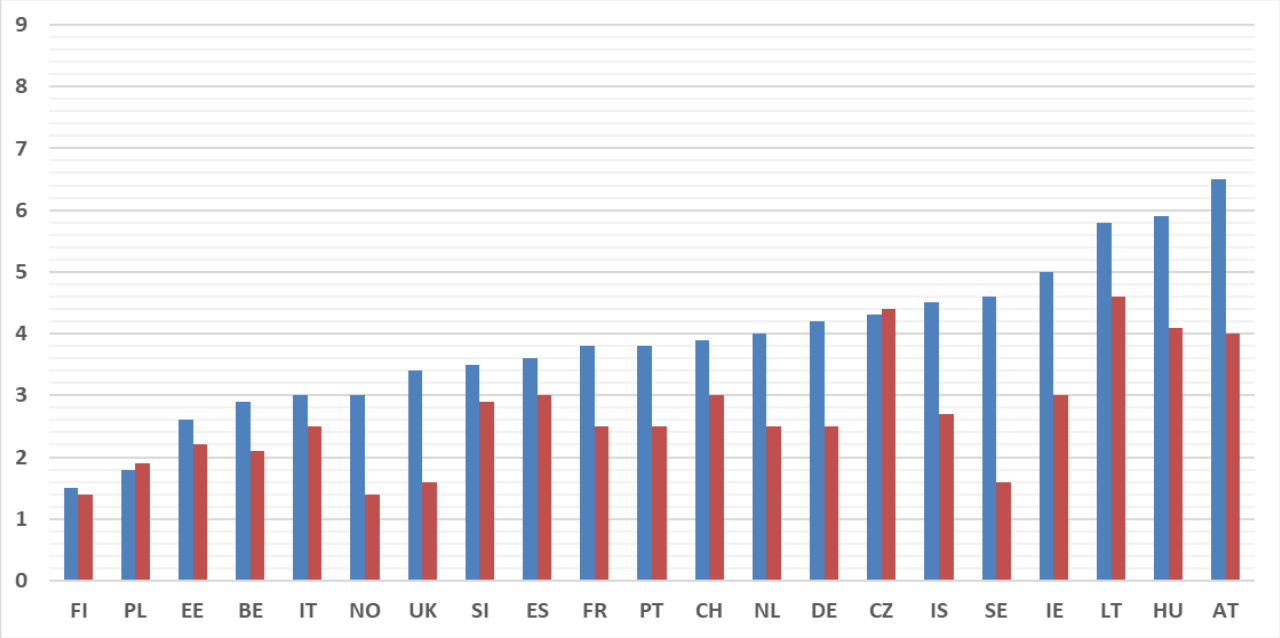
* Number of countries: ESS 8: 21; ESS 9: 25.
Household size: no information available in LFS for CH, FI, IS, NO, and SE. 16 and 21 countries remaining in ESS 8 and 9, respectively.
Marital status: no information available for PT in ESS 8 and LV in ESS 9.
** Relative reduction in average D, after applying post-stratification weight.

The level of improvement when applying post-stratification weights varies between countries (see Figure 2). Except from two countries, the introduction of post-stratification weights reduces the average size of differences between ESS and LFS in all countries of ESS 8 and ESS 9. The exceptions are Czechia and Poland in ESS 8, which show a slight increase in the average level of misrepresentation after applying post-stratification weights.

In ESS 8, the relative reduction in the mean index of dissimilarity across the six variables is largest in Sweden (-65%) and smallest in Finland (-7%). In ESS 9, the reduction is largest in Bulgaria (-73%) and smallest in Spain (-7%). In absolute terms, the most noticeable reduction refers to Sweden in ESS 8 (minus 3.0 percentage points) and to Bulgaria and Latvia in ESS 9 (minus 4.0 percentage points in both countries).

Figure 2: Average level of dissimilarity (mean D across six variables);
design weighted (blue bars) and post-stratification weighted (red bars) data;
ESS 8 and ESS 9

ESS 8



ESS 9

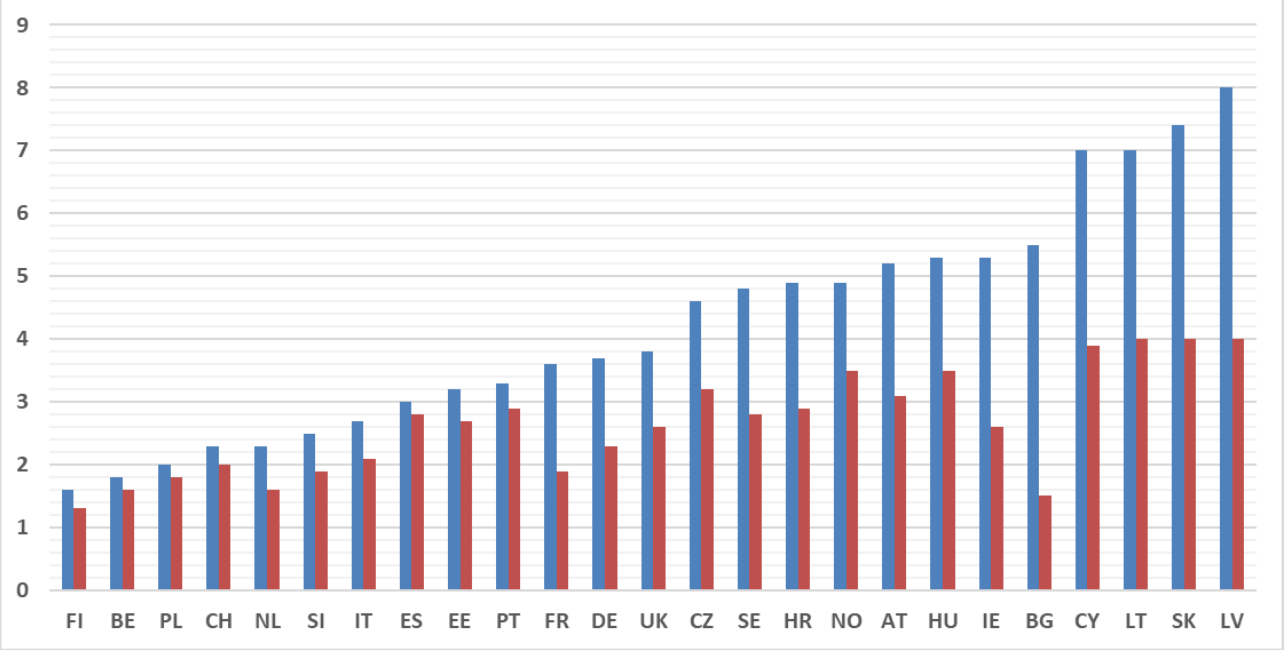


Figure 3 shows design weighted and post-stratification weighted results separately for the six variables of our analysis. Again, we see that the effect of the post-stratification weight varies both between countries and between variables. Usually, the post-stratification weight decreases the level of misrepresentation. Every now and then, however, countries exhibit an increase of ESS-LFS differences when the post-stratification weight has been used. In the following, we highlight the basic patterns for the six variables.

For the variable gender we observe a decrease in dissimilarity in practically all countries. After applying post-stratification weights, the difference between ESS and LFS estimates is close to zero in nearly every country.

For the variable age we observe a decrease in nearly all countries, albeit at a varying degree. In two countries in ESS 8, and four countries in ESS 9, however, the application of post-stratification weights is associated with a slight increase in dissimilarity. Each of these countries has a D below average without applying post-stratification weights.

For the variable marital status, we again find a decrease in dissimilarity in most countries. The magnitude of decrease, however, varies considerably between countries. In a few countries (both from round 8 and round 9), post-stratification weights slightly increase dissimilarity. This mainly occurs in countries exhibiting small deviations for the design weighted data.

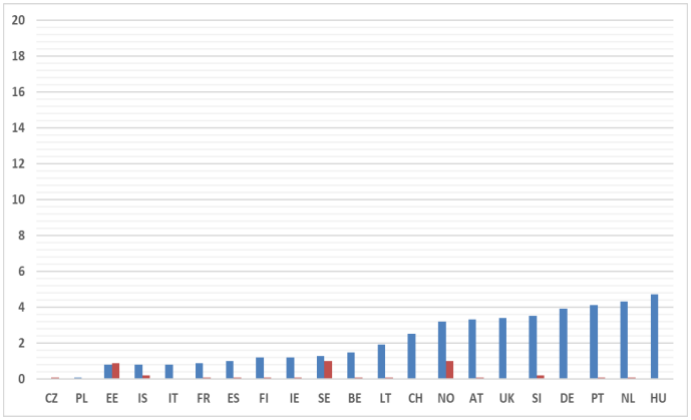
For the variable work status, we observe a decrease in dissimilarity in around half of the countries. In the other countries, however, post-stratification weights increase the difference to the LFS. This also happens in countries, where the difference to the LFS has been above average before applying post-stratification weights. Czechia and Hungary in ESS 8 are the main examples in that regard.

For the remaining two variables (nationality and household size), applying post-stratification weights has a rather small effect in nearly all countries. The main exception is Bulgaria in ESS 9, where post-stratification weighting reduces the misrepresentation for the variable household size considerably.

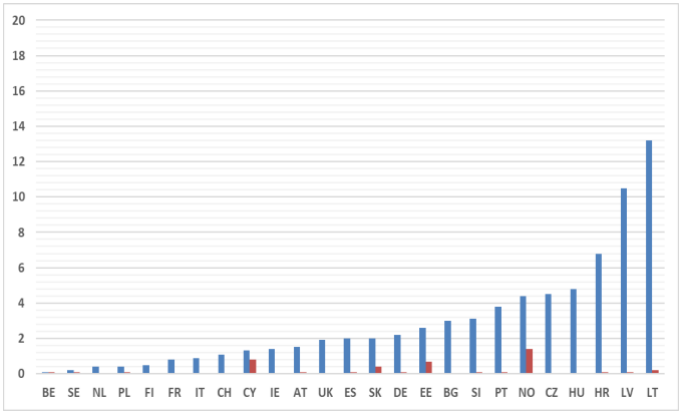
Finding out about the reasons for the different effects of post-stratification weighting, both between countries and between variables, is beyond the scope of the present paper. Such an endeavour requires country-specific insights into the relationship between the interesting variables, adjustment variables, and response propensities. To reduce nonresponse bias effectively, the adjustment variables need to be correlated with both the response propensity and the interesting variables (in the terminology of Groves 2006, the ‘common cause model’ has to apply).

Figure 3: Indices of dissimilarity ESS vs. LFS;
d-weighted (blue bars) and PS-weighted
(red bars) data; ESS 8 and ESS 9

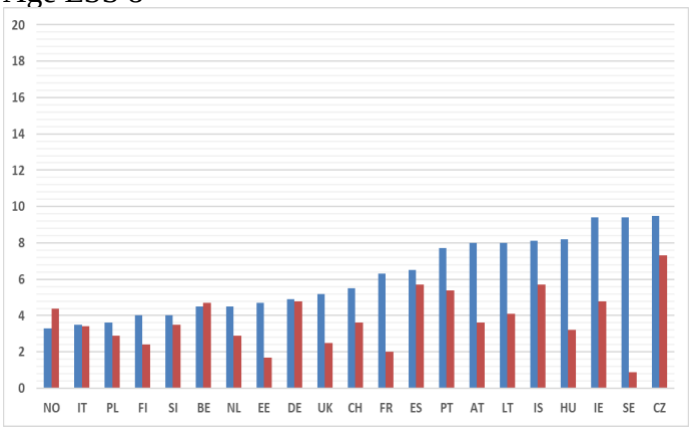
Gender ESS 8



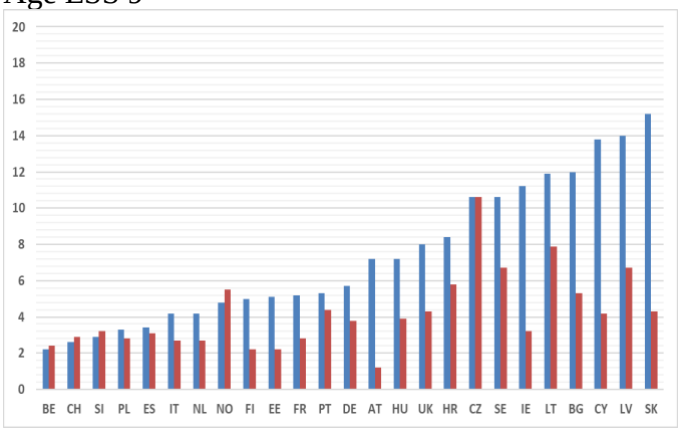
Gender ESS 9



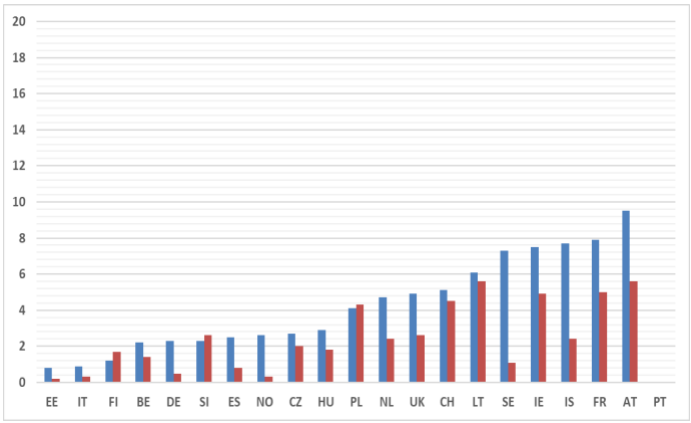
Age ESS 8



Age ESS 9



Marital status ESS 8



Marital status ESS 9

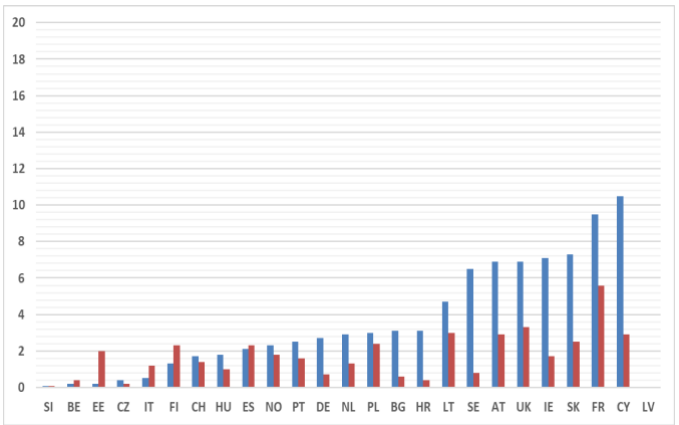
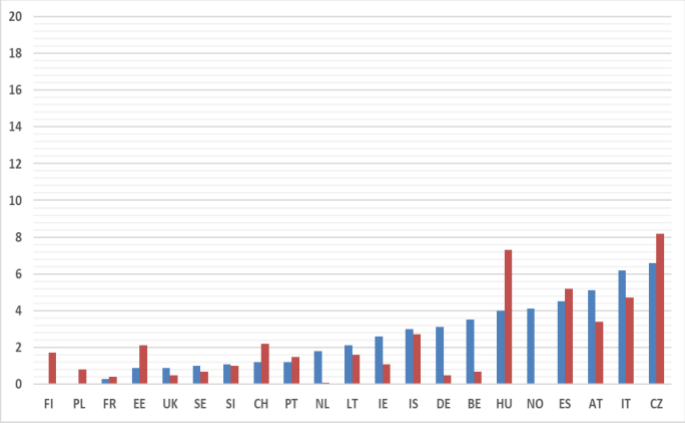
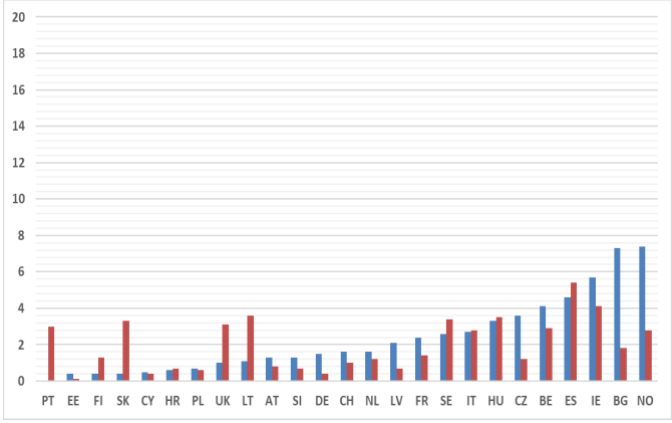


Figure 3: continued

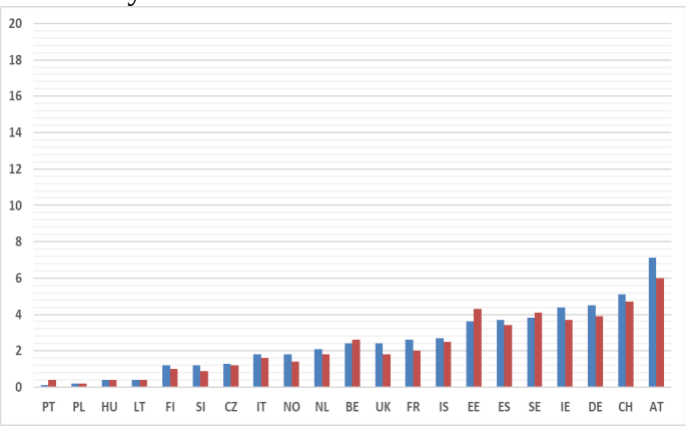
Work status ESS 8



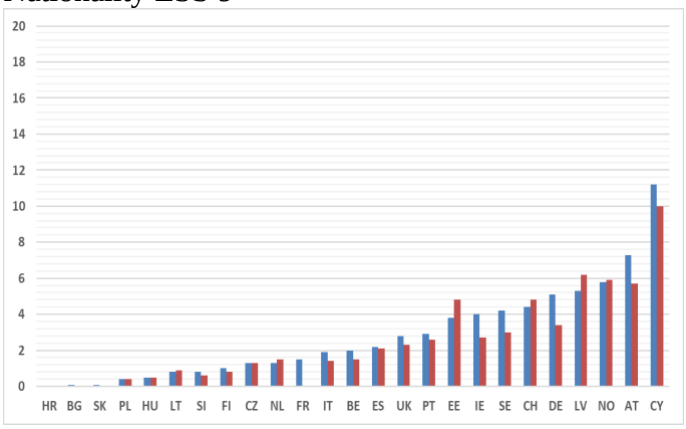
Work status ESS 9



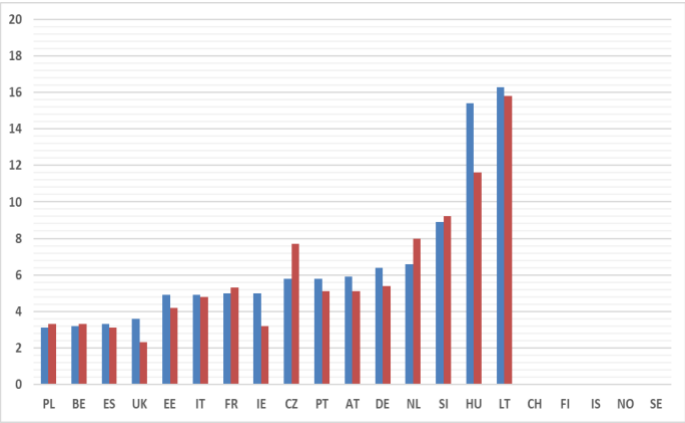
Nationality ESS 8



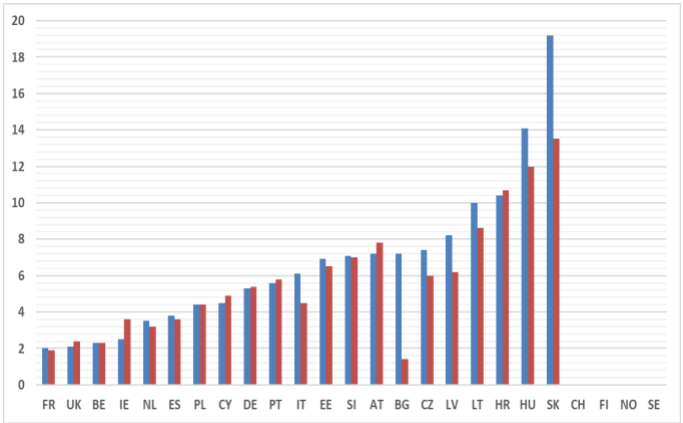
Nationality ESS 9



Household size ESS 8



Household size ESS 9



7.1. High levels of misrepresentation and effect of post-stratification weights

Positive effects of post-stratification weighting are especially beneficial in situations, where the level of misrepresentation is high. Table 14 displays indices of dissimilarity for the six variables in ESS 8 and ESS 9, before and after weighting with the ESS post-stratification weight. Higher values of D are highlighted in two different gradations. In the following, we again concentrate on indices of dissimilarity of 10.0 or larger.

In ESS 8, only 2 of the 120 measurements show a $D \geq 10.0$. Both cases concern the variable household size. In neither of the two countries, the post-stratification weight leads to a reduction in dissimilarity below 10.0.

In ESS 9, 16 out of 145 measurements exhibit a $D \geq 10.0$. In two countries, this refers to the variable gender. In both countries, the post-stratification adjustment reduces the dissimilarity close to zero.

Eight of the 25 countries in ESS 9 have a $D \geq 10.0$ at the variable age. In 7 countries, the post-stratification weight reduces the dissimilarity for age considerably (by around 1/3 to 2/3). However, in one country (Czechia) no reduction at all takes place. In Section 7.2, we provide a closer look at this country.

One country (Cyprus) shows a $D \geq 10.0$ both at the variable marital status and at the variable nationality. The post-stratification adjustment reduces the dissimilarity substantially for marital status. For nationality, however, the improvement is minimal.

In four countries of round 9, the household size distribution in ESS is very different ($D \geq 10.0$) from the results of the LFS. Only in one country (Slovakia), the post-stratification adjustment leads to a noticeable improvement (reducing the dissimilarity by around 1/3). However, even in that case the deviation after weighting stays high ($D \geq 10.0$).

Taken together, we can conclude that the ESS post-stratification weight reduces large discrepancies for the variables gender, age, and marital status in many cases considerably. For nationality and household size, however, only small, if any, improvements can be observed. This result mirrors the pattern we already observed for the variables in general (see above, p. 19ff).

Table 14: Effect of post-stratification weights on large discrepancies; ESS 8 and ESS 9*

ESS 8	gndr	gndr psp	age	age psp	marital	marital psp	work	work psp	national	national psp	hhsize	hhsize psp	mean	mean psp
AT	3.3	0.1	8.0	3.6	9.5	5.6	5.1	3.4	7.1	6.0	5.9	5.1	6.5	4.0
BE	1.5	0.1	4.5	4.7	2.2	1.4	3.5	0.7	2.4	2.6	3.2	3.3	2.9	2.1
CH	2.5	0.0	5.5	3.6	5.1	4.5	1.2	2.2	5.1	4.7			3.9	3.0
CZ	0.0	0.1	9.5	7.3	2.7	2.0	6.6	8.2	1.3	1.2	5.8	7.7	4.3	4.4
DE	3.9	0.0	4.9	4.8	2.3	0.5	3.1	0.5	4.5	3.9	6.4	5.4	4.2	2.5
EE	0.8	0.9	4.7	1.7	0.8	0.2	0.9	2.1	3.6	4.3	4.9	4.2	2.6	2.2
ES	1.0	0.1	6.5	5.7	2.5	0.8	4.5	5.2	3.7	3.4	3.3	3.1	3.6	3.0
FI	1.2	0.1	4.0	2.4	1.2	1.7	0.0	1.7	1.2	1.0			1.5	1.4
FR	0.9	0.1	6.3	2.0	7.9	5.0	0.3	0.4	2.6	2.0	5.0	5.3	3.8	2.5
HU	4.7	0.0	8.2	3.2	2.9	1.8	4.0	7.3	0.4	0.4	15.4	11.6	5.9	4.1
IE	1.2	0.1	9.4	4.8	7.5	4.9	2.6	1.1	4.4	3.7	5.0	3.2	5.0	3.0
IS	0.8	0.2	8.1	5.7	7.7	2.4	3.0	2.7	2.7	2.5			4.5	2.7
IT	0.8	0.0	3.5	3.4	0.9	0.3	6.2	4.7	1.8	1.6	4.9	4.8	3.0	2.5
LT	1.9	0.1	8.0	4.1	6.1	5.6	2.1	1.6	0.4	0.4	16.3	15.8	5.8	4.6
NL	4.3	0.1	4.5	2.9	4.7	2.4	1.8	0.1	2.1	1.8	6.6	8.0	4.0	2.5
NO	3.2	1.0	3.3	4.4	2.6	0.3	4.1	0.0	1.8	1.4			3.0	1.4
PL	0.1	0.0	3.6	2.9	4.1	4.3	0.0	0.8	0.2	0.2	3.1	3.3	1.8	1.9
PT	4.1	0.1	7.7	5.4			1.2	1.5	0.1	0.4	5.8	5.1	3.8	2.5
SE	1.3	1.0	9.4	0.9	7.3	1.1	1.0	0.7	3.8	4.1			4.6	1.6
SI	3.5	0.2	4.0	3.5	2.3	2.6	1.1	1.0	1.2	0.9	8.9	9.2	3.5	2.9
UK	3.4	0.0	5.2	2.5	4.9	2.6	0.9	0.5	2.4	1.8	3.6	2.3	3.4	1.6
# of countries with														
5.0 ≤ D ≤ 9.9	0	0	12	4	7	3	3	3	2	1	8	7	4	0
D ≥10.0	0	0	0	0	0	0	0	0	0	0	2	2	0	0

ESS 9	gndr	gndr psp	age	age psp	marital	marital psp	work	work psp	national	national psp	hhsize	hhsize psp	mean	mean psp
AT	1.5	0.1	7.2	1.2	6.9	2.9	1.3	0.8	7.3	5.7	7.2	7.8	5.2	3.1
BE	0.1	0.1	2.2	2.4	0.2	0.4	4.1	2.9	2.0	1.5	2.3	2.3	1.8	1.6
BG	3.0	0.0	12.0	5.3	3.1	0.6	7.3	1.8	0.1	0.0	7.2	1.4	5.5	1.5
CH	1.1	0.0	2.6	2.9	1.7	1.4	1.6	1.0	4.4	4.8			2.3	2.0
CY	1.3	0.8	13.8	4.2	10.5	2.9	0.5	0.4	11.2	10.0	4.5	4.9	7.0	3.9
CZ	4.5	0.0	10.6	10.6	0.4	0.2	3.6	1.2	1.3	1.3	7.4	6.0	4.6	3.2
DE	2.2	0.1	5.7	3.8	2.7	0.7	1.5	0.4	5.1	3.4	5.3	5.4	3.7	2.3
EE	2.6	0.7	5.1	2.2	0.2	2.0	0.4	0.1	3.8	4.8	6.9	6.5	3.2	2.7
ES	2.0	0.1	3.4	3.1	2.1	2.3	4.6	5.4	2.2	2.1	3.8	3.6	3.0	2.8
FI	0.5	0.0	5.0	2.2	1.3	2.3	0.4	1.3	1.0	0.8			1.6	1.3
FR	0.8	0.0	5.2	2.8	9.5	5.6	2.4	1.4	1.5	0.0	2.0	1.9	3.6	1.9
HR	6.8	0.1	8.4	5.8	3.1	0.4	0.6	0.7	0.0	0.0	10.4	10.7	4.9	2.9
HU	4.8	0.0	7.2	3.9	1.8	1.0	3.3	3.5	0.5	0.5	14.1	12.0	5.3	3.5
IE	1.4	0.0	11.2	3.2	7.1	1.7	5.7	4.1	4.0	2.7	2.5	3.6	5.3	2.6
IT	0.9	0.0	4.2	2.7	0.5	1.2	2.7	2.8	1.9	1.4	6.1	4.5	2.7	2.1
LT	13.2	0.2	11.9	7.9	4.7	3.0	1.1	3.6	0.8	0.9	10.0	8.6	7.0	4.0
LV	10.5	0.1	14.0	6.7			2.1	0.7	5.3	6.2	8.2	6.2	8.0	4.0
NL	0.4	0.0	4.2	2.7	2.9	1.3	1.6	1.2	1.3	1.5	3.5	3.2	2.3	1.6
NO	4.4	1.4	4.8	5.5	2.3	1.8	7.4	2.8	5.8	5.9			4.9	3.5
PL	0.4	0.1	3.3	2.8	3.0	2.4	0.7	0.6	0.4	0.4	4.4	4.4	2.0	1.8
PT	3.8	0.1	5.3	4.4	2.5	1.6	0.0	3.0	2.9	2.6	5.6	5.8	3.3	2.9
SE	0.2	0.1	10.6	6.7	6.5	0.8	2.6	3.4	4.2	3.0			4.8	2.8
SI	3.1	0.1	2.9	3.2	0.1	0.1	1.3	0.7	0.8	0.6	7.1	7.0	2.5	1.9
SK	2.0	0.4	15.2	4.3	7.3	2.5	0.4	3.3	0.1	0.0	19.2	13.5	7.4	4.0
UK	1.9	0.0	8.0	4.3	6.9	3.3	1.0	3.1	2.8	2.3	2.1	2.4	3.8	2.6
# of countries with														
5.0 ≤ D ≤ 9.9	1	0	9	6	6	1	3	1	4	3	9	8	8	0
D ≥10.0	2	0	8	1	1	0	0	0	1	1	4	3	0	0

* Large values of D highlighted: yellow: $5.0 \leq D \leq 9.9$; red: $D \geq 10.0$

7.2. Effect of post-stratification weights: the example of age in Czechia and Bulgaria in ESS 9

Czechia exhibits a rather high index of dissimilarity (D = 10.6) for the variable age (see upper panel in Table 15). Applying the post-stratification weight does not reduce the dissimilarity. D stays at the same level in Czechia, despite the age variable is one of the adjustment variables included in the ESS post-stratification weight. Bulgaria exhibits a dissimilarity which is similar in size to the one of Czechia (D = 12.0 for the variable age). For Bulgaria, however, applying the post-stratification weight reduces the discrepancy notably (from D = 12.0 to D = 5.3). A closer look at the data reveals the reason for this difference in effects.

In Czechia, the direction of difference between ESS and LFS varies within the three age weighting classes used for the post-stratification weight. The first weighting class, for instance, includes persons aged 15-34 years. This age group is only slightly stronger represented in ESS than in LFS (lower panel of Table 15: 27.4% vs. 26.5 %, difference = 0.9). However, when we look at the more detailed categorisation of age, we see that within this weighting class only persons aged 15-24 years are overrepresented (upper panel: 14.7% vs. 10.9%, difference = 3.8), whereas persons aged 25-34 years are underrepresented (upper panel: 12.7% vs. 15.6 %, difference = -2.9). Similar reverse patterns can be observed for the other two weighting classes. Consequently, the age variable shows a much smaller discrepancy to the LFS data for the broad categorisation applied in the post-stratification weight than for the fine graded categorisation we used (D = 4.0 vs. D = 10.6). As intended, the post-stratification weight brings the already small discrepancy of the broad categorisation close to zero (D = 0.2, lower panel). However, this does not apply to the finer categorisation, since here the differences in opposite directions remain present when the post-stratification weight is used (D sums up to 10.6 again, upper panel).

In Bulgaria, in contrast, the differences for the detailed age categories are in the same direction within the three weighting classes (upper panel). Consequently, the dissimilarity remains the same for the collapsed categorisation of age used for the post-stratification weight (D = 12.0, lower panel). When the post-stratification weight is applied, the discrepancy comes close to zero for the broad categorisation of age (D = 0.2, lower panel). In contrast to Czechia, however, the discrepancy is also considerably reduced for the finer age categorisation when the post-stratification weight is applied (D = 5.3, upper panel), as the differences were in the same direction within the weighting classes.

Table 15: Differential effects of post-stratification weights for the age variable

			CZ					BG		
	ESS 9 (dweight)	*difference*	LFS 2018	*difference*	ESS 9 (psp_weight)	ESS 9 (dweight)	*difference*	LFS 2018	*difference*	ESS 9 (psp_weight)
Categorisation of present analysis										
15-24y.	14.7	*3.8*	10.9	*3.4*	14.3	8.6	*-1.9*	10.5	*1.4*	11.9
25-34y.	12.7	*-2.9*	15.6	*-3.5*	12.1	8.1	*-7.1*	15.2	*-1.6*	13.6
35-44y.	16.2	*-3.2*	19.4	*-4.4*	15.0	14.5	*-2.8*	17.3	*-1.5*	15.8
45-54y.	22.6	*6.3*	16.3	*4.4*	20.7	16.1	*-0.2*	16.3	*1.5*	17.8
55-64y.	15.3	*0.5*	14.8	*1.9*	16.7	17.8	*1.8*	16.0	*-2.2*	13.8
65-74y.	13.2	*-0.8*	14.0	*0.9*	14.9	21.0	*6.6*	14.4	*1.9*	16.3
75+y.	5.2	*-3.7*	8.9	*-2.6*	6.3	13.9	*3.6*	10.3	*0.4*	10.7
		D = 10.6		D = 10.6			D = 12.0		D = 5.3	
Categorisation used for PS-weighting										
15-34y.	27.4	0.9	26.5	-0.1	26.4	16.7	-9.0	25.7	-0.2	25.5
35-54y.	38.8	3.1	35.7	0.0	35.7	30.6	-3.0	33.6	0.0	33.6
55+y.	33.7	-4.0	37.7	0.2	37.9	52.7	12.0	40.7	0.1	40.8
		D = 4.0		D = 0.2			D=12.0		D=0.2	

The ESS weighting guide says: “In a weighted analysis using post-stratification weights all the three mentioned errors (coverage, sampling and nonresponse) are completely corrected with respect to the post-stratification variables, and any other estimates are error free to the extent that such estimates are correlated with these variables.” (Kaminska 2020, p. 2f).

The Czech example shows that even for variables used for building the post-stratification weight this assertion does not hold in all circumstances. There are instances, where the post-stratification weight does not correct existing errors for the post-stratification variables. And for other estimates, the use of the post-stratification weight may even make matters worse, as we have seen in a few countries especially for the variable work status (see Figure 3 above).

Generally spoken, we must be aware that when the response behaviour varies within the weighting classes, the post-stratification weight will not correct existing errors completely. In order to eliminate nonresponse bias in a specific survey variable, this variable needs to be related to the variables used for post-stratification weighting and the respondents within a weighting class must have the same expected value on the target variable as the nonrespondents (in other words, nonresponse has to be “missing at random within classes”). Unfortunately, the latter assumption is usually untestable.

8. Summary and conclusions

(1) This paper used external benchmark data to assess the socio-demographic sample composition in ESS 8 and 9 with respect to gender, age, marital status, work status, nationality, and household size. We found indications of misrepresentation of demographic groups in the samples of the ESS. The level of misrepresentation varied between countries and variables. The basic pattern of results was rather similar in ESS 8 and 9 (e.g. underrepresentation of younger age groups and of non-nationals; overrepresentation of females and of married persons).

A detailed follow-up of the reasons for the differences in sample quality between individual countries was not part of the present task. The most obvious explanation for differences in sample quality between countries is that countries differ with respect to the response propensities of socio-demographic subgroups. If, for instance, a certain group is particularly difficult to contact or less willing to consent with a survey request, then an under-representation of this subgroup will occur.

In addition, we cannot preclude that the behavior of interviewers also affects sample composition. If interviewers, for instance, do not always adhere to the survey standards set in the ESS (e.g. by not following the ESS contact schedule or by substituting reluctant target persons by persons more willing to participate), this might also contribute to demographic misrepresentation.

(2) Applying post-stratification weights is, in principle, a cost-efficient approach to correct for demographic misrepresentation in sample surveys. The ESS has been providing standard post-stratification weights for some time. These weights have been constructed using information on gender, age group, education, and region (mainly from the LFS). Re-running the analyses with the ESS post-stratification weights revealed that the level of discrepancies between ESS and LFS usually decreases when the weights were applied. The size of the reduction, however, differs between countries and variables. In a few countries, differences to the LFS data even

increased for some variables when the post-stratification weights were applied. Thus, the standard ESS post-stratification weight is no panacea to deal with demographic misrepresentation.

Up till now, the ESS used a standard post-stratification weight, constructed in the same way for all countries. To improve the efficiency of the weight across countries, a more customised design of the weight might be considered. Such a design could use the same set of control variables for all countries, however, rely on different combinations and categorisations of these variables per country.

(3) The ESS strives to implement a policy of quality improvement from round to round. In light of the present findings, aiming for balanced response rates during fieldwork continues to be an important goal. ESS National Coordinators should be aware of the specific patterns of misrepresentation in their country. They should discuss potential reasons of misrepresentation and consider measures to improve in the upcoming round. To that end, the present results were fed back to the countries of ESS round 10 (see the feedback document in the Appendix). A few basic suggestions were provided on how to achieve (better) balanced response rates by administrating targeted survey procedures to population subgroups with generally low response rates.

References:

Biemer, P. P., Murphy, J., Zimmer, S., Berry, C., Deng, G. & Lewis, K. (2018). Using Bonus Monetary Incentives to Encourage Web Response in Mixed-mode Household Surveys. Journal of Survey Statistics and Methodology, 6, 240–261

Duncan, O. D. & Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. American Sociological Review, 20, 210-217

Eckman, S. & Koch, A. (2019). Interviewer involvement in sample selection shapes the relationship between response rates and data quality. Public Opinion Quarterly, 83(2), 313-337

European Social Survey (2014). Documentation of ESS Post-Stratification Weights. www.europeansocialsurvey.org, 25th April 2014

European Social Survey (2018). Round 9 Survey Specification for ESS ERIC Member, Observer and Guest Countries, v3. London: Centre for Comparative Social Surveys, City University London

Eurostat (2018a). Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2016. 2018 edition. eurostat Statistical working papers

Eurostat (2018b). Quality report of the European Union Labour Force Survey 2016. 2018 edition. eurostat Statistical Reports

Eurostat (2019a): Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2018. 2019 edition. eurostat Statistical reports

Eurostat (2019b): EU Labour Force Survey Database, User Guide, Version: November 2019

Eurostat (2020). Quality report of the European Union Labour Force Survey 2018. 2020 edition. eurostat Statistical Reports

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly, 70, 646-675

Kaminska, O. (2020). Guide to Using Weights and Sample Design Indicators with ESS Data. Institute for Social and Economic Research, University of Essex. V1.1

Koch, A. (2016). Assessment of socio-demographic sample composition in ESS Round 6. Mannheim: European Social Survey, GESIS

Koch, A. (2018): Assessment of socio-demographic sample composition in ESS Round 7. GESIS – Leibniz Institute for the Social Sciences, Mannheim/Germany

Koch, A., Halbherr, V., Stoop, I.A.L. & Kappelhof, J.W.S. (2014). Assessing ESS sample quality by using external and internal criteria. Mannheim: European Social Survey, GESIS

Köhne-Finster, S. & Lingnau, A. (2009). Untersuchung der Datenqualität erwerbsstatistischer Angaben im Mikrozensus. Ergebnisse des Projekts "Nachbefragung im Mikrozensus/LFS". Wirtschaft und Statistik, 12/2008, 1067-1088

Kölln, A.-K., Ongena, Y. P. & Aarts, K. (2019). The Effects of Sampling Frame Designs on Nonresponse and Coverage Error: Evidence from the Netherlands. Journal of Survey Statistics and Methodology, 7(3), 422–439

Lynn, P. (2019): ESS Round 8 Sample Design Data File: User Guide. Institute for Social and Economic Research, University of Essex

Lynn, P. & Anghelescu. G. (2018): European Social Survey Round 8 Weighting Strategy. Institute for Social and Economic Research, University of Essex

Ortmanns, V. & Schneider, S. L. (2016). Can we assess representativeness of cross-national surveys using the education variable? Survey Research Methods, 10(3), 189-210

Thomsen, I. & Villund, O. (2011). Using Register Data to Evaluate the Effects of Proxy Interviews in the Norwegian Labour Force Survey. Journal of Official Statistics, 27, 87-98

Wuyts, C. & Loosveldt, G. (2019): Quality matrix for the European Social Survey, Round 8. Overall fieldwork and data quality report. European Social Survey

Zühlke, S. (2008). Auswirkungen von Proxy-Interviews auf die Datenqualität des Mikrozensus. P. 3-10 in: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (ed.): Statistische Analysen und Studien NRW, Band 53

Appendix:

Individualised feedback to countries

Countries planning to participate in ESS round 10 were fed back the results of the sample composition assessment on an individual basis in the year 2020. The following countries received the attached document from their respective country contact in time to be taken into account when planning fieldwork for ESS 10:

AT – Austria

BE – Belgium

BG – Bulgaria

CH – Switzerland

CY – Cyprus

CZ – Czechia

DE – Germany

EE – Estonia

ES – Spain

FI – Finland

FR – France

HR – Croatia

HU – Hungary

IE – Ireland

IT – Italy

LT – Lithuania

LV – Latvia

NL - Netherlands

NO – Norway

PL – Poland

PT – Portugal

SE – Sweden

SI – Slovenia

SK – Slovakia

UK - United Kingdom

Assessment of demographic sample composition in ESS 9:

Individualised feedback to countries

Dummy country

Achim Koch & Jan-Lucas Schanze
GESIS Leibniz Institute for the Social Sciences, Mannheim, Germany
(July 2020)

Contents Page

1. Country-specific results

The ESS aims to produce high-quality data on social structure, attitudes, values and behaviour patterns in Europe. An important aspect of survey quality refers to the quality of the realised samples in terms of representation of the target population. In order to assess sample quality and the degree of nonresponse bias, data from ESS 9 were compared with external benchmark data from the European Union Labour Force Survey (LFS). The sample composition was assessed with respect to six core demographic variables:

- Gender (male / female)
- Age (10-year age cohorts)
- Marital status (married (incl. registered partnership): y/n)
- Work status (in paid work (for at least one hour): y/n)
- Nationality (national / non-national of a country)
- Household size (1 / 2 / 3 / 4 / 5+ persons)[14]

1.1 Procedure

Comparisons between ESS and LFS are possible for variables which were either measured in an identical way or, if this was not the case, where the measurements could be recoded to a common standard.[15] Independent of the timing of ESS Round 9 fieldwork in a country, LFS 2018 data (edition 2019) were used for the comparison. The ESS interviews persons aged 15 years and over resident within private households, regardless of their nationality, citizenship or language. In order to achieve comparable target populations, we excluded persons under 15 years in the LFS.[16] In addition, persons living in an institutional household (which were surveyed in a few LFS countries) were excluded.

ESS data were weighted with the design weight (DWEIGHT). This weight corrects for differences in selection probabilities between sampling units in a country. LFS data were weighted with the standard weight variable COEFF as recommended by Eurostat. COEFF corrects for differences in selection probabilities. In addition, it includes a post-stratification adjustment to adapt the LFS data to known population characteristics.[17]

[14] In Finland, Norway, Sweden and Switzerland no data on household size were available in the LFS.

[15] To learn more about procedures and variables, you may look at the respective reports for previous rounds of ESS (Koch 2016; Koch 2018; Koch et al. 2014).

[16] In Estonia, Norway and Sweden, persons aged 75 years and older were not (or only incompletely) included in the LFS. For these countries, both the data from the LFS and the data from the ESS were restricted to the population aged 15 to 74 years.

[17] In (nearly) all LFS countries data on gender, age and region were used for the adjustment. Several countries included additional variables in weighting, like information on unemployment or nationality (see Eurostat 2019). Using weighted data for the LFS thus should reduce both sampling errors and errors due to nonresponse or noncoverage – at least for the variables included in the weighting procedure.

Our estimates compare the ESS data prior to any adjustment for nonresponse or (non-) coverage with the results from the LFS. Thus, the differences provide the best measure of how the interviewed population differs from the true population (on condition that the LFS data represent the true population).

Assuming comparable target populations and comparable measurements, differences between ESS and LFS estimates can arise from sampling, coverage and/or nonresponse errors. To get an indication of whether the differences between ESS 9 and LFS 2018 estimates are within the limits of sampling error, complex standard errors can be estimated for the ESS estimates. In the final report on sample composition for ESS 9 (and ESS 8) we will provide this information.[18]

1.2 Results

Table 1 shows the results of the ESS-LFS comparison for your country in detail. For ESS 9, the ESS and LFS estimates and the difference between them are shown. For the purpose of comparison, the respective differences for previous rounds are also included (insofar as they were available).

What should you do with these results? Next steps:

- Check the plausibility of results:
  Do the ESS results follow typical patterns known also from other similar surveys in your country?
- Pay particular attention to very large differences and / or consistent differences across rounds
- Discuss / analyse potential reasons for the discrepancies (see section 3 below)
- Consider measures to reduce the over-/underrepresentation of certain groups in ESS 10 (see section 3 below)
- Discuss the results and potential measures with your country contact, with your SWEP expert and/or with the fieldwork team

[18] The report will be available by the end of spring 2021.

Table 1: Dummy country

Country abbr.	Variable	ESS 9	LFS 2018	Diff ESS-LFS ESS 9		Diff ESS-LFS ESS 8	Diff ESS-LFS ESS 7	Diff ESS-LFS ESS 6	Diff ESS-LFS ESS 5
ES	% female								
ES	% age 15-24								
ES	% age 25-34								
ES	% age 35-44								
ES	% age 45-54								
ES	% age 55-64								
ES	% age 65-74								
ES	% age 75+								
ES	% married								
ES	% in paid work								
ES	% non-national								
ES	% in 1pers hh								
ES	% in 2pers hh								
ES	% in 3pers hh								
ES	% in 4pers hh								
ES	% in 5+pers hh								

Note: Section 5 below provides more details on the data sources used.

2. Overall results

Table 2 on the next page provides a summary of the differences between ESS and LFS data for all countries in ESS 9 (edition 2.0).[19] For dichotomous variables (gender, marital status, work status, nationality), the differences for only one category are shown. For age and household size, differences for all categories are provided. Green cells indicate an overrepresentation of the respective category in a country in the ESS, while red cells indicate an underrepresentation. Thus, it can easily be checked whether the structure of demographic misrepresentation is similar across countries.

According to Table 2, the following patterns of under-/overrepresentation prevail across countries:

Underrepresented are
- Young age groups (15-44 years)
- Non-nationals
- Persons living in larger households

Overrepresented are:
- Females
- Middle-aged persons (45-74 years)
- Married persons
- Persons living in 2-person households

Mixed pattern:
- Oldest age group (75+ years)
- People in paid work
- Persons living in 1-person households

[19] Among the 27 countries included in ESS 9 (data edition 2.0) only Montenegro and Serbia are missing, as the LFS 2018 (edition 2019) does not include Montenegrin and Serbian data.

Table 2: Differences between ESS 9 and LFS 2018 estimates (in percentage points)*

Cntry	Female	Age							Married	In paid work	Non-national	HH-size				
		15-24y.	25-34y.	35-44y.	45-54y.	55-64y.	65-74y.	75+y.				1p-hh	2p-hh	3p-hh	4p-hh	5+p-hh
AT	1.5	-3.7	-3.2	1.1	3.2	1.2	1.6	-0.3	6.9	1.3	-7.3	-1.4	7.1	-1.4	-0.8	-3.6
BE	-0.1	0.2	-0.6	-1.0	0.6	0.6	0.8	-0.5	0.2	4.1	-2.0	-1.5	1.0	-0.7	1.3	-0.1
BG	3.0	-1.9	-7.1	-2.8	-0.2	1.8	6.6	3.6	3.1	-7.3	0.1	3.0	4.2	-2.7	-3.2	-1.3
CH	-1.1	1.0	-0.9	-0.1	0.2	0.4	1.0	-1.6	1.7	1.6	-4.4					
CY	-1.3	-5.2	-8.3	-0.3	2.0	5.6	2.2	3.9	10.5	-0.5	-11.2	-2.5	3.1	-1.5	-0.5	1.4
CZ	4.5	3.8	-2.9	-3.2	6.3	0.5	-0.8	-3.7	0.4	3.6	-1.3	-6.8	-0.5	5.3	2.1	-0.1
DE	-2.2	1.1	-1.3	-1.5	-0.9	2.3	2.2	-2.0	2.7	1.5	-5.1	-5.3	1.3	1.0	1.4	1.6
EE	2.6	-1.3	-3.2	-0.7	0.6	1.2	3.2		-0.2	-0.4	-3.8	-6.0	3.3	-0.9	2.0	1.6
ES	-2.0	1.2	-1.0	-1.3	0.5	1.2	0.5	-1.1	-2.1	4.6	-2.2	-2.2	1.3	1.6	0.9	-1.6
FI	0.5	-2.2	-2.8	-0.1	0.4	1.9	2.2	0.4	1.3	-0.4	-1.0					
FR	0.8	-2.4	-2.5	-0.3	1.5	1.5	2.2	0.0	9.5	2.4	-1.5	0.0	0.7	-1.3	1.4	-0.6
HR	6.8	-0.9	-2.5	-0.8	2.9	3.0	2.5	-4.1	3.1	0.6	0.0	2.3	7.2	0.9	-5.8	-4.5
HU	4.8	-3.3	-1.7	-2.2	1.1	0.6	2.8	2.7	1.8	3.3	-0.5	8.4	5.7	-0.4	-6.6	-7.1
IE	1.4	-4.9	-4.8	-1.4	2.0	3.3	4.8	1.1	7.1	-5.7	-4.0	-1.1	1.3	-0.3	-1.1	1.2
IT	0.9	0.3	-1.2	-1.7	-1.2	0.7	2.7	0.5	-0.5	2.7	-1.9	1.8	4.3	-1.1	-3.4	-1.5
LT	13.2	-4.9	-4.5	-2.5	0.8	6.3	4.5	0.3	4.7	-1.1	-0.8	0.0	10.0	-1.5	-6.2	-2.3
LV	10.5	-4.5	-5.8	-3.7	4.6	2.5	4.6	2.3		-2.1	-5.3	1.1	7.1	-0.9	-4.5	-2.8
NL	-0.4	-0.9	-2.2	-0.1	1.1	1.8	1.2	-1.1	2.9	1.6	-1.3	-2.8	0.9	-0.6	1.4	1.2
NO	-4.4	-1.2	-3.0	-0.5	1.9	2.0	0.9		2.3	7.4	-5.8					
PL	0.4	1.6	-1.1	-0.8	-1.3	0.3	1.4	-0.1	-3.0	-0.7	-0.4	2.3	1.9	-1.6	0.2	-2.7
PT	3.8	-1.3	-0.8	-0.8	3.3	1.5	0.5	-2.4	2.5	0.0	2.9	-1.1	0.5	-4.5	1.1	3.9
SE	0.2	-5.0	-3.3	-1.2	-1.0	3.4	7.3		6.5	2.6	-4.2					
SI	3.1	1.4	-1.6	-0.5	-0.7	-0.1	1.4	0.0	-0.1	-1.3	-0.8	-3.6	3.9	-2.6	-0.9	3.1
SK	2.0	-4.8	-6.5	-3.9	0.3	5.2	9.0	0.7	7.3	0.4	-0.1	6.8	12.4	-4.9	-7.3	-6.9
UK	1.9	-5.0	-2.7	0.4	1.3	3.8	2.5	-0.3	6.9	1.0	-2.8	0.6	-0.7	-0.7	1.5	-0.6

* green cells = overrepresentation; red cells = underrepresentation;

EE, NO, SE: persons 75 years or older not included; CH, FI, NO, SE: no LFS data on HH-size available; LV: marital status missing in ESS due to routing error

3. Further considerations on misrepresentation

The following remarks provide some further pieces of information on the misrepresentation of demographic groups in the ESS. They deal with the reasons for misrepresentation, emphasize the role of the sample design, and provide a few hints on how to improve in the upcoming round of ESS. Note that the deliberations do not intend to provide an exhaustive discussion of the topic.

3.1 Reasons for misrepresentation

Checking for the reasons for the observed differences between ESS and LFS estimates is a useful first step. Analytically, we may distinguish three different types of causes: (1) differential response propensities of demographic subgroups, (2) interviewer behaviour, and (3) other causes.[20]

(1) Differential response propensities

Demographic groups may differ in response propensities. Certain groups may be more difficult to interview, as they are

- more difficult to contact (young people; persons living alone),
- less able to participate due to language or health reasons (non-nationals and older persons),
- less willing to consent with a survey request.

In order to tackle the under-/overrepresentation of a specific group, it is helpful to have a grasp about which processes led to the pattern observed. Checking the ESS quality report for ESS 9 or running some dedicated analyses with the contact form data of your country can provide insights in that respect.

(2) Interviewer behaviour

Interviewers may contribute to the patterns of over-/underrepresentation observed. The number and timing of interviewers' call attempts, or the efforts interviewers exert to convince initially reluctant target persons will affect sample composition. If, for instance, certain groups of people are difficult to contact (like persons working full-time) it is useful to check whether the number and timing of contact attempts in a country is adequate. ESS standards lay down that contact attempts are not only made during daytime. To increase the chances of getting in touch with groups which are at home less often, also contact attempts in the evening and at weekends are required.

[20] We do not take sampling error into account here.

In other instances, interviewers may cut corners and deviate from prescribed sampling procedures. Some interviewers may preferentially interview households and persons who are cooperative and at home, in order to keep their response rate high and to reduce the number of visits required. If interviewers, for instance, tend to substitute a reluctant male target person by his cooperative wife when selecting a respondent within a household, this will lead to an overrepresentation of women in the final sample.

(3) Other reasons

The present analyses assume that LFS data are correct, and that ESS and LFS use comparable measurement instruments. If the LFS data themselves are in error, the present feedback will under- or overstate the level of discrepancies. This also holds when difficulties with respect to the comparability of ESS and LFS data exist, which were not taken into account here.[21] Finally, for the sake of completeness, it should also be mentioned that (differential) undercoverage of certain regions or groups in ESS and LFS may in principle contribute to the results.

3.2 The role of the sample design

The sample design used in a country is important in two respects. First, as analyses show, the average level of discrepancies between ESS and LFS is larger in countries, where a sample of households/addresses has been used than in countries where a sample of named individuals from a register has been used (Eckman & Koch 2019). The explanation for this result may lie in the interviewer involvement in sample selection. Interviewers play an important role in sample selection when a sample of households/addresses is used. In such a design, some interviewers may cut corners and substitute 'difficult' sample units by households/persons who are more accessible and/or more willing to participate.

Second, countries using a sample of individuals from a register usually have a richer sample frame than countries using a sample of households/addresses. Often, samples of individuals contain, for instance, information on gender and age of the sample unit. This information can be used to stratify the sample, to find out about the reason for any misrepresentation (is it mainly a problem of accessibility or of amenability?) and for applying targeted measures in the upcoming survey round (see below).

Thus, samples of individuals offer clear advantages compared to samples of households/ addresses in the given context.

[21] In a few countries, for instance, rather large differences in the proportion of married persons can be observed between ESS and LFS. In these countries it might be worthwhile to check in detail whether persons living in a registered partnership are treated in the same way in ESS and LFS.

3.3 Measures to improve

Broadly speaking, there are two approaches to deal with demographic misrepresentation: (1) Applying post-stratification weights and/or (2) aiming for balanced response rates.[22]

(1) Post-stratification weighting

The demographic variables we investigated can be used for the construction of post-hoc weighting variables. ESS has been providing standard post-stratification (PS-) weights for some time (Lynn & Anghelescu 2018). These weights have been constructed using information on gender, age group, education and region (mainly from the LFS). Analyses with the ESS PS-weights for the variables included in the present investigation, show that the level of discrepancies between ESS and LFS usually decreases when these weights were applied (Koch 2018). The size of the reduction, however, differs between countries and variables. For the variables not included as control in the weights, there is only a moderate decrease in the size of discrepancies. In some countries differences for some variables even increased when PS-weights were applied. Thus, the standard ESS PS-weight is no panacea to deal with that issue.[23]

(2) Aiming for balanced response rates

Against this backdrop, aiming for balanced response rates during fieldwork becomes important. (Better) balanced response rates can be achieved by administrating targeted survey procedures to population subgroups with generally low response rates (Haan/Ongena 2014; Lynn 2014, 2017). In order to implement such a targeted design, the relevant subgroups have to be identified and a decision on the treatment has to be made. Design features that can be targeted include:

- incentives (higher incentives to groups of sample members with low cooperation propensity),
- field time (prioritising cases: difficult cases to be worked at the beginning of field time),
- contact schedule (differences in number and timing of contact attempts of inter-viewers),
- content and design of communications like advance letters and brochures (e.g., specific advance letter for non-nationals; or using a standard advance letter, however mentioning prominently that the cooperation of non-nationals is important),
- differential interviewer payments (higher payment rates for low response propensity sample units, e.g. higher rates in urban areas),

[22] In ESS round 10, oversampling of low-response strata will be permitted. Oversampling will increase the sample size in the respective strata; however, it will not eliminate demographic misrepresentation.

[23] Generally spoken, PS-weights do not take account of any bias arising within weighting classes. In addition, one has to be aware that PS-weights do not provide improvements in precision.

- allocation of interviewers to sample cases (best interviewers to work on cases with lowest response propensity; possibly also re-allocation of interviewers during fieldwork in the context of a responsive survey design).

Targeting of measures is reasonable, when the underlying mechanism for misrepresentation is nonresponse. In case there are hints that interviewers do not adhere to ESS standards like the prescribed call pattern or that explicit misconduct of interviewers plays a role (like undocumented substitution), better briefing, monitoring and back-checking of interviewers, and – as the case may be – also better interviewer payment might be considered.

4. References

Eurostat (2019): Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2018. Statistical reports, eurostat

Eckman, S. & Koch, A. (2019): Interviewer involvement in sample selection shapes the relationship between response rates and data quality. Public Opinion Quarterly 83 (2): 313-337. doi: http://dx.doi.org/10.1093/poq/nfz012

Lynn, P. & Anghelescu, G. (2018): European Social Survey Round 8 weighting strategy. Institute for Social and Economic Research, University of Essex.

Similar exercises for previous rounds of ESS:

(Reports available at:
https://www.europeansocialsurvey.org/methodology/ess_methodology/data_quality.html)

Koch, A. (2018). Assessment of socio-demographic sample composition in ESS Round 7. Mannheim: European Social Survey, GESIS

Koch, A. (2016): Assessment of socio-demographic sample composition in ESS Round 6. Mannheim: European Social Survey, GESIS

Koch, A., Halbherr, V., Stoop, I.A.L. & Kappelhof, J.W.S. (2014): Assessing ESS sample quality by using external and internal criteria. Mannheim: European Social Survey, GESIS

A few further readings on targeted survey procedures:

Haan, M. & Ongena, Y. (2014): Tailored and targeted designs for hard-to-survey populations. Pp. 555-574 in R. Tourangeau, B. Edwards, T. P. Johnson, K. M. Wolter, & N. Bates (eds.), Hard-to-Survey Populations. Cambridge, UK: Cambridge University Press

Lynn, P. (2014): Targeted response inducement strategies on longitudinal surveys. Pp. 322-338 in U. Engel, B. Jann, P. Lynn, A. Scherpenzeel, & P. Sturgis (eds.), Improving Survey Methods: Lessons from Recent Research. New York: Routledge

Lynn, P. (2017): From standardised to targeted survey procedures for tackling non-response and attrition. Survey Research Methods, 11(1), 93-103

5. Data sources

ESS 9, data file edition 2.0, data weighted by DWEIGHT
LFS 2018, data file edition 2019 (released in July 2019), data weighted by COEFF

ESS 8, data file edition 2.1, data weighted by DWEIGHT
LFS 2016, data file edition 2018, data weighted by COEFF

ESS 7, data file edition 2.1, data weighted by DWEIGHT
LFS 2014, data file edition 2016, data weighted by COEFF
LFS 2015, data file edition 2016, data weighted by COEFF

ESS 6, data file edition 2.1, data weighted by DWEIGHT
LFS 2012, data file edition 2014, data weighted by COEFF
LFS 2013, data file edition 2014, data weighted by COEFF

ESS 5, data file edition 3.0, data weighted by DWEIGHT
LFS 2010, data file edition 2012, data weighted by COEFF

In the comparisons made for ESS 9, ESS 8 and ESS 5, the LFS data used for all countries refer to the 'official' survey year of the respective ESS round (2018, 2016 and 2010).
In the comparisons made for ESS 7 and ESS 6, in contrast, the LFS data were taken from two different survey years, depending on the year in which the seventh or sixth round of ESS was actually fielded in a specific country. For ESS 7, either LFS 2014 or LFS 2015 data were used. For ESS 6, either LFS 2012 or LFS 2013 data were used.

In case of questions and comments regarding the data and analyses, please get in touch with Achim Koch or Jan-Lucas Schanze at GESIS in Mannheim:

achim.koch@gesis.org Phone: +49-621-1246-401

Jan-Lucas.Schanze@gesis.org Phone: +49-621-1246-405