European
Social
Survey

Assessment of socio-demographic sample composition

in ESS Round 7

Achim Koch

GESIS Leibniz Institute for the Social Sciences, Mannheim

achim.koch@gesis.org

October 2018

Assessment of socio-demographic sample composition in ESS Round 7
Executive Summary

Approach:
- Comparison of six demographic variable distributions (gender, age, marital status, work status, nationality, and household size) from ESS 7 with external benchmark data from the EU LFS.
- Comparison of results for ESS 7 with those of analogous assessments for ESS 5 and 6.
- Analysis of the relation between the level of misrepresentation and the type of sample used and the response rate achieved.
- 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 7 are rather similar to the patterns observed for ESS 5 and 6 (e.g. underrepresentation of younger age-groups, of non-nationals, and of persons living alone; overrepresentation of females and of married persons).
- On average, countries using a sample of named individuals show lower levels of misrepresentation than countries using a sample of households or addresses.
- A negative correlation between the level of misrepresentation and the response rate (as one usually would expect) can only be observed for countries using a sample of individuals.
- 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 PS-weights increases the dissimilarity between ESS and LSF.

Interpretation and recommendation:
- Generally speaking, 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.
- This reasoning alone, however, cannot explain the findings as regards the relation between misrepresentation and response rates and sample type. An explanation might be found when we assume that interviewers do not always adhere to the survey standards set in the ESS, e.g. by substituting reluctant target persons by persons more willing to participate.
- 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.
- In light of these 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 with the CST and consider measures to improve in the upcoming round. To that end, the present results were fed back to the countries of ESS round 9 in spring 2018.

Assessment of socio-demographic sample composition in ESS Round 7[1]

Achim Koch

GESIS – Leibniz Institute for the Social Sciences, Mannheim/Germany
October 2018

Contents

Appendix:

Individualised feedback to countries

---

[1] The CST of the ESS requests that the following citation for this document should be used: Koch, A. (2018). Assessment of socio-demographic sample composition in ESS Round 7. 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 errors and bias due to nonresponse and over-/undercoverage 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 or not 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 take steps to 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 7, which was fielded in the years 2014 and 2015. 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 and ESS 6 (Koch et al. 2014; Koch 2016). The present analysis carries on and extends this exercise. For the first time, we analyse whether the use of post-stratification weights in the ESS is an effective way to reduce deviations from the LFS.

---

[2] The ESS 7 Specifications for participating countries state: "All countries are expected to aim for the 70% response rate or – at least – plan for a higher response rate than in the previous round." (European Social Survey 2013, p. 23)

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 often used to construct post-stratification weights. From 2014 onwards, post-stratification weights are also provided for the ESS (European Social Survey 2014).

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 33 European countries. These include the 28 Member States of the European Union, three EFTA countries (Iceland, Norway, and Switzerland), and two EU candidate countries (the Former Yugoslav Republic of Macedonia 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 various countries vary between 0.2% and 5.0%.

---

[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 2015a, 2015b, 2016a, 2017). 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).

The present comparison with the LFS is conducted for the seventh survey round of ESS. In total, 21 countries participated in ESS 7. Among these countries, 20 countries also participated in the LFS. Only Israel was 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 2014 for ESS 7). Unfortunately, not all countries managed to adhere to this schedule. In seven out of the 20 countries included in our analyses, all interviews or the majority of ESS 7 interviews were completed only in 2015 (see Table 1). These countries were Austria, the Czech Republic, Hungary, Lithuania, Poland, Portugal, and Spain. For these countries, we used LFS 2015 data for the comparison. For all the other countries, data from LFS 2014 were used.

Table 1: Timing of fieldwork in ESS 7

Country	ESS 7: % of interviews completed in year	
	2014	2015
AT	15.1	84.9
BE	95.6	4.4
CH	95.0	5.0
CZ	32.7	67.3
DE	80.5	19.5
DK	98.2	1.8
EE	100.0	0.0
ES	0.0	100.0
FI	91.8	8.2
FR	82.4	17.6
HU	0.0	100.0
IE	76.3	23.7
LT	0.0	100.0
NL	95.8	4.2
NO	100.0	0.0
PL	0.0	100.0
PT	0.0	100.0
SE	99.3	0.7
SI	97.1	2.9
UK	73.6	26.4

Source: ESS 7, ed. 2.1, variable 'inwyys' (start of interview, year)
Highlighted: Countries with all interviews or the majority of interviews completed in the year 2015

Table 2 documents a few basic parameters (participation compulsory, response rate[4] and rate of proxy interviews) for LFS 2014 and 2015. For the sake of comparison, ESS 7 response rates are also included.

Table 2: Basic characteristics of LFS 2014, LFS 2015 and ESS 7*

| Country | LFS 2014 | | | LFS 2015 | | | ESS 7 |
| | Partici-pation compul-sory | Response rate | Proxy rate among 15-74 years old respondents | Partici-pation compul-sory | Response rate | Proxy rate among 15-74 years old respondents | Response rate |
		(%)	(%)		(%)	(%)	(%)
AT	yes	94.3	25.2	yes	92.3	23.2	51.6
BE	yes	72.2	19.0	yes	73.3	17.8	57.0
CH	no	81.3	1.9	no	81.7	2.7	52.7
CZ	no	79.4	44.6	no	79.5	44.1	67.9
DE	yes	97.7	26.0	yes	96.6	25.6	31.4
DK	no	53.8	4.8	no	53.0	5.8	51.9
EE	no	68.7	36.4	no	71.9	35.7	59.9
ES	yes	84.8	52.5	yes	87.6	52.2	67.9
FI	no	72.0	4.1	no	70.8	4.2	62.7
FR	yes	79.1	27.8	yes	79.7	27.9	50.9
HU	no	82.8	44.0	no	82.8	42.7	52.7
IE	no	76.1	50.5	no	74.9	50.3	60.7
LT	no	80.5	34.5	no	79.7	34.5	68.9
NL	no	57.4	46.5	no	54.5	46.2	58.6
NO	yes	80.1	16.7	yes	79.7	16.5	53.9
PL	no	68.6	40.0	no	65.1	37.8	65.8
PT	yes	85.2	47.7	yes	84.6	47.9	43.0
SE	no	64.4	2.6	no	60.0	2.7	50.1
SI	no	78.7	55.6	no	78.7	53.4	52.3
UK	no	60.2	35.5	no	52.6	35.0	43.6
mean		75.9	30.8		75.0	30.3	55.2

* 20 countries which took part both in ESS 7 and in LFS
Source: Eurostat (2015a, 2015b, 2016a, 2017); ESS website: Fieldwork summary (10.07.2017)

Among the 20 countries, participation in the LFS was mandatory in 7 countries. The LFS response rates vary between 52.6% (UK 2015) and 97.7% (Germany 2014). Accordingly, the LFS, too, has a 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 (often considerably) higher than the ESS response rate. On average, the response rate in the LFS is almost 20 percentage points higher than in the ESS (average rate 75.9% in LFS 2014 and 75.0% in LFS 2015 vs. 55.2% in ESS 7). 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 2015a, 2016a). Several LFS countries

---

[4] 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).

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 are conducted. The percentage of proxy interviews varies between less than 2% (Switzerland in 2014) and 56% (Slovenia in 2014). On average across all 20 countries, around 30% of the interviews were proxy interviews (both in LFS 2014 and LFS 2015). 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 7 data (edition 2.1)[5] and anonymised EU LFS 2014 and 2015 data (edition 2016)[6]. 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.[7] 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 in our comparison, 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.

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 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 three 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

---

[5] European Social Survey Round 7 Data (2014). Data file edition 2.1. 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.

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

[7] The focus here is on comparability between the general standards set in the LFS and the ESS. However, one has to note that the comparability of measurements between countries within the LFS might also be an issue. The LFS sets various standards to ensure that the national surveys provide data that are compatible with the EU definitions. However, the leeway for differences in national questions is larger than in the ESS. Accordingly, the quality report for LFS 2014 states: "As a general conclusion it emerges that, in spite of the progress regarding the adherence to the EU regulations, principles and guidelines (i.e. the explanatory notes), the national questionnaires still largely differ even in the collection of key variables such as WSTATOR (Labour status in the reference week)." (Eurostat 2015b, p. 23)

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 2015a, 2016a). 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.

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)

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. With the exception of one variable (nationality), this could be achieved. Table 4 shows the minimum and maximum values among the 20 countries in the LFS 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 20 countries, the percentage of non-nationals is 2.0% maximum. These are (in ascending order) Poland, Hungary, Lithuania, Czech Republic, and Portugal. 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 (Denmark, Finland, Norway,

Sweden, and Switzerland).[8] Thus, the analyses with the variable household size had to be restricted to the remaining 15 countries.

Table 4: Minimum and maximum values of the variables analysed
           (20 countries, LFS 2014/15)

Variable / category	# of countries	Minimum (%)	Maximum (%)
Female	20	49.1	54.9
15-24 y.	20	11.5	17.2
25-34 y.	20	13.7	19.1
35-44 y.	20	14.4	19.9
45-54 y.	20	14.8	19.2
55-64 y.	20	13.4	17.7
65-74 y.	20	9.5	14.8
75+ y.	17	6.9	11.7
Married	20	41.8	59.8
In paid work	20	45.9	68.4
Non-national	20	0.2	23.2
1p-hh	15	9.9	23.0
2p-hh	15	26.0	38.6
3p-hh	15	15.3	28.8
4p-hh	15	14.5	24.5
5+p-hh	15	6.4	19.2

5. Patterns of misrepresentation

Which socio-demographic groups are over- or underrepresented in the ESS samples? Table 5 displays the direction and size of differences between ESS 7 and LFS 2014/15 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, taking the geographical clustering of the samples into account where relevant.[9] The confidence intervals were estimated using the complex sample procedure of SPSS. When the

---

[8] For Norway and Switzerland, no household information at all is available in the LFS. For Denmark, Finland, and Sweden, data is available only in separate country-specific files for a special household subsample. See Eurostat 2016b, p. 41.

[9] The majority of countries in ESS 7 (14 out of 20) used a geographically clustered sample. Only six countries used an unclustered design. These were Denmark, Estonia, Finland, the Netherlands, Sweden, and Switzerland.

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.[10]

For each of the six variables, significant differences between ESS and LFS estimates show up in at least half of the countries. Broadly speaking, we can state that − in case significant differences occur − the following patterns prevail:

- Females tend to be overrepresented (7 countries).

- Younger age groups (persons 15-24, 25-34 and 35-44 years old) tend to be underrepresented (6, 13 and 7 countries, respectively).
  Also the oldest age group (75 years or older) tends to be underrepresented (6 countries).
  Conversely, persons around and over 50 years (45-54, 55-64 and 65-74 years) tend to be overrepresented (4, 7 and 13 countries, respectively).

- Married persons (including persons living in a registered partnership) tend to be overrepresented (11 countries).

- As regards persons in paid work, significant deviations in both directions occur. In 6 countries, persons in paid work are underrepresented; in another 6 countries they are overrepresented.

- Non-nationals tend to be underrepresented (13 countries).

- When it comes to household size, the picture is diverse. The most notable pattern is that persons living in one-person households tend to be underrepresented (7 out of the 15 countries for which this information is available), whereas persons living in two-person households tend to be overrepresented (6 countries).

Patterns of misrepresentation across ESS rounds

An obvious question is whether the patterns we observed for ESS 7 are an idiosyncratic feature of this survey round. To answer this question, we can compare the patterns of misrepresentation in ESS 7 with the respective information from ESS 5 and ESS 6. With the exception of Austria, the relevant information from the two previous ESS surveys is available for all countries which participated in ESS 7. Table 6 provides the differences between ESS and LFS estimates for ESS rounds 5, 6 and 7 (the analyses conducted for ESS 5 and 6 are described in Koch et al. 2014 and Koch 2016). Again, the direction of the difference is indicated by colour. Green cells indicate an overrepresentation of this category in a specific ESS country compared to the LFS data, red cells indicate an underrepresentation. In Table 6, we do not differentiate whether a difference is significant or not.

---

[10] We could not estimate the sampling errors of the LFS estimates. Due to the rather large sample size (on average, approximately 140.000 persons 15 years and older per country), they tend to be small (see the examples in Eurostat 2015b, p. 13). 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 7 and LFS 2014/2015 estimates (in percentage points)*

	Female	15-24 y.	25-34 y.	35-44 y.	Age 45-54 y.	55-64 y.	65-74 y.	75+ y.	Married	In paid work	Non-national	HH-size 1p-hh	2p-hh	3p-hh	4p-hh	5+p-hh
AT	0.9	-4.7	0.0	1.9	-1.4	1.3	3.4	-0.4	-4.1	-0.6	-5.0	11.6	2.6	-5.1	-4.3	-4.8
BE	-2.0	2.0	-1.8	-0.6	0.9	1.2	0.7	-2.3	-0.5	4.2	-1.5	-0.4	-1.6	0.0	0.5	1.5
CH	-0.9	1.7	-2.6	-0.2	0.1	0.7	2.0	-1.5	3.8	-0.6	-3.9					
CZ	0.9	4.5	-3.0	0.9	3.0	1.3	-1.3	-5.4	5.0	2.7	-0.8	-5.2	-3.1	7.7	2.5	-2.0
DE	-1.9	0.1	-1.7	-1.5	1.8	2.5	1.6	-2.8	3.8	2.7	-3.7	-5.6	1.8	-0.2	1.1	2.8
DK	-2.6	-1.3	-0.6	-0.9	1.1	1.3	0.8	-0.5	3.2	2.2	-3.4					
EE	4.8	-3.6	-0.7	-1.2	-0.3	3.2	2.6		3.8	3.5	4.1	-1.0	3.7	1.7	-1.8	-2.6
ES	-2.5	-0.2	-0.3	-2.1	1.4	1.1	-1.0	1.1	1.6	2.2	-4.1	-3.6	2.5	-0.6	1.9	-0.1
FI	-0.6	-3.2	-2.6	-1.7	-0.3	3.0	3.7	1.0	2.8	-3.0	-0.3					
FR	-0.2	-1.7	0.3	6.2	2.7	-2.2	-1.9	-3.4	13.0	4.8	-1.3	-7.5	-7.4	0.4	7.5	7.0
HU	4.5	-2.2	-2.5	-2.2	1.1	1.1	3.6	0.9	4.1	2.5	-0.5	3.3	7.0	0.9	-6.1	-5.1
IE	3.1	-0.7	-3.9	-2.4	1.4	4.3	2.3	-1.1	4.7	-4.8	-5.3	1.2	0.9	-1.0	1.2	-2.4
LT	4.3	1.9	-3.2	0.6	3.3	0.2	2.5	-5.4	1.7	0.9	-0.3	-8.7	2.2	6.6	2.3	-2.3
NL	4.2	-1.1	-2.8	-0.8	1.5	2.6	1.4	-0.7	7.4	-3.7	-1.1	-6.6	-0.5	2.3	4.1	0.5
NO	-3.0	0.0	-3.8	-0.9	0.6	1.4	2.6		1.5	2.5	-0.5					
PL	2.3	0.6	-1.7	-0.1	-1.2	0.4	1.2	0.8	-2.6	-1.1	-0.2	0.0	-1.7	0.8	0.6	0.2
PT	0.0	-0.8	-3.1	-3.4	-0.4	0.7	5.5	1.5	3.3	-6.6	0.6	0.3	4.1	-5.0	-0.1	0.7
SE	0.1	-2.2	-1.0	-1.0	-0.1	1.2	3.1		1.8	1.8	-2.5					
SI	3.2	0.0	-2.4	-2.9	-1.3	3.6	3.5	-0.5	0.4	-8.1	-1.5	-4.3	2.0	-1.4	-1.0	4.8
UK	2.3	-4.9	-3.6	1.0	1.4	2.8	2.8	0.3	6.3	-3.8	-2.5	3.4	-0.5	-2.4	-0.3	-0.1
# sign. diff.	7+ / 2-	2+ / 6-	0+ / 13-	2+ / 7-	4+ / 0-	7+ / 1-	13+ / 1-	0+ / 6-	11+ / 2-	6+ / 6-	1+ / 13-	3+ / 7-	6+ / 2-	3+ / 3-	4+ / 2-	3+ / 5-

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

10

As can be seen, there is great stability in the *direction* of differences across the three survey rounds. For (nearly) all variables and categories distinguished, half or more of the countries show differences from the LFS in the same direction in all three rounds of ESS. This does not preclude that every now and then, the *size* of differences in a specific country for a specific variable varies considerably between the three rounds. In Portugal, for instance, the underrepresentation of persons in paid work in ESS 5, 6 and 7 is within the range of -6.6 to -15.2 percentage points. In France, the overrepresentation of married persons is within the range of +6.7 and +13.0 percentage points. Only in a minority of countries, the direction of differences changes between rounds. In these cases, the differences from the LFS are usually rather modest in both directions.

For a few categories, we even find that the overwhelming majority of differences in *all* countries in all three rounds of ESS are in the same direction. This pertains, for instance, to the category of '25-34 years old persons'. This age category is underrepresented in the three rounds of ESS in 56 of the 58 observations we have. A very uniform picture also shows up with respect to the categories of '55-64 years old persons' and '65-74 years old persons' (both of which are overrepresented), and with respect to the category of 'non-nationals' (under-represented).

The most obvious explanation for the patterns of under- and overrepresentation we observe will have to recur to differential response propensities of demographic subgroups. 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 old persons),
- less willing to consent with a survey request.

Interviewers may also contribute to the patterns of misrepresentation 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 might be useful to check whether the number and timing of contact attempts in a country is adequate. Investigating these processes in detail requires micro-level analyses of the response behaviour, either on the level of individual sample units or on the level of individual (interviewer) contact attempts for each country. This is not part of the present task. In section 7 of this paper, however, we will provide a basic analysis of two correlates of the size of dissimilarity between ESS and LFS data on the macro-level of countries.

In spring 2018, each of the 20 countries from ESS 7 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 9 of ESS 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.

Table 6: Differences between ESS and LFS estimates; ESS 5, 6 and 7 (in percentage points)

	Female			15-24y.			25-34y.			35-44y.			45-54y.			55-64y.			65-74y.			75+y.		
ESS	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7
AT			0.9			-4.7			0.0			1.9			-1.4			1.3			3.4			-0.4
BE	0.5	-0.1	-2.0	1.1	0.7	2.0	-0.9	-1.5	-1.8	0.5	-0.2	-0.6	-0.5	1.0	0.9	1.6	0.6	1.2	0.6	0.3	0.7	-2.2	-0.8	-2.3
CH	-2.5	-1.1	-0.9	0.4	1.1	1.7	-2.6	-2.3	-2.6	-1.7	-0.8	-0.2	1.0	0.2	0.1	2.2	1.5	0.7	1.3	1.3	2.0	-0.4	-1.0	-1.5
CZ	-2.3	-2.3	0.9	2.2	3.0	4.5	-1.5	-5.2	-3.0	0.0	1.8	0.9	3.9	5.0	3.0	0.0	0.4	1.3	-0.8	-0.4	-1.3	-3.8	-4.6	-5.4
DE	-2.8	-1.4	-1.9	2.4	1.7	0.1	-0.9	-0.9	-1.7	-1.4	-1.3	-1.5	2.3	1.7	1.8	1.1	1.0	2.5	-0.6	1.1	1.6	-2.9	-3.4	-2.8
DK	-2.1	-1.2	-2.6	-0.5	0.3	-1.3	-4.5	-3.5	-0.6	-0.4	-1.9	-0.9	2.1	0.8	1.1	1.6	2.9	1.3	3.0	1.8	0.8	-1.3	-0.3	-0.5
EE	4.7	3.6	4.8	-2.3	-1.0	-3.6	-3.3	-3.3	-0.7	-0.4	-1.1	-1.2	-0.6	-0.2	-0.3	1.8	1.5	3.2	2.8	2.1	2.6	1.9	2.2	
ES	0.5	0.5	-2.5	1.2	-0.8	-0.2	-1.5	-1.0	-0.3	-0.4	0.9	-2.1	2.0	1.0	1.4	-0.4	1.1	1.1	0.2	0.4	-1.0	-1.1	-1.4	1.1
FI	0.0	-0.4	-0.6	-1.0	-2.6	-3.2	-0.8	-1.3	-2.6	-1.3	-0.5	-1.7	-0.7	0.5	-0.3	1.1	1.1	3.0	2.4	2.6	3.7	0.4	0.3	1.0
FR	0.1	1.9	-0.2	-0.5	-3.0	-1.7	-3.3	-3.7	0.3	-0.3	1.3	6.2	3.3	3.4	2.7	2.1	1.6	-2.2	0.2	1.1	-1.9	-1.4	-0.6	-3.4
HU	0.9	1.9	4.5	-1.7	-0.8	-2.2	-1.5	-2.6	-2.5	0.7	1.3	-2.2	-0.5	1.2	1.1	2.5	1.2	1.1	0.2	0.4	3.6	0.3	-0.8	0.9
IE	2.8	0.9	3.1	4.4	1.0	-0.7	-4.8	-3.0	-3.9	-0.4	-1.1	-2.4	-0.1	0.7	1.4	0.7	3.0	4.3	1.4	1.3	2.3	-1.1	-1.7	-1.1
LT	11.6	1.6	4.3	-4.4	2.5	1.9	-3.0	-1.2	-3.2	0.7	1.7	0.6	-0.3	0.7	3.3	4.2	0.0	0.2	1.0	2.2	2.5	1.9	-5.9	-5.4
NL	2.3	2.4	4.2	-3.3	-3.2	-1.1	-3.2	-3.2	-2.8	0.5	0.7	-0.8	3.7	2.0	1.5	1.7	1.9	2.6	1.1	1.9	1.4	-0.6	-0.1	-0.7
NO	0.5	-2.5	-3.0	0.3	-0.5	0.0	-1.2	-2.7	-3.8	-0.6	-0.6	-0.9	-0.5	1.7	0.6	0.3	2.6	1.4	1.8	-0.4	2.6			
PL	-0.6	0.0	2.3	3.0	0.5	0.6	-1.4	-0.6	-1.7	-1.4	-1.2	-0.1	0.1	-0.2	-1.2	-0.4	0.5	0.4	0.4	1.0	1.2	-0.5	0.0	0.8
PT	7.4	7.0	0.0	-2.2	-0.1	-0.8	-6.4	-2.6	-3.1	-4.1	-2.7	-3.4	0.2	1.0	-0.4	3.6	2.5	0.7	5.6	1.8	5.5	3.1	0.0	1.5
SE	2.8	-0.3	0.1	-1.5	-1.4	-2.2	-2.1	-0.9	-1.0	-1.4	-1.7	-1.0	-0.6	0.4	-0.1	1.4	0.6	1.2	4.2	3.0	3.1			
SI	2.7	3.2	3.2	0.0	0.7	0.0	-1.2	-3.4	-2.4	-1.1	0.4	-2.9	-1.3	-1.1	-1.3	2.9	2.0	3.6	0.9	0.8	3.5	-0.3	0.7	-0.5
UK	3.4	5.5	2.3	-1.1	-2.1	-4.9	-2.4	-3.8	-3.6	-0.6	-1.2	1.0	1.1	0.7	1.4	2.6	2.6	2.8	1.5	2.7	2.8	-1.1	1.2	0.3

green = overrepresentation
red = underrepresentation

Table 6, continued: Differences between ESS and LFS estimates; ESS 5, 6 and 7 (in percentage points)

ESS	Mar-ried			In paid work			Non-natio-nal			1p-hh			2p-hh			3p-hh			4p-hh			5p-hh		
	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7	5	6	7	5	6	6	5	6	7
AT			-4.1			-0.6			-5.0			11.6			2.6			-5.1		-4.3				-4.8
BE	-0.3	0.3	-0.5	3.0	3.4	4.2	-2.8	-2.1	-1.5	-1.5	2.2	-0.4	0.6	-1.3	-1.6	-1.3	-2.2	0.0	0.8	1.0	0.5	1.2	0.4	1.5
CH	3.6	3.5	3.8	0.6	2.3	-0.6	-6.4	-6.1	-3.9															
CZ	-1.6	4.9	5.0	1.8	3.0	2.7	-0.4	-0.7	-0.8	-1.8	-5.3	-5.2	-2.2	-4.5	-3.1	4.4	0.8	7.7	1.8	8.2	2.5	-2.1	0.9	-2.0
DE	1.3	1.8	3.8	1.1	2.8	2.7	-3.2	-4.5	-3.7	-5.0	-5.3	-5.6	-1.1	-0.8	1.8	0.0	1.0	-0.2	3.6	2.8	1.1	2.6	2.2	2.8
DK	3.3	4.8	3.2	1.1	2.3	2.2	-1.9	-3.3	-3.4															
EE	0.5	2.0	3.8	0.5	-0.9	3.5	-5.5	-0.7	4.1	5.7	5.9	-1.0	-0.7	-0.2	3.7	-3.7	-1.1	1.7	-0.6	-2.3	-1.8	-0.7	-2.3	-2.6
ES	-1.9	2.1	1.6	2.7	1.8	2.2	-5.5	-2.9	-4.1	2.1	-2.9	-3.6	2.1	1.4	2.5	-1.2	1.7	-0.6	-1.2	0.0	1.9	-1.8	-0.1	-0.1
FI		2.4	2.8	-3.1	-1.1	-3.0	0.1	-0.1	-0.3															
FR	6.7	9.3	13.0	1.9	1.6	4.8	-2.3	-1.9	-1.3	-2.2	-3.1	-7.5	0.7	0.9	-7.4	0.7	-1.1	0.4	0.5	1.5	7.5	0.4	1.9	7.0
HU	0.2	-3.3	4.1	3.7	2.5	2.5	0.0	-0.6	-0.5	3.9	11.7	3.3	3.1	2.3	7.0	0.2	-1.5	0.9	-5.7	-7.1	-6.1	-1.5	-5.4	-5.1
IE	-2.8	3.5	4.7	-11.1	-7.0	-4.8	-0.4	-5.0	-5.3	1.6	-1.3	1.2	0.7	-1.0	0.9	0.0	0.0	-1.0	-1.5	2.3	1.2	-0.8	0.0	-2.4
LT	-6.6	2.1	1.7	-2.4	2.4	0.9	0.4	-0.2	-0.3	1.7	-6.6	-8.7	6.8	2.0	2.2	1.1	5.8	6.6	-5.5	2.0	2.3	-4.1	-3.1	-2.3
NL	9.3	7.8	7.4	-0.3	-0.6	-3.7	-1.6	-1.9	-1.1	-5.8	-3.6	-6.6	1.1	0.2	-0.5	-1.7	2.6	2.3	5.1	0.3	4.1	1.3	0.6	0.5
NO	0.6	3.8	1.5	1.5	4.2	2.5	1.7	1.8	-0.5															
PL	-3.0	-0.4	-2.6	1.2	1.9	-1.1	0.2	-0.1	-0.2	1.6	2.0	0.0	0.4	0.2	-1.7	-1.5	0.6	0.8	1.5	-0.4	0.6	-2.1	-2.4	0.2
PT	0.3	0.9	3.3	-15.2	-7.7	-6.6	-0.5	-0.3	0.6	0.8	2.1	0.3	9.0	3.3	4.1	-3.9	-0.4	-5.0	-5.0	-3.9	-0.1	-0.8	-1.0	0.7
SE	-0.8	3.8	1.8	1.1	-0.6	1.8	-2.1	-3.0	-2.5															
SI	0.9	0.7	0.4	-6.8	-8.6	-8.1	-0.3	-0.2	-1.5	-2.6	-3.8	-4.3	-0.8	1.6	2.0	-0.4	-1.9	-1.4	-0.1	0.6	-1.0	4.0	3.6	4.8
UK	4.1	6.9	6.3	-2.7	-6.8	-3.8	-2.4	-3.7	-2.5	-1.0	-2.1	3.4	0.0	0.5	-0.5	0.4	-0.1	-2.4	0.3	-0.1	-0.3	0.3	1.8	-0.1

green = overrepresentation
red = underrepresentation

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 in our analysis:

$$D = \frac{1}{2} \sum_i^n |ESSi - LFSi|$$

with n = number of categories,
ESSi = percentage in category i of ESS,
LFSi = percentage in category i of LFS.

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 exactly 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 7: Index of dissimilarity between ESS 7 and LFS 2014/2015 variable distributions

Country	Gender	Age	Marital status	Work status	Nation-ality	House-hold size	mean
AT	0.9	6.6	4.1	0.6	5.0	14.2	5.2
BE	2.0	4.8	0.5	4.2	1.5	2.0	2.5
CH	0.9	4.4	3.8	0.6	3.9		2.7
CZ	0.9	9.7	5.0	2.7	0.8	10.3	4.9
DE	1.9	6.0	3.8	2.7	3.7	5.8	4.0
DK	2.6	3.3	3.2	2.2	3.4		2.9
EE	4.8	5.8	3.8	3.5	4.1	5.4	4.6
ES	2.5	3.6	1.6	2.2	4.1	4.4	3.1
FI	0.6	7.8	2.8	3.0	0.3		2.9
FR	0.2	9.2	13.0	4.8	1.3	14.9	7.2
HU	4.5	6.8	4.1	2.5	0.5	11.2	4.9
IE	3.1	8.1	4.7	4.8	5.3	3.4	4.9
LT	4.3	8.6	1.7	0.9	0.3	11.1	4.5
NL	4.2	5.5	7.4	3.7	1.1	7.0	4.8
NO	3.0	4.7	1.5	2.5	0.5		2.4
PL	2.3	3.0	2.6	1.1	0.2	1.7	1.8
PT	0.0	7.7	3.3	6.6	0.6	5.1	3.9
SE	0.1	4.3	1.8	1.8	2.5		2.1
SI	3.2	7.1	0.4	8.1	1.5	6.8	4.5
UK	2.3	8.4	6.3	3.8	2.5	3.4	4.4
mean	2.2	6.3	3.8	3.1	2.2	7.1	3.9
min	0.0	3.0	0.4	0.6	0.2	1.7	1.8
max	4.8	9.7	13.0	8.1	5.3	14.9	7.2

The size of D varies both between countries and between variables. In ESS 7, the largest dissimilarity pertains to the variable household size in France with a D of 14.9 (see Table 7). The mean value of D across all variables and countries is 3.9.[11] This means that – on average – less than 4% of respondents in ESS would have to change categories in order to achieve the same distribution as in the LFS. D is highest for the variables age (mean 6.3) and household size (mean 7.1). To some extent, this is the consequence of these two variables having a larger number of categories than the remaining variables. The smallest D refers to the variables gender (mean 2.2) and nationality (mean 2.2). The latter is – at least in part – related to the highly skewed distribution of this variable.

The mean value of D across the six variables varies between a low of 1.8 in Poland and a high of 7.2 in France (see Figure 1). Countries with a rather high average D typically show values well above-average in several variables (see Table 7).

Figure 1: Average level of dissimilarity (mean D across six variables); ESS 7

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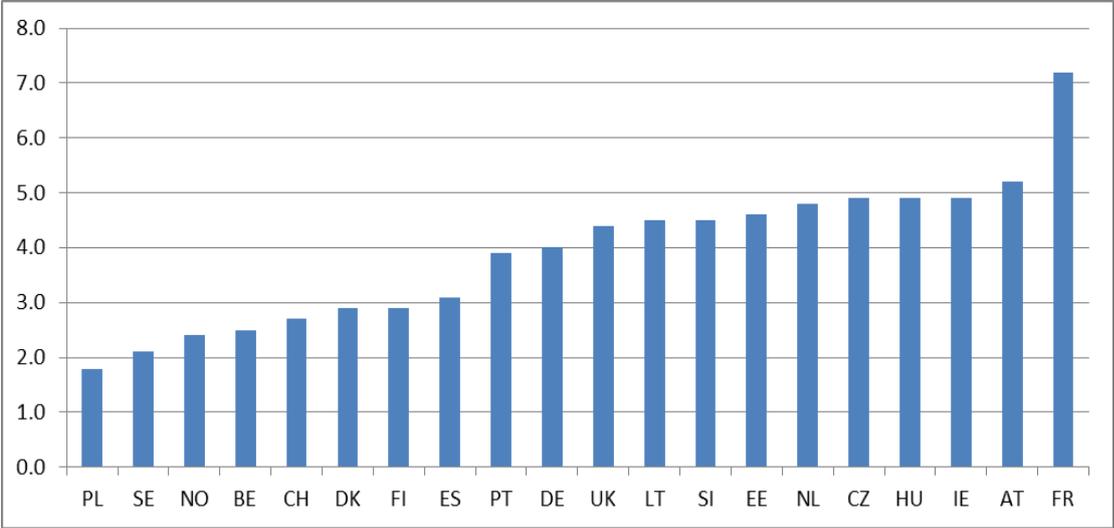
[11] The index of dissimilarity for household size is not available in five countries (Denmark, Finland, Norway, Sweden, and Switzerland). For these countries, the average value of D is based on the remaining five variables.

Level of dissimilarity across ESS rounds

For all countries except Austria, we can compare the dissimilarity indices for the six variables in ESS 7 with the respective information from ESS 5 and 6. Table 8 displays the mean dissimilarity indices for ESS 5, 6 and 7 for the 19 countries which participated in ESS 5, 6 and 7. On average across all countries and variables, we do observe a slight increase in the level of dissimilarity between ESS and LFS from round 5 to 7. The mean D across the six variables and 19 countries increases from 3.4 in ESS 5 to 3.8 in ESS 7. In particular, the variables 'marital status' and 'household size' reveal larger differences on average in round 7 than in round 5.

Table 8: Average level of dissimilarity of individual variables (mean D across countries) by ESS round; ESS 5, 6 and 7; 19 countries*

Variable	ESS 5	ESS 6	ESS 7
Gender	2.7	2.0	2.3
Age	5.5	5.3	6.3
Marital status	2.7	3.4	3.8
Work status	3.3	3.2	3.2
Nationality	2.0	2.1	2.0
Household size	5.2	5.6	6.6
Mean D across 6 variables	3.4	3.5	3.8

* 19 countries participating in ESS 5, 6 and 7. Information on marital status missing for FI in ESS 5. No information on household size available in LFS for CH, DK, FI, NO, and SE in each survey round

Figure 2 shows that in the majority of countries, the average D across the six variables is rather stable across the three survey rounds. The two most noticeable exceptions are Portugal and France. In Portugal, the average D for round 6 and 7 is considerably lower than in round 5. In France, the average D for round 7 is much higher than the average D in round 5 and 6. Accordingly, we might say that in Portugal quality has improved with respect to the demographic sample composition, whereas in France demographic sample quality became worse over the last three ESS rounds. One might speculate whether in both countries changes in the survey organisation appointed and/or the sample design used have contributed to these changes in the level of misrepresentation.

Figure 2: Average level of dissimilarity (mean D across six variables) by ESS round; ESS 5, 6 and 7; 20 countries
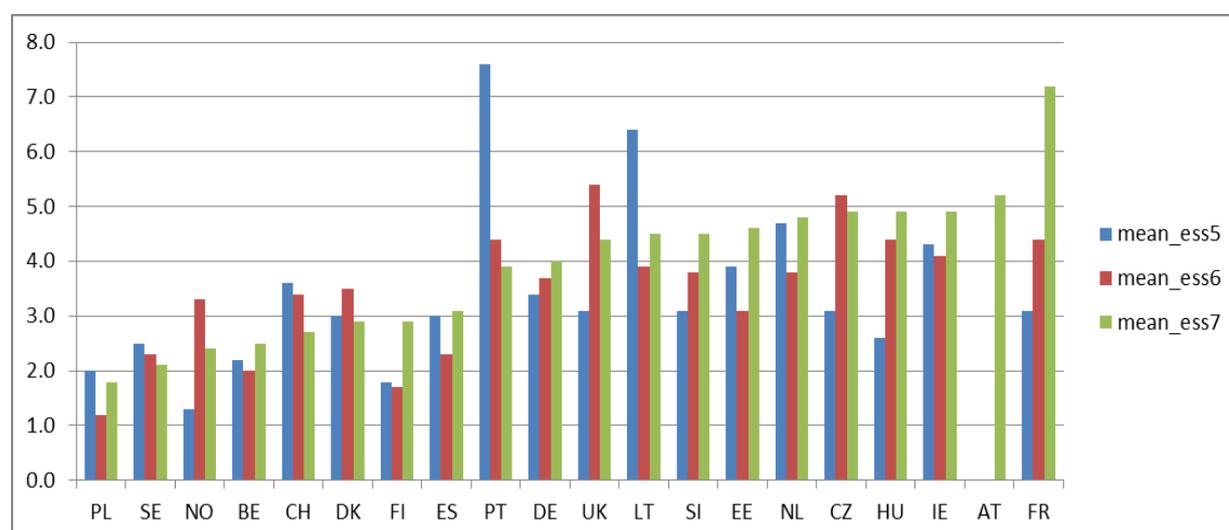
Figure 3 allows for a quick overview on stability and change in the level of differences for each variable separately. The prevailing picture is stability. For many countries, the indices of dissimilarity for the six variables do not differ much between the three survey rounds. As regards the abovementioned changes in average D in Portugal and France, a closer look at the dissimilarity indices for individual variables reveals the source of these changes. In Portugal, the D for age, work status and household size is much higher in ESS 5 than in ESS 6 and ESS 7. In addition, the D for gender in Portugal is higher in ESS 5 and 6 than in ESS 7. In France, in particular the household size variable exhibits a higher level of dissimilarity from the LFS data for ESS 7 than for ESS 5 and 6. To a somewhat smaller degree, this also applies to the marital status variable.

Beside these shifts in Portugal and France, there are some further instances where countries exhibit considerable changes in the size of D for individual variables. In Lithuania, the index of D for gender is much larger in ESS 5 than in ESS 6 and 7. In Ireland, the index of D for work status decreased considerably from round 5 to round 6 and 7. While these changes reflect an improvement in consistency over time between ESS and LFS data, there are other instances where we see a change in the opposite direction. The index of D for nationality in Ireland, for example, is higher in ESS 6 and 7 than in ESS 5. In Czech Republic and Hungary, the index of D for household size in ESS 6 and 7 is higher than in ESS 5. A change for the worse can also be observed for the variable age in Czech Republic and in Finland.

The country-specific results which are documented in the Appendix allow checking further details of the changes. A cursory inspection reveals that these shifts typically reflect a change in the level (and not in the direction) of misrepresentation. In Ireland, for instance, persons in paid work were underrepresented in each round of ESS, however to a different degree. While in ESS 5 the underrepresentation in relation to the LFS data was minus 11 percentage points, the underrepresentation was only minus 7 percentage points in ESS 6 and minus 5 percentage points in ESS 7.

Different reasons might have contributed to such changes in the level of discrepancies. Dedicated efforts of countries to improve sample composition by implementing targeted fieldwork efforts might be one of them. As regards the abovementioned example of Ireland, one might wonder whether the better representation of the working population was the result of a change in the number and/or timing of interviewers' call attempts. 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. Finding out about the concrete reasons usually will require 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.), and often additional analyses, e.g. of the ESS contact forms data, will be advisable.

Figure 3: Indices of dissimilarity ESS vs. LFS;
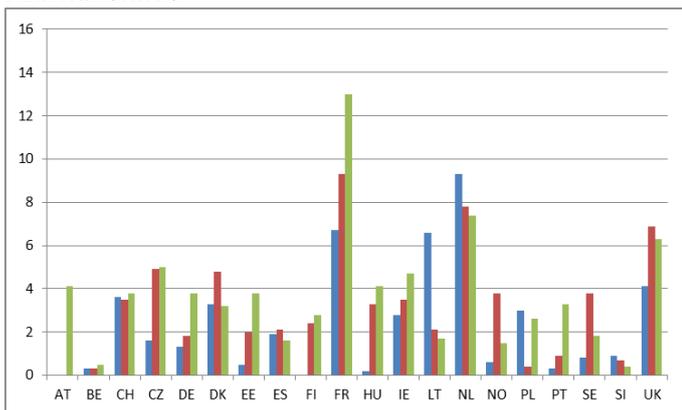ESS 5 (blue), ESS 6 (red), ESS 7 (green)
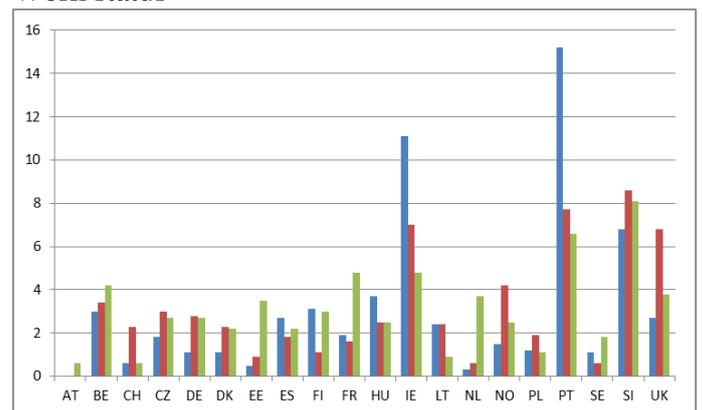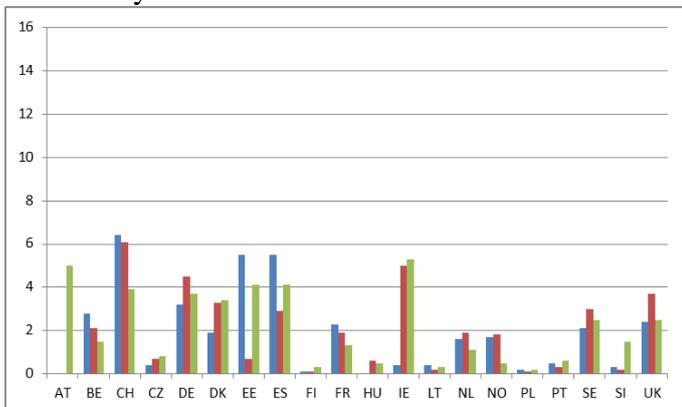(20 countries in ESS 7)
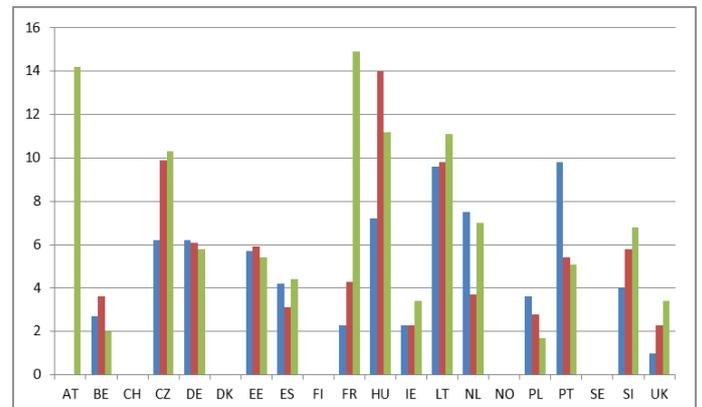
Gender

Age
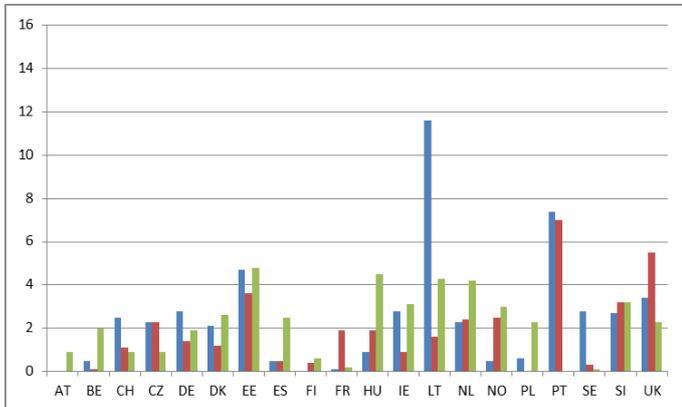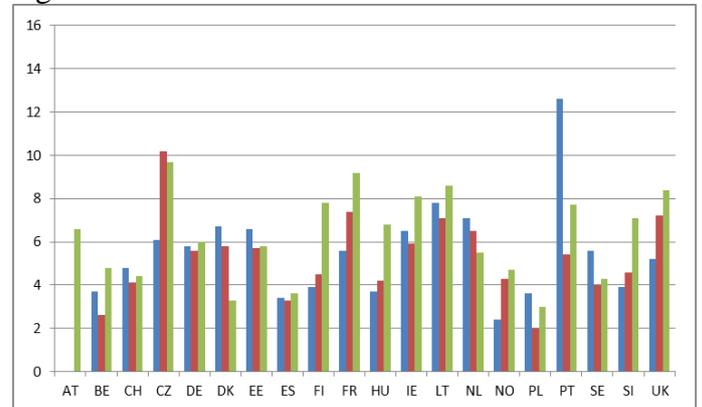
Marital status

Work status

Nationality

Household size

7. Correlates of misrepresentation

The previous sections revealed that the size of the differences between the ESS and the LFS data varies between countries. In the following, we will briefly analyse on the country-level whether the extent of misrepresentation is related to two basic survey parameters: the response rate achieved and the sample design used. Since the results we obtain for ESS 7 are very similar to those of previous analyses for ESS 5 and 6 (see Koch 2016), we pool the data from ESS 5, 6 and 7. That way, the number of cases for our analyses can be increased to 67 country-round combinations (23 countries from ESS 5, 24 countries from ESS 6 and 20 countries from ESS 7).

Misrepresentation and response rates

For several decades, the response rate achieved in a survey has been used as a proxy for the degree of nonresponse bias and the quality of the realised sample (Kreuter 2013). With respect to the present analysis, we therefore should expect that higher response rates come along with smaller ESS-LFS differences. Empirically, this is not the case. Figure 4 provides no evidence for a negative relationship between the average index of dissimilarity across the six variables and the response rate achieved in a country (Pearson's r = .07).

Figure 4: Average level of dissimilarity (mean D across six variables) by response rate (in %); ESS 5, 6 and 7 (n=67)
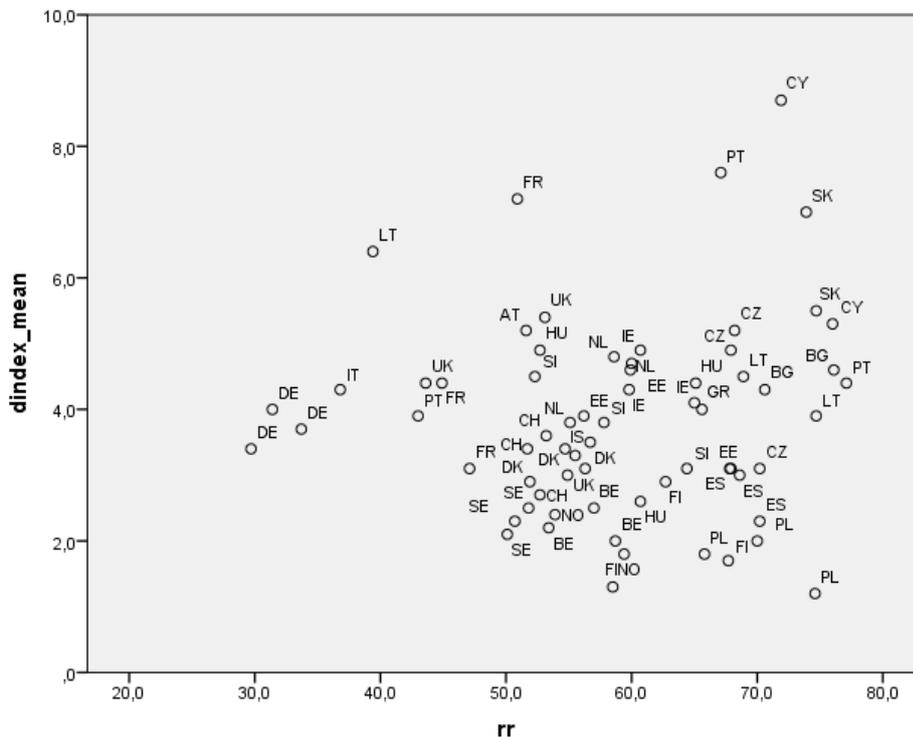
Misrepresentation and type of sample

More pronounced differences can be found when we turn to the type of sample used (see Figure 5). The ESS prefers countries to use a sample of named individuals from a register (European Social Survey 2013). If such a sample is used, the sample is drawn without any involvement of the interviewers. Where a sampling frame of individuals is not available,

countries may use a sampling frame of households or addresses. This can take various forms, differing, among other things, in the degree of interviewer involvement required. One possibility is that the households/addresses come from a list, like a registry of telephone numbers or the customer directory of an electricity provider. If such a list is not available, the survey organisation and their interviewers have to enumerate the households, usually before fieldwork starts. In all household/address designs, the interviewers have to select a respondent in the household, since the ESS interviews only one person per household. Among the 67 country-round combinations included in our analyses, 39 cases used a sample of individuals and 28 cases used a sample of households or addresses. Figure 5 shows the average index of dissimilarity, separately for countries using a sample of individuals and for countries using a household/address sample. In the group of countries using a sample of households/addresses, the average size of D is considerably higher (mean = 4.9) than in the group of countries using a sample of individuals (mean = 3.0).

Figure 5: Average level of dissimilarity (mean D across six variables) by type of sample; ESS 5, 6 and 7 (n = 67)

Interaction between type of sample and response rate

When the relationship between the response rate and the index of dissimilarity is analysed separately for countries with a sample of individuals on the one hand, and countries with an address/household sample on the other hand, an interesting pattern emerges (see Figure 6). For countries with samples of individuals, a moderate negative relationship between the response rate and D can be observed (Pearson's r = -.41). In line with conventional wisdom, the sample composition of countries with a higher response rate corresponds closer with the LFS data than the sample composition of countries with a lower response rate. For countries using a sample of addresses/households, however, this is not true. Here, a weak tendency to larger discrepancies with higher response rates shows up (Pearson's r = .14).
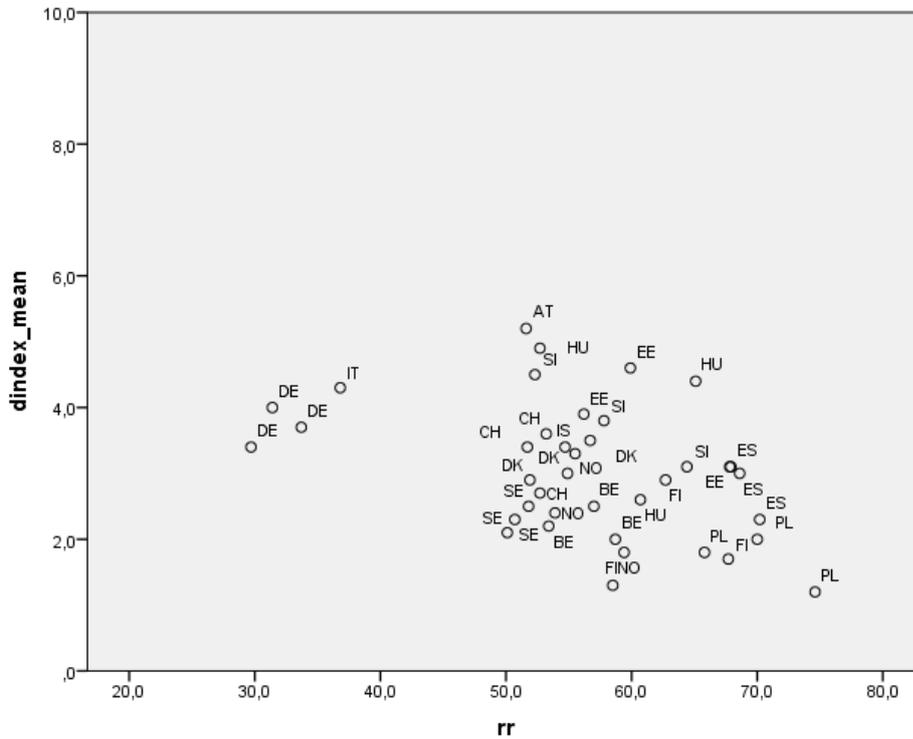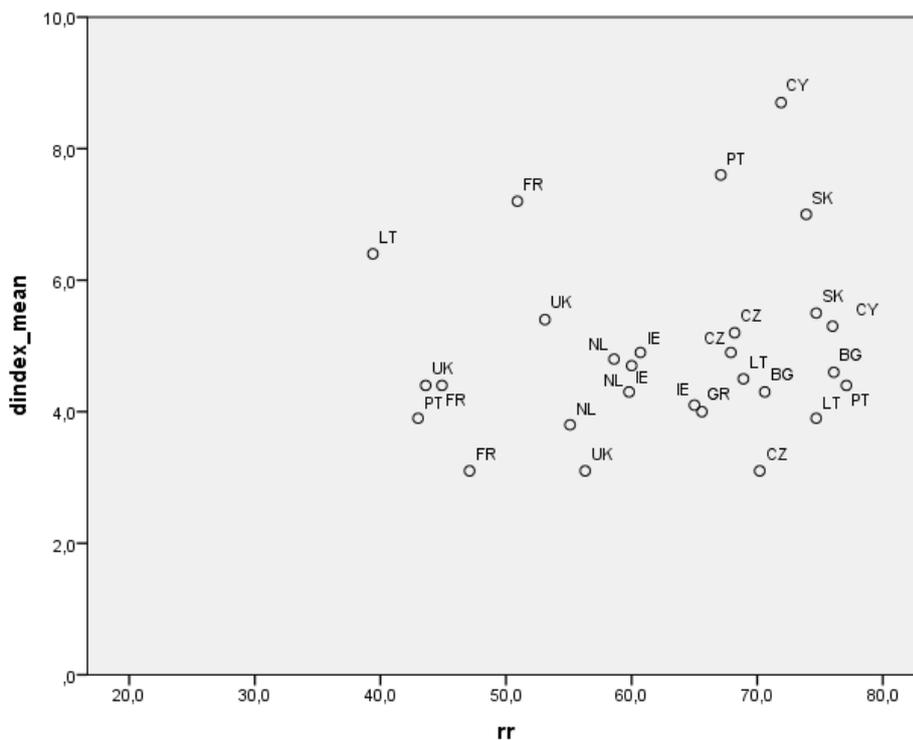
Figure 6: Average level of dissimilarity (mean D across six variables) by response rate (in %),
separately for countries with samples of individuals and countries with samples of
households/addresses; ESS 5, 6 and 7

Sample of individuals: Pearson's r = -.41 (n = 39)

Sample of households/addresses: Pearson's r = .14 (n = 28)

The same basic pattern can be observed when the correlation with the response rate is calculated separately for each of the six variables (see Table 9). Among the country-round combinations using a sample of individuals, we find a negative correlation between the response rate and the index of dissimilarity for each of the six variables. Among the countries using a sample of households/addresses, a positive correlation shows up for all variables except one (marital status).

Table 9: Correlation between response rate and index of dissimilarity by type of sample, separately for each variable (Pearson's r); ESS 5, 6 and 7

Variable	Sample of individuals (n=39)	Sample of households/addresses (n=28)	Total (n=67)
Gender	-.10	.12	.13
Age	-.44	.31	.18
Marital status*	-.11	-.46	-.13
Work status	-.13	.19	.13
Nationality	-.42	.12	-.11
Household size**	-.21	.12	-.01
Mean D across 6 variables	-.41	.14	.07

* Sample of individuals: n = 38; total: n = 66
** Sample of individuals: n = 23; total: n = 51

In section 5, we hypothesised that differential response propensities of demographic subgroups in combination with the effort interviewers exert were the main determinants of the level and patterns of misrepresentation we find in the data. This reasoning alone, however, cannot explain the present findings. An explanation might be found when we assume that interviewers do not always adhere to the survey standards set in the ESS. As already mentioned, interviewers play an important role in sample selection when a sample of households/addresses is used. If interviewers do not follow the rules of probability sampling properly – for instance during the listing of households before fieldwork, or during the selection of target persons within households – this might contribute to the patterns of results we observe. Some interviewers may preferentially select 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 (see Kohler 2007; Sodeur 1997). In addition, such a misconduct of interviewers might not only increase sample bias, but will also undermine the reliability of the response rate calculated. Undocumented substitution of reluctant target persons, as this behaviour might be termed, can lead to both demographic misrepresentation *and* to inflated response rates. Hence, this process might explain why sometimes (seemingly) high response rates come along with large biases in sample composition. Finding out whether and how often undocumented substitution actually occurs requires having in place a dedicated system of interviewer back-checks (as regards the within-household selection of respondents, see Koch 2018).[12]

---

[12] A general note of caution is in place when interpreting these macro-level results. The data are of an observational (non-experimental) nature only. Therefore, we cannot rule out that some countries are more strongly affected by differential nonresponse of certain population subgroups than other countries (independently from the type of sample they use and the response rate they achieve). It seems rather unlikely, however, that such country-specific differences led to the patterns of results we found.

8. Post-stratification adjustment as a solution?

Using post-stratification adjustments is, in principle, a cost-efficient approach to improve survey representativeness. Well-designed post-stratification (PS) weights can correct for sampling, coverage, and nonresponse errors. Applying PS-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 PS-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 (see Berzelak et al. 2016). The post-stratification weights are obtained by adjusting the ESS design weights 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. 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 10 shows the average indices of dissimilarity for the six variables, both without and with applying the ESS PS-weights. On average across all countries and variables, the level of dissimilarity between ESS and LFS is reduced by about one-third when the PS-weights are used (mean D of 3.9 vs. mean D of 2.6).

Table 10: Effect of using PS-weights: Average level of dissimilarity (mean D across 20 countries*); ESS 7

Variable	ESS 7 (dweight)	ESS 7 (PS-weight)	Relative reduction in average D, after applying PS-weight
Gender	2.2	0.2	91%
Age	6.3	4.3	32%
Marital status	3.8	2.5	34%
Work status	3.1	1.8	42%
Nationality	2.2	1.9	14%
Household size	7.1	6.2	13%
Mean D across 6 variables	3.9	2.6	33%

* Household size: 15 countries; no information available in LFS for CH, DK, FI, NO, and SE.

The level of reduction varies between variables. The largest relative reduction pertains to the variable 'gender' (-91%), followed by the variables 'work status' (-42%), 'marital status' (-34%), and 'age' (-32%). The smallest reductions pertain to the variables 'nationality' (-14%) and 'household size' (-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

---

[13] For ESS 7, information from LFS 2014 has been used. In a few countries, information on education (Ireland, Poland, Switzerland and United Kingdom) or region (Estonia, Lithuania) was not included in the weighting procedure. In Austria, Finland, and the Netherlands, the information on region came not from the LFS, but from other sources (see Berzelak et al. 2016, p. 5ff).

PS-weight. One usually would expect that the variables included as control in the PS-weight will show a more or less perfect fit with the benchmark data.[14] Also, the variable 'age' has been used as a control for the PS-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 PS-weight, only three different age groups have been distinguished. In the present comparison, however, we use a more detailed categorisation with seven age groups.

Figure 7 shows that the level of improvement when applying PS-weights also varies between countries. In all countries but one, the introduction of PS-weights reduces the average size of differences between ESS and LFS. In relative terms, the reduction in the mean index of dissimilarity across the six variables is largest in Finland (-66%) and smallest in Spain (-13%). In France, which is the country with the largest average deviation when applying only the ESS design weight, we see the most noticeable reduction in absolute terms: the average index of dissimilarity decreases from 7.2 to 4.2 after applying the PS-weights. In Austria, however, the use of PS-weights even slightly increases the average level of misrepresentation (+1.9%, in relative terms). As a consequence, Austria is the country with the largest average deviation from the LFS data when applying the ESS PS-weight.

Figure 7: Average level of dissimilarity (mean D across six variables); dweighted (blue bars) and PS-weighted (red bars) data; ESS 7
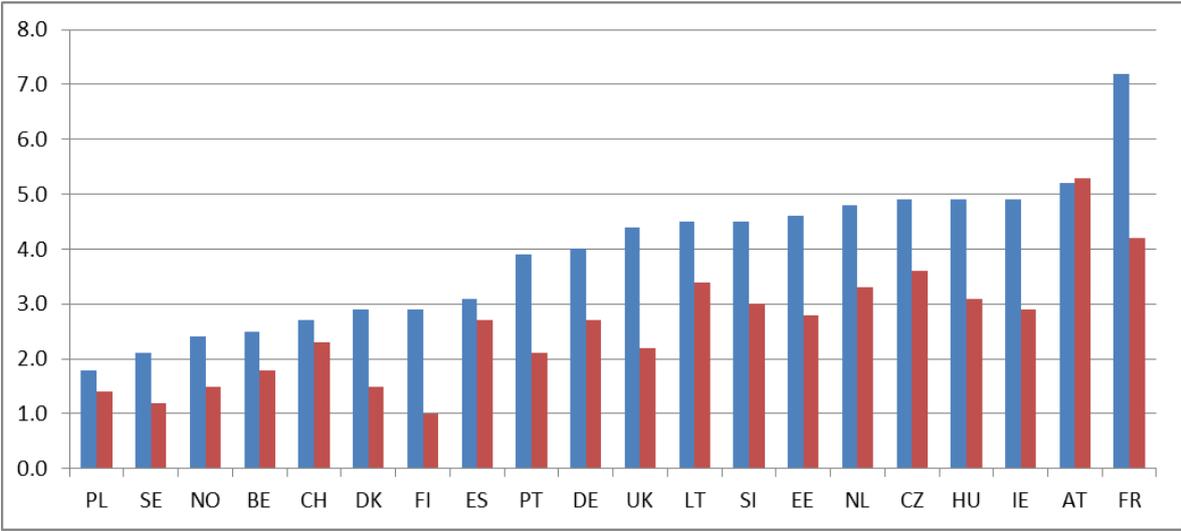

Figure 8 shows dweighted and PS-weighted results separately for the six variables of our analysis. Again, we see that the effect of the PS-weight varies both between countries and between variables. Usually, the PS-weight decreases the level of misrepresentation. Every now and then, however, countries exhibit an increase of ESS-LFS differences at a certain variable when the PS-weight has been used.
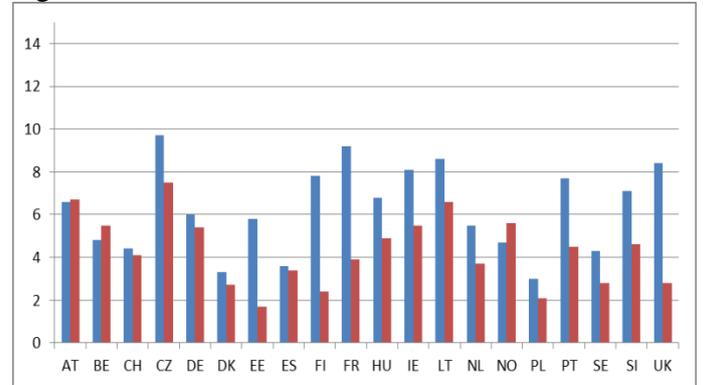
---

[14] In the present case, slight differences still may arise due to differences in the timing of the variables (PS-weight: always LFS 2014; present benchmark data: LFS 2014 vs. LFS 2015, depending on the timing of ESS 7 fieldwork).

Figure 8: Indices of dissimilarity ESS vs. LFS;
dweighted (blue bars) and PS-weighted
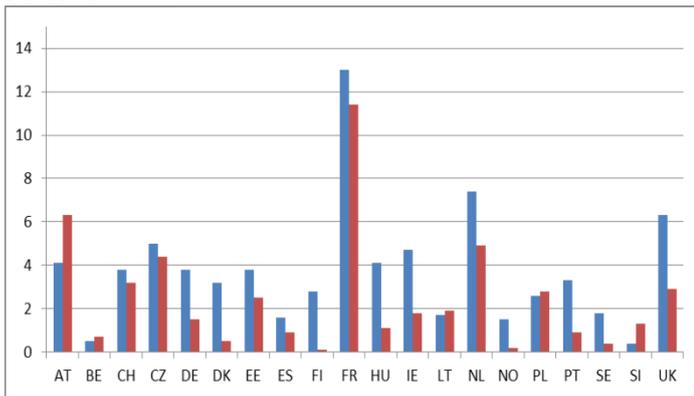(red bars) data; 20 countries in ESS 7
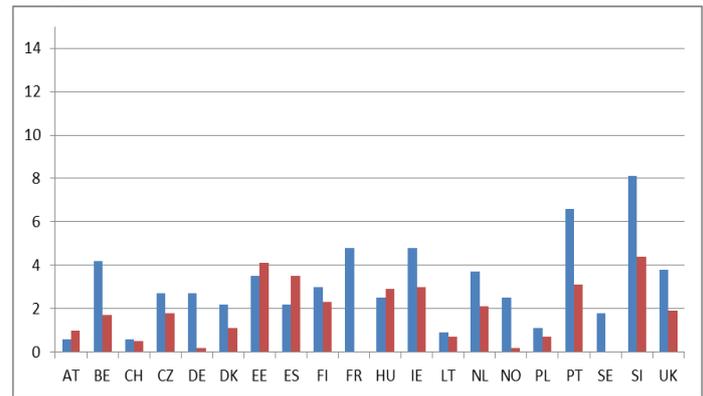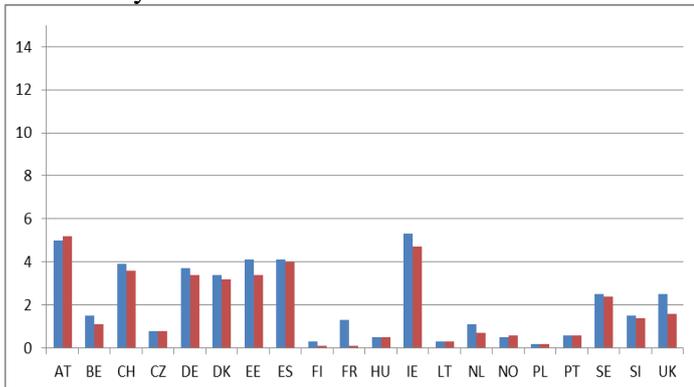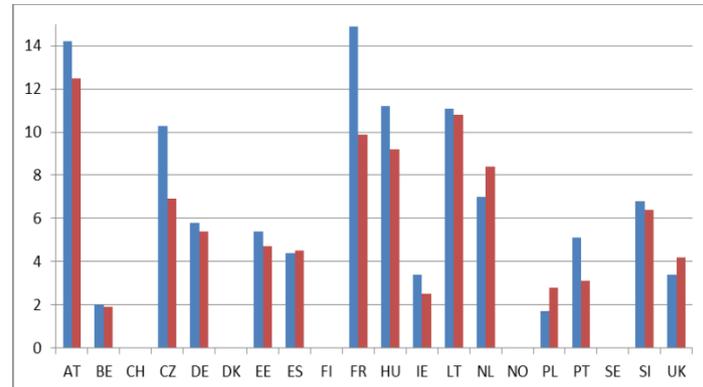
Gender

Age
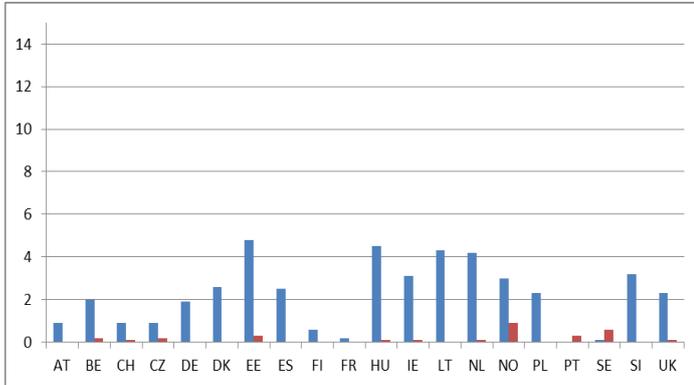
Marital status

Work status

Nationality

Household size

Countries differ in the patterns of demographic misrepresentation and in the relations that exist between the control variables included in the PS-weight and the variables of interest. These differences may lead to differences in the effects of the ESS PS-weight between countries, as can be seen here. All in all, this little exercise shows that the standard ESS PS-weight is no panacea to deal with the issue of demographic misrepresentation in ESS.

9. Summary

(1) This paper used external benchmark data to analyse the socio-demographic sample composition in ESS 7. In several ESS countries, the comparison of six demographic variable distributions (gender, age, marital status, work status, nationality, and household size) with data from the LFS revealed only small differences. At the same time, large differences were observed in a number of other countries – sometimes only with respect to one or two variables, sometimes with respect to most of the variables examined. The basic pattern of results for ESS 7 was rather similar to the one received from an analogous exercise for ESS 5 and 6.

A detailed follow-up of the reasons for the differences in sample quality between individual countries was not part of the present task. Generally speaking, 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 in a country, then an underrepresentation of this subgroup will occur.

(2) A correlational analysis including ESS 7 plus ESS 6 and ESS 5 data showed that the type of sample used in a country helps in predicting the size of deviation from the LFS. ESS countries relying on a sample of individuals exhibited smaller discrepancies from the LFS on average than countries using a household/address sample. When the relationship between the response rate and sample quality was analysed separately for countries with a sample of individuals on the one hand, and countries with a sample of households/addresses on the other hand, different patterns were observed. While for countries with a sample of individuals sample representativeness was positively related to the response rate (as one would usually expect), no such correlation could be observed for countries with a sample of households/addresses.

This result cannot be explained sufficiently by referring exclusively to potential differences in subgroups' response propensities. A satisfactory explanation might be achieved by taking the interviewer and his/her behaviour into account. Interviewers play an important role in sample selection when a sample of households/addresses is used. If interviewers do not follow the rules of probability sampling properly (e.g. by substituting 'difficult' sample units by sample units which can be interviewed more easily), this might have contributed to the patterns of results we observed.

(3) 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 (PS) 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 PS-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 PS-weights were applied. Thus, the standard ESS PS-weight is no panacea to deal with demographic misrepresentation. In light of these findings, aiming for balanced response rates during fieldwork continues to be an important goal.

(4) The ESS strives to implement a policy of quality improvement from round to round. The results of the sample composition assessment for ESS 7 were fed back to the countries on an individual basis in spring 2018. That way, the information could be taken into account when countries drew up their national fieldwork plans for ESS round 9. Countries were invited to discuss and analyse potential reasons for the discrepancies observed, and to consider measures to improve in ESS 9. 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 (see Appendix).

References:

Berzelak, N., Slavec, A. & Vehovar, V. (2016): ESS Round 7 Weighting Procedures. University of Ljubljana, Faculty of Social Sciences, Centre for Social Informatics; November 2016

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

European Social Survey (2013). Round 7 Specification for Participating Countries. London: Centre for Comparative Social Surveys, City University London

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

Eurostat (2015a). Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2014. 2015 edition. eurostat Statistical working papers

Eurostat (2015b). Quality report of the European Union Labour Force Survey – 2014. 2015 edition. eurostat Statistical working papers

Eurostat (2016a). Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2015. 2016 edition. eurostat Statistical working papers

Eurostat (2016b): EU Labour Force Survey Database, User Guide, Version: November 2016

Eurostat (2017). Quality report of the European Union Labour Force Survey – 2015. 2017 edition. eurostat Statistical Reports

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

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

Koch, A. (2018) (Forthcoming). Within-Household Selection of Respondents. In T. P. Johnson, B. Pennell, I. A. L. Stoop, and B. Dorer (eds.), Advances in Comparative Survey Methods: Multinational, Multiregional and Multicultural Contexts (3MC). Wiley Series in Survey Methodology. New York: John Wiley & Sons

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

Kohler, U. (2007). Surveys from inside: An assessment of unit nonresponse bias with internal criteria. Survey Research Methods, 1, 55-67

Kreuter, F. (2013). Facing the Nonresponse Challenge. The ANNALS of the American Academy of Political and Social Science, 645, 23-35

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

Sodeur, W. (1997). Interne Kriterien zur Beurteilung von Wahrscheinlichkeitsauswahlen. ZA-Information, 41, 58-82

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

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

20 countries were included in the present analyses (see list below). As of April 2018, each of these countries planned to participate in ESS 9, either as ESS ERIC Member, Observer or Guest country. The countries were sent the following document by their respective country contact in time to be taken into account when planning fieldwork for ESS 9.

AT - Austria

BE - Belgium

CH - Switzerland

CZ - Czech Republic

DE - Germany

DK - Denmark

EE - Estonia

ES – Spain

FI - Finland

FR - France

HU - Hungary

IE - Ireland

LT - Lithuania

NL - Netherlands

NO - Norway

PL - Poland

PT - Portugal

SE - Sweden

SI - Slovenia

UK - United Kingdom

Assessment of demographic sample composition in ESS 7:

Individualised feedback to countries

Dummy country

Achim Koch

GESIS Leibniz Institute for the Social Sciences, Mannheim, Germany

(April 2018)

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 7 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)[15]

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. [16] Depending on the timing of ESS Round 7 fieldwork in a country, either LFS 2014 or LFS 2015 data 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.[17] 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.[18]

---

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

[16] To learn more about procedures and variables, you may look at the respective report for ESS 6 ("ESS 6 Sample Composition Assessment") where a similar exercise was implemented (Koch 2016).

[17] In Estonia, Norway and Sweden persons aged 75 years and older were not included in the LFS. For these countries, the data from the ESS also had to be restricted to the population aged 15 to 74 years.

[18] In (nearly) all LFS countries data on gender, age and region were used for the adjustment. A number of countries included additional variables in weighting, like information on unemployment or nationality (see Eurostat 2015). 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 (assuming 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. 95% confidence intervals for the ESS estimates were calculated to indicate whether the difference between ESS and LFS is within the limits of sampling error.[19] When the confidence intervals do not overlap with the estimate from the LFS, we interpret this as an indication of a significant over- or underrepresentation with respect to that specific estimate.

1.2 Results

Table 1 shows the results of the ESS-LFS comparison for your country in detail. For ESS 7, the respective estimates and the upper and lower limits of the 95% confidence intervals are shown. Differences between ESS 7 and LFS 2014/2015 were highlighted when the LFS data are outside the 95% confidence interval of the ESS estimate. For the purpose of comparison, the respective differences for ESS 5 and 6 are also included.

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 improve in ESS 9 (see section 3 below)

- Discuss the results and potential measures with your country contact, with your SWEP expert and/or with the fieldwork team

---

[19] The geographical clustering of the sample was taken into account in these calculations, as the majority of countries in ESS 7 (14 out of 20) used a geographically clustered sample. Only six countries used an unclustered design; these were Denmark, Estonia, Finland, the Netherlands, Sweden and Switzerland.

Table 1: Dummy country

Category	ESS 7			LFS 2014/15	Diff ESS-LFS ESS 7		Diff ESS-LFS ESS 6	Diff ESS-LFS ESS 5
	Lower CI limit	Estimate	Upper CI limit					
	(%)	(%)	(%)	(%)	(%)		(%)	(%)
Female								
Age 15-24 y. Age 25-34 y. Age 35-44 y. Age 45-54 y. Age 55-64 y. Age 65-74 y. Age 75+ y.								
Married								
In paid work								
Non-national								
1pers hh 2pers hh 3pers hh 4pers hh 5+pers hh								

Highlighted difference: LFS 2014/15 estimate outside 95% confidence interval of ESS 7 estimate

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 7.[20] 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. Dark green or dark red cells, respectively, indicate that the LFS estimate is outside the 95% confidence interval of the ESS estimate. Thus, it can easily be checked whether the structure of demographic misrepresentation is similar across countries.

According to Table 2, significant differences between ESS and LFS estimates show up at each of the six variables in at least half of the 20 countries which were included in the comparison. Broadly speaking, we can state that – *in case significant differences occur* – the following patterns of under-/overrepresentation prevail:

Underrepresented are	- Young age groups (15-44 years) and the oldest age group (75+ years) - Non-nationals - Persons living in 1-person households
Overrepresented are:	- Females - Middle-aged persons (45-74 years) - Married persons - Persons living in 2-person households
Mixed pattern:	- People in paid work (underrepresented in some countries, overrepresented in others)

---

[20] In total, 21 countries participated in ESS 7. Israel was not part of the LFS, and thus could not be included in the present comparison.

Table 2: Differences between ESS 7 and LFS 2014/2015 estimates (in percentage points)*

	Female	15-24 y.	25-34 y.	35-44 y.	Age 45-54 y.	55-64 y.	65-74 y.	75+ y.	Married	In paid work	Non-national	HH-size 1p-hh	2p-hh	3p-hh	4p-hh	5+p-hh
AT	0.9	-4.7	0.0	1.9	-1.4	1.3	3.4	-0.4	-4.1	-0.6	-5.0	11.6	2.6	-5.1	-4.3	-4.8
BE	-2.0	2.0	-1.8	-0.6	0.9	1.2	0.7	-2.3	-0.5	4.2	-1.5	-0.4	-1.6	0.0	0.5	1.5
CH	-0.9	1.7	-2.6	-0.2	0.1	0.7	2.0	-1.5	3.8	-0.6	-3.9					
CZ	0.9	4.5	-3.0	0.9	3.0	1.3	-1.3	-5.4	5.0	2.7	-0.8	-5.2	-3.1	7.7	2.5	-2.0
DE	-1.9	0.1	-1.7	-1.5	1.8	2.5	1.6	-2.8	3.8	2.7	-3.7	-5.6	1.8	-0.2	1.1	2.8
DK	-2.6	-1.3	-0.6	-0.9	1.1	1.3	0.8	-0.5	3.2	2.2	-3.4					
EE	4.8	-3.6	-0.7	-1.2	-0.3	3.2	2.6		3.8	3.5	4.1	-1.0	3.7	1.7	-1.8	-2.6
ES	-2.5	-0.2	-0.3	-2.1	1.4	1.1	-1.0	1.1	1.6	2.2	-4.1	-3.6	2.5	-0.6	1.9	-0.1
FI	-0.6	-3.2	-2.6	-1.7	-0.3	3.0	3.7	1.0	2.8	-3.0	-0.3					
FR	-0.2	-1.7	0.3	6.2	2.7	-2.2	-1.9	-3.4	13.0	4.8	-1.3	-7.5	-7.4	0.4	7.5	7.0
HU	4.5	-2.2	-2.5	-2.2	1.1	1.1	3.6	0.9	4.1	2.5	-0.5	3.3	7.0	0.9	-6.1	-5.1
IE	3.1	-0.7	-3.9	-2.4	1.4	4.3	2.3	-1.1	4.7	-4.8	-5.3	1.2	0.9	-1.0	1.2	-2.4
LT	4.3	1.9	-3.2	0.6	3.3	0.2	2.5	-5.4	1.7	0.9	-0.3	-8.7	2.2	6.6	2.3	-2.3
NL	4.2	-1.1	-2.8	-0.8	1.5	2.6	1.4	-0.7	7.4	-3.7	-1.1	-6.6	-0.5	2.3	4.1	0.5
NO	-3.0	0.0	-3.8	-0.9	0.6	1.4	2.6		1.5	2.5	-0.5					
PL	2.3	0.6	-1.7	-0.1	-1.2	0.4	1.2	0.8	-2.6	-1.1	-0.2	0.0	-1.7	0.8	0.6	0.2
PT	0.0	-0.8	-3.1	-3.4	-0.4	0.7	5.5	1.5	3.3	-6.6	0.6	0.3	4.1	-5.0	-0.1	0.7
SE	0.1	-2.2	-1.0	-1.0	-0.1	1.2	3.1		1.8	1.8	-2.5					
SI	3.2	0.0	-2.4	-2.9	-1.3	3.6	3.5	-0.5	0.4	-8.1	-1.5	-4.3	2.0	-1.4	-1.0	4.8
UK	2.3	-4.9	-3.6	1.0	1.4	2.8	2.8	0.3	6.3	-3.8	-2.5	3.4	-0.5	-2.4	-0.3	-0.1
# sign. diff.	7+ / 2-	2+ / 6-	0+ / 13-	2+ / 7-	4+ / 0-	7+ / 1-	13+ / 1-	0+ / 6-	11+ / 2-	6+ / 6-	1+ / 13-	3+ / 7-	6+ / 2-	3+ / 3-	4+ / 2-	3+ / 5-

* green cells = overrepresentation;

red cells = underrepresentation;

dark green / dark red cells = LFS estimate outside 95% confidence interval of ESS estimate;

EE, NO, SE: persons 75 years or older not included; CH, DK, FI, NO, SE: no LFS data on HH-size available

3. Further considerations on misrepresentation

The following remarks provide some further pieces of information on 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 behavior, and (3) other causes.[21]

(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 Leuven quality report for ESS 7 (Beullens et al. 2016) or running some dedicated analyses with the contact form data of your country can provide insights in that respect.

(2) Interviewer behavior

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.

In other instances, interviewer 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

---

[21] We do not take account of sampling error here.

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.[22] 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 (Koch 2016). 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 at least information on gender and age of the sample unit. This information can be used both 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.

---

[22] 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.

(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 (Berzelak et al. 2016). These weights have been constructed using information on gender, age group, education and region (mainly from the LFS). Analyses with the ESS PS-weights show that the level of discrepancies between ESS and LFS usually decreases when these weights were applied. 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),
- allocation of interviewers to sample cases (best interviewers to work on cases with lowest response propensity).

---

[23] In addition, one has to be aware that PS-weights neither provide improvements in precision, nor do they take account of any bias arising within weighting classes.

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

Berzelak, N., Slavec, A. & Vehovar, V. (2016): ESS Round 7 Weighting Procedures. University of Ljubljana, Faculty of Social Sciences, Centre for Social Informatics

Beullens, K., Loosveldt, G., Denies, K. & Vandenplas, C. (2016): Quality matrix for the European Social Survey, round 7. Leuven, KU Leuven, Centre for Sociological Research http://www.europeansocialsurvey.org/methodology/ess_methodology/data_quality.html

Eurostat (2015): Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2014. eurostat Statistical working papers

Similar exercises for ESS 6 and ESS 5:

Koch, A. (2016): Assessment of socio-demographic sample composition in ESS Round 6. Mannheim: European Social Survey, GESIS www.europeansocialsurvey.org/docs/round6/methods/ESS6_sample_composition_assessment.pdf

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 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, data weighted by DWEIGHT
LFS 2010, data file edition 2012, data weighted by COEFF

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

Email: achim.koch@gesis.org

Phone: +49-621-1246-401