



***Guide to Using Weights and Sample Design
Indicators with ESS Data***

Olena Kaminska

INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH, UNIVERSITY OF ESSEX

07 July 2020

V1.1

Contents

	Page Number
1. Introduction	1
2. Importance of using weights with ESS data	2
3. How to choose the correct weight	4
4. Data Preparation and analysis	5
4.1 Sample design indicators	5
4.2 Analysis weight	5
4.3 Specifying sample design	6
5. Impact of weighting and taking account of sample design	10
Acknowledgment	13
References	13

1. Introduction

This User Guide provides information on using ESS data with weights. We discuss why weighting data is important, what weights do to the estimates, how the ESS weights are calculated, and follow this with practical examples for taking weights into account alongside stratification and clustering in different statistical software packages.

2. Importance of using weights with ESS data

In order to represent a population an analyst needs to take into account the survey design correctly. Sample designs differ greatly across countries – they are designed to achieve a minimum effective sample size (which results in similarity of confidence intervals across countries) in the most cost effective way for each country, taking into account each country context. Thus, countries differ in the magnitude of selection probabilities, the variation among them, as well as clustering and stratification. The variation in selection probabilities is particularly important. In most ESS rounds half, or more, of the participating countries use an address-based sample, meaning that one person is selected at random at each address in a random sample of addresses. Consequently, people living alone have twice the selection probability of people living in a household containing two adults, three times the probability of people living in a household containing three adults, and so on. Failing to correct for this using weights will result in samples being heavily skewed towards people living alone and estimates consequently being biased towards the opinions of such people, for address-based countries.

Furthermore, countries differ in nonresponse processes. This is reflected not only in national differences in response rates, but in differences in the demographic and other composition among respondents. Post-stratification weights reduce the impact of nonresponse error. Additionally, they have the advantage of correcting for coverage and sampling errors with respect to the post-stratification variables.

Currently, ESS post-stratification is based on gender, age, education and geographical region. 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.

Without weights your estimates may be biased. This includes point estimates, such as means and proportions, measure of association such as regression coefficients, as well as estimates of variance and any quantities based upon them such as confidence intervals, hypothesis tests and therefore conclusions about group differences.

If weights and clustering are not specified, statistical software tends to assume that the data comes from a simple random sample with 100% response rate. This would imply that we spent many more millions on data collection than actually were spent, which would be misleading. But more importantly estimates based on such assumptions often have unpredictable bias. In particular, sample clustering and variation in selection probabilities both tend to increase the standard errors of estimates – often considerably. Failing to take this into account will lead to seriously under-estimated standard errors and, consequently, over-fitted models and biased hypothesis tests.

It is therefore crucial to use correct estimation procedures and always take into account weighting and clustering (note, stratification is optional and excluding it makes your standard error estimates conservative).

3. How to choose the correct weight

It is recommended that by default you should always use `anweight` (analysis weight) as a weight in all analysis. This weight is suitable for all types of analysis, including when you are studying just one country, when you compare across countries, or when you are studying groups of countries.

`anweight` corrects for differential selection probabilities within each country as specified by sample design, for nonresponse, for noncoverage, and for sampling error related to the four post-stratification variables, and takes into account differences in population size across countries. It is constructed by first deriving the design weight, then applying a post-stratification adjustment, and then a population size adjustment. Further details of how the weights are derived are documented in the round-specific report on the production of weights. Starting from Round 9, `anweight` is provided for you in the integrated data file. If you are using data from earlier ESS rounds, you can derive `anweight` yourself. We explain below how to do this.

For advanced users who may want to model a nonresponse correction themselves we provide a design weight (`dweight`). This weight corrects only for differential selection probabilities. Please note that we do not recommend use of this weight without non-response correction.

We also provide the post-stratified design weight (without the population size correction), `pspwght`. This weight can be used for single country analysis or for cross-country comparisons and for these types of analyses should give you the same estimates as `anweight`.

4. Data preparation and analysis

4.1 Sample design indicators

From round 9 onwards, all the necessary sample design indicators and weights are already included in the integrated (second release) data file, but if you are working with data from earlier rounds you will first need to merge the sample design indicators on to the main data file. For rounds 7 and 8, the sample design indicators are in the integrated SDDF (sample design data file), so you need to merge this file with the main integrated (questionnaire data) file. For rounds 1 to 6, sample design indicators are stored in a separate file for each country (and files are missing for some countries in some rounds), so you would need to merge several files. Furthermore, for these rounds the indicators `psu` and `stratify` have not been recoded in a manner suitable for cross-country analysis, so you will need to do this if you are analysing data from more than one country. Follow the guidance in section 2 of Kaminska & Lynn (2017) and ensure that each value is exclusive to one country.

4.2 Analysis weight

The analysis weight, `anweight`, is already included on the integrated main data file from round 9 onwards. For rounds 1 to 8, users should create this weight themselves, as follows. This weight is suitable for all analysis, including studying a single country, comparing across multiple countries or studying multiple countries as a group.

Box 1: Creating analysis weight in Stata

```
gen anweight=pspwght*pweight
```

Box 2: Creating analysis weight in R

```
library(survey)      # estimation  
  
data1[, anweight := pspwght * pweight * 10e3]
```

Box 3: Creating analysis weight in SPSS

```
compute anweight=pspwght*pweight.
```

4.3 Specifying sample design

Many commonly-used statistical software packages allow the user to specify the parameters of a survey sample design. Once specified, all subsequent analysis correctly takes into account clustering, stratification and weighting.

In Stata you tell the software that your data come from a specific sample design by specifying this with the `svyset` command. After this for any analysis you run within the same session you need to add the prefix `'svy:'` to the command. See also the next section for more examples in Stata. Similarly, sample design can be specified in R. For this, use the package called `'survey'` with the `svydesign` function. And in SPSS you would use the Complex Samples functions.

The clustering variable in ESS is called `'psu'`, stratification is indicated by `'stratum'`, and weighting by `'anweight'`. Boxes 4, 5 and 6 demonstrate how to specify the sample design in Stata, R, and SPSS respectively.

For SPSS users who prefer to use the menu rather than syntax, you can find everything under Analyze-Complex Sample. At first you have to specify your design under Analyze-Complex Sample-Prepare for Analysis (where you will save a new csplan file which will contain sample design information – so give a new name under 'browse'). After clicking 'next' you will be prompted to select strata (stratum), clusters (psu) and sample weight (anweight). You can now use this csplan file when analysing data.

For example, to estimate a proportion using the variable netusoft you can follow Analyze – Complex Sample – Frequencies - and specify csplan file under Plan File, click 'continue' and select variable (e.g. netusoft) for a frequency tables. Then you can select 'statistics' where you can choose 'table percent' and 'standard error', click 'continue' and 'ok' and you should now see estimates of proportions for netusoft.

Similarly to get an estimate of a mean follow Analyze – Complex Sample – Descriptives - and specify csplan file under Plan File, click 'continue' and select the variable (e.g. netustm). Then select 'statistics', choose 'mean' and 'standard error', click 'continue' and 'ok' and you should now see estimates of means for netustm.

Box 4: Specifying sample design in Stata

```
svyset psu [pweight=anweight], strata(stratum)

svy: /*INCLUDE YOUR ANALYSIS HERE*/

/*for example */

svy: proportion netusoft

svy: mean netustm
```

Box 5: Specifying sample design in R

```
library(survey)      # estimation

# Define sample design

design <- svydesign(ids = ~psu, strata = ~stratum, weights = ~anweight,
                  data = data1)

# Examples
# Values of netusoft
data1[, .N, keyby = .(netusoft)]
# Convert to factor (categorical variable)
data1[, netusoft := factor(netusoft)]
# Estimate proportions
svymean(x = ~netusoft, design = design)
# Estimate mean
svymean(x = ~netustm, design = design, na.rm = T)
```

Box 6: Specifying sample design in SPSS

```
* Analysis Preparation Wizard.
CSPLAN ANALYSIS
/PLAN FILE='I:\myfile.csaplan'
/PLANVARS ANALYSISWEIGHT=anweight
/SRSESTIMATOR TYPE=WOR
/PRINT PLAN
/DESIGN STRATA=stratum CLUSTER=psu
/ESTIMATOR TYPE=WR.
* Complex Samples Frequencies.
CSTABULATE
/PLAN FILE='I:\myfile.csaplan'
/TABLES VARIABLES=netusoft
/CELLS TABLEPCT
/STATISTICS SE
/MISSING SCOPE=TABLE CLASSMISSING=EXCLUDE.
* Complex Samples Descriptives.
CSDESCRIPTIVES
/PLAN FILE='I:\myfile.csaplan'
/SUMMARY VARIABLES=netustm
/MEAN
/STATISTICS SE
/MISSING SCOPE=ANALYSIS CLASSMISSING=EXCLUDE.
```

5. Impact of weighting and taking account of sample design

This section demonstrates how omitting weighting or sample design specification in general influences point estimates and their confidence intervals. The examples are based on Round 8 data from Austria, but the results are general and extend to all ESS rounds and countries.

We present estimates of proportion of internet usage frequency for AT in round 8 using 'proportion' command in Stata. Specifically we look at three different scenarios: a correct sample design specification that takes into account clustering, stratification and weighting; unweighted sample specification that takes into account clustering and stratification but not weighting; and no sample design specification which lets Stata assume simple random sample (SRS) design with 100% response rate.

Box 7: Specifying 3 scenarios in Stata

```
svyset psu [pweight=anweight], strata(stratum) // takes into acct full sample design, incl
weighting
svy: proportion netusoft if cntry=="AT"

svyset psu, strata(stratum) // takes into account clustering and stratification, but not weighting
svy: proportion netusoft if cntry=="AT"

proportion netusoft if cntry=="AT" // sample design is not taken into account
```

Table 1. Estimates of proportion of internet usage in AT, ESS round 8, under different sample design specifications

	full sample design Scenario 1			strat and psu, but no weighting Scenario 2			nothing taken into account Scenario 3		
	[95% Conf. Interval]			[95% Conf. Interval]			[95% Conf. Interval]		
	Proportion	Conf.	Interval]	Proportion	Conf.	Interval]	Proportion	Conf.	Interval]
netusoft									
Never	16.6%	14.6%	18.8%	19.5%	17.6%	21.4%	19.5%	17.8%	21.3%
Only									
occasionally	5.5%	4.5%	6.8%	6.4%	5.4%	7.7%	6.4%	5.4%	7.6%
A few times									
a week	8.5%	7.2%	10.0%	9.7%	8.3%	11.2%	9.7%	8.4%	11.0%
Most days	11.5%	9.9%	13.4%	12.0%	10.5%	13.6%	12.0%	10.6%	13.5%
Every day	57.8%	55.0%	60.5%	52.5%	50.0%	54.9%	52.5%	50.3%	54.6%

The results are presented in table 1. In general weighting influences point estimates and standard errors (therefore, confidence interval), while clustering and stratification influences only standard error but not point estimates. We therefore observe that Scenario 2 and Scenario 3 have the same point estimates but Scenario 2 has larger confidence interval than Scenario 3. This slight increase is related to lower efficiency of clustered design in comparison to SRS design (though remember this gives cost efficiency). It can therefore be seen that failing to take into account the sample design results, for clustered sample designs, in under-estimation of the width of confidence intervals.

One can observe that weighting (Scenario 1) changes point estimates in comparison to either of Scenario 2 or 3. Importantly, estimated frequency of internet usage changes with weighting – we estimate fewer of those who never use internet and more of those who use internet each day. This change can be related to either unequal selection probabilities in sample design (where they are present)

or differential nonresponse related to internet usage frequency, or both.

Importantly, the confidence interval of the correct 'Never' proportion estimate (between 14.6% and 18.8%) does not include the wrongly estimated proportion of 19.5% when weighting is not taken into account. Thus, the estimation bias introduced by not taking weighting into account can potentially lead to wrong conclusions.

Weighting (Scenario 1) also increases standard errors which leads to widened confidence interval. This is an important reflection of a less efficient sample design with unequal selection probabilities (though remember cost savings), and increased uncertainty from nonresponse. It is worth mentioning that estimation of correct confidence intervals is also very important, especially in analysis for group or country comparisons. Mis-estimating leads to too narrow confidence intervals which may not overlap where the correct estimation would show an overlap. The analyst may therefore be misled into thinking that a significant difference is present between groups while in fact there is not enough statistical power for such a conclusion (and in fact true population quantities may not be different).

Kaminska and Lynn (2017) provide further examples of mis-specification effects in cross-national survey analysis.

Remember therefore to always take into account full sample design, including stratification, clustering and weighting.

Acknowledgment

I am grateful to Mārtiņš Liberts for providing the examples of R syntax and to Peter Lynn for comments and suggestions.

Reference

Kaminska, O. & Lynn, P. (2017) 'Survey-based cross-country comparisons where countries vary in sample design: issues and solutions'. *Journal of Official Statistics*, 33(1): 123-136. <https://doi.org/10.1515/jos-2017-0007>.